ММТК
МОНГОЛЫН МЭДЭЭЛЛИЙН ТЕХНОЛОГИЙН КОНСЕРЦИУМ

МОНГОЛЫН МЭДЭЭЛЛИЙН ТЕХНОЛОГИ

ММТ-2026

СӨРӨН ТЭСВЭРЛЭХ ЧАДАВХ

МОНГОЛЫН МЭДЭЭЛИЙН ТЕХНОЛОГИЙН 13 ДАХЬ УДААГИЙН ЭРДЭМ ШИНЖИЛГЭЭНИЙ ХУРЛЫН ЭМХЭТГЭЛ

Монгол улс - Чадавхаа сэргээх 2026

УЛААНБААТАР ХОТ
2026

МОНГОЛЫН МЭДЭЭЛЛИЙН ТЕХНОЛОГИЙН 13 ДАХЬ УДААГИЙН ХУРАЛ

ҮНДЭСНИЙ БАТЛАН ХАМГААЛАХЫН ИХ СУРГУУЛЬ

ММТ 2026

СӨРӨН ТЭСВЭРЛЭХ ЧАДАВХ

Монголын Мэдээллийн Технологи 2026
эрдэм шинжилгээний хурлын эмхэтгэл

Улаанбаатар хот
2026

МОНГОЛЫН МЭДЭЭЛЛИЙН ТЕХНОЛОГИ-2026

Монгол улс, Улаанбаатар хот 2026 оны 05 сарын 22-ны өдөр

©Copyright 2026, ММТ

Энэхүү эмхэтгэлийг хуулбарлах, олшруулах, хэвлэхийг
“Зохиогчийн эрхийн тухай хууль”-аар хориглоно.

www.mmt.edu.mn

Хавтасны дизайн: К.Гэрэлчимэг
Хэвлэлийн эх бэлтгэсэн: К.Гэрэлчимэг, Б.Лхагвадорж
Редактор: Б.Лхагвадорж, Д.Идэшноров

“Totopress LLC”-д хэвлэв.

ISBN 978-9919-521-81-3

ЗОХИОН БАЙГУУЛАГЧ

ХАМТРАН ЗОХИОН БАЙГУУЛАГЧ

ОРОЛЦОГЧ БАЙГУУЛЛАГУУД

ХАМРАХ ХҮРЭЭ

1. Сөрөн тэсвэрлэх чадавх
2. Холбоо, сүлжээ
3. Эх хэл боловсруулалт
4. Кибер аюулгүй байдал
5. Өгөгдлийн шинжилгээ
6. Дохио, медиа боловсруулалт, автоматжуулалт
7. Хиймэл оюун
8. Компьютерын хараа

МЭНДЧИЛГЭЭ



Эрхэм хүндэт эрдэмтэн судлаачид, профессор багш нар, салбарын мэргэжилтнүүд та бүхний амар амгаланг айлтан мэндчилье.

Монголын мэдээллийн технологийн консорциумаас санаачлан эхлүүлсэн “Монголын мэдээллийн технологи” эрдэм шинжилгээний цуврал хурал нь мэдээллийн технологийн салбарын судалгаа, инновац, хамтын ажиллагаанд суурилсан үндэсний хэмжээний нэр хүндтэй эрдмийн индэр болон төлөвшжээ.

Энэ удаагийн хурлыг Үндэсний батлан хамгаалахын их сургууль түүхт 105 жилийн ойн хүрээндээ анх удаа зохион байгуулж буй нь бидний хувьд нэр төрийн хэрэг төдийгүй өндөр хариуцлага, цэрэг, иргэний салбарын судалгаа хөгжүүлэлтийн хамтын ажиллагааг шинэ шатанд гаргах том боломж гэж үзэж байна.

Өнөө цагт мэдээллийн технологи, дижитал шилжилт, хиймэл оюун, их өгөгдөл, кибер орчны хурдацтай өөрчлөлт хөгжил нь улс орны нийгэм, эдийн засгийн бүхий л салбарт хүчтэй нөлөөлж, үндэсний аюулгүй байдлын баталгаа, батлан хамгаалах чадавхыг илэрхийлэх гол хүчин зүйл болж байгааг олон улсын цэрэг, улс төрийн байдал харуулж байна. Дэлхийн улс орнууд мэдээллийн технологиор ноёрхох, хамгаалах, сөрөн зогсох чадавхаа бэхжүүлэхийн төлөө бүхий л чармайлтаа төвлөвлөрүүлж, үүнд шинжлэх ухаан, технологийн судалгаа, сургалт чиглэх боллоо.

Манай улс ч “Үндэсний сөрөн тэсвэрлэх чадавхыг бэхжүүлэх стратеги”-тэй болох асуудлыг төр, засгийн хэмжээнд хөндөж, салбар бүрийн төлөөлөл багтсан олон талт судалгаа, шинжилгээний ажил явагдаж байгаа бөгөөд энэ хүрээнд мэдээллийн технологийг хөгжүүлэх, тэр дундаа кибер аюулгүй байдлын асуудлыг чухлаар авч үзэж байна. Ийм ч учраас энэ удаагийн хурлын үндсэн чиглэлийг “сөрөн тэсвэрлэх чадавх”-хэмээн тодорхойлсон бөгөөд шинжлэх ухаан, технологийн хөгжил, тэр дундаа мэдээллийн технологийн ололт, амжилт, түүний онол, практикийг уялдуулсан судалгааны үр дүнгээр шинжлэх ухаанд суурилсан үндэслэл, шинэлэг санаачилгад тулгуурласан бодит санал, шийдлүүдийг боловсруулан, хамтын судалгааны санаачилга өрнүүлж, улс орныхоо аюулгүй байдал, дижитал хөгжил, мэдээллийн орчин дахь сөрөн тэсвэрлэх чадавхийг бэхжүүлэх үйлсэд үнэтэй хувь нэмэр оруулна гэдэгт итгэл төгс байна.

Салбарынхаа хөгжил дэвшил, шинэчлэл өөрчлөлт бүрийн төлөө өөрийн мэдлэг, туршлага, ур чадвар, чин сэтгэл, уйгагүй зүтгэлээ харамгүй зориулан ирсэн эрхэм хүндэт Та бүхэнд талархал илэрхийлье.

“Монголын мэдээллийн технологи-2026” эрдэм шинжилгээний хурлын үйл ажиллагаанд амжилт хүсч, Та бүхний эрдэм судлалын ажил өрнүүн дэлгэрч, шинжлэх ухааны ололт амжилтаар арвин байхыг хүсэн ерөөе!

ХҮНДЭТГЭСЭН

ҮНДЭСНИЙ БАТЛАН ХАМГААЛАХЫН ИХ СУРГУУЛИЙН ЗАХИРАЛ,
БРИГАДЫН ГЕНЕРАЛ Т.БУЯНБАТ

ӨМНӨТГӨЛ

Сайн байцгаана уу?
Эрхэм хүндэт судлаачид аа



Та бүхний амар амгаланг эрэн мэндчилье. Бидний жил бүрийн тэсэн ядан хүлээдэг ММТ - эрдэм шинжилгээний хурал маань хүрэлцэн ирлээ. Энэ жил хурал зохион байгуулагч шинэ тоглогчтой боллоо. Энэ бол Үндэсний батлан хамгаалахын их сургууль юм. ММТК-иас тавьсан хурал зохион байгуулах саналыг халуун дотноор хүлээн авсан их сургуулийн удирдлагад сургуулийнх нь 105 жилийн ойн мэндийг дэвшүүлэн сургалт, судалгааны ажлын өндөр амжилтыг хүсье.

Хурлын бэлтгэлийг хангуулахад ихээхэн хүчин чармайлт гаргасан ҮБХИС-ийн Эрдэм шинжилгээ, инновац, хамтын ажиллагаа хариуцсан дэд захирал Д.Бат-Эрдэнэ, ҮБХИС-ийн Эрдэм шинжилгээ, инновацын хэлтэсийн дарга Д.Идэшноров, ҮБХИС-ийн Эрдэм шинжилгээ, инновацын хэлтэсийн мэргэжилтэн Б.Лхагвадорж нарыг онцлон тэмдэглэмээр байна.

ММТ хурал жил бүр нэг сэдвийг онцолдог уламжлалтай. Энэ жилийн онцлох сэдэв бол “Сөрөн тэсвэрлэх чадавх” юм. Сөрөн тэсвэрлэх чадавх гэдэг зөвхөн цэрэг, батлан хамгаалахын асуудал биш ээ. Дэлхий нийтийн өрсөлдөөнт зах зээлийн орчинд дэвшилтэд технологид суурилсан сөрөн тэсвэрлэх чадавхаа бэхжүүлж, хөгжүүлэх нь нэн чухал юм. Манай судлаачид үүнийг гүнээ ухамсарлан ойлгож, хуралд маш олон тооны бүтээлээ ирүүллээ.

Сүүлийн жилүүдэд ММТ хурлын нэр хүнд өсөж, чанар нь дээшилж байгаа жишгийг энэ жилийн хурал дахин нотолж 135 бүтээл ирсэн. Олон тооны бүтээлээс шилэн сонгож, хуралдаанд оруулна гэдэг хөтөлбөрийн хорооны хувьд амаргүй ажил байлаа. Хөтөлбөрийн хорооны гишүүд үнэхээр сайн ажилласан.

Хөтөлбөрийн хорооноос сонгосон 50 бүтээлийг аман, 53 бүтээлийг үзүүлэн илтгэлийн хэлбэрээр бид өнөөдөр сонсож, хэлэлцэнэ. ММТ-2026 эрдэм шинжилгээний хуралд бүтээлээ ирүүлсэн нийт судлаачдад ММТК -ын нэрийн өмнөөс болон хувиасаа чин сэтгэлийн талархал илэрхийлж, судалгааны ажилдаа шинэ шинэ оргилд хүрэхийг хүсэн ерөөе! Мэдээллийн технологийн хөгжлөөр Монгол орныхоо сөрөн тэсвэрлэх чадавхыг бэхжүүлж, сайн сайхан амьдрал, хөгжлийн манлайд хүргэцээ!

Хурлын ажилд амжилт хүсье.

ХҮНДЭТГЭСЭН,

ДОКТОР (PH.D), ПРОФ.А.ЭРДЭНЭБААТАР
МОНГОЛЫН МЭДЭЭЛЛИЙН ТЕХНОЛОГИЙН КОНСЕРЦИУМЫН
ГҮЙЦЭТГЭХ ЗАХИРАЛ

ММТ 2026 ЭРДЭМ ШИНЖИЛГЭЭНИЙ ХУРАЛ ЗОХИОН БАЙГУУЛАХ БҮРЭЛДЭХҮҮН

ХУРЛЫН ДАРГА

Д.Бат-Эрдэнэ /ҮБХИС-ийн эрдэм шинжилгээ, инновац, хамтын ажиллагаа хариуцсан дэд захирал/

ХӨТӨЛБӨРИЙН ХОРОО

Дарга: Б.Наранчимэг /МУИС, МТЭС, МКУТ-ийн дэд профессор/
Орлогч дарга: Х.Загарзүсэм /ШУТИС, МХТС-ийн ЭНБдарга/
Нарийн бичгийн дарга: Б.Гэрэлмаа /МУИС, МТЭС, ЭХИТ-ийн дэд профессор/

ХӨТӨЛБӨРИЙН ХОРООНЫ ГИШҮҮД

Б.Ганбат /МУИС, МТЭС, захирал/
А.Эрдэнэбаатар /ШУТИС, МХТС, Компьютерын ухааны салбар, зөвлөх профессор/
Б.Лхагвадорж /ҮБХИС, ЭШИХхэлтэс мэргэжилтэн/
К.Гэрэлчимэг /ҮБХИС, ЭШИХэлтэс мэргэжилтэн/
Н.Оюун-Эрдэнэ /МУИС, МТЭС, МКУТ-ийн профессор/
Д.Цэдэвсүрэн /МУБИС, СЦШХ-ийн дарга/
А.Алтангэрэл /ШУТИС, МХТС, Мэдээллийн технологийн тэнхимийн эрхлэгч/
Б.Дорж /ШУТИС, МХТС, Электроникийн тэнхимийн эрхлэгч/
П.Нямсүрэн /ШУТИС, МХТС, Холбооны инженерчлэлийн тэнхимийн эрхлэгч/
А.Хүдэр /ШУТИС, МХТС, Компьютерын ухааны тэнхимийн эрхлэгч/
Б.Мөнхбаяр /ШУТИС, МХТС, Кибер аюулгүй байдлын тэнхимийн эрхлэгч/
П.Жаргалбат /АШУҮИС, ФМЗ-ийн тэнхимийн эрхлэгч/
Ю.Отгонбаатар /Хүрээ МХТДС, Тэнхимийн эрхлэгч/
Б.Баянмөнх /Монголын эрүүл мэндийн мэдээллийн технологийн нийгэмлэг, дарга/
Д.Уганбаатар /ШУА-ийн МТТХ-ийн ЭША, дэд профессор/
Д.Болормаа/СЭЗИС, АТС-ийн профессор/
Б.Сувдаа /МУИС, МТЭС, МКУТ-ийн дэд профессор/
У.Ганбаяр /МУИС, МТЭС, ЭХИТ-ийн ахлах багш/
Р.Мижиддорж /МУБИС, МЗТ-ийн дэд профессор/
Х.Золзаяа /МУИС, МТЭС, ЭХИТ-ийн ахлах багш/
Р.Амартүвшин /МУИС, МТЭС, ЭХИТ-ийн дэд профессор/
С.Баттуяа /ХААИС, ИТС, Ахлах багш/
Ц.Ганбат /Отгонтэнгэр ИС, ХЭЗС, БПХТ-ийн дэд профессор/
М.Мэнд-Амар /МУИС, МТЭС, ХМТ-ийн дэд профессор/
Ц.Жавзмаа /АШУҮИС, Био-Анагаахын сургууль, Физик, мэдээлэл зүйн тэнхимийн дэд профессор

ЗОХИОН БАЙГУУЛАХ ХОРОО

Дарга: Д.Идэшноров /ҮБХИС, ЭШИХэлтсийн дарга/
Нарийн бичгийн дарга: Б.Лхагвадорж /ҮБХИС, ЭШИХ, мэргэжилтэн/

ЗОХИОН БАЙГУУЛАХ ХОРООНЫ ГИШҮҮД

Л.Бямбажаргал /ҮБХИС, ЭШИХэлтэс, ахлах мэргэжилтэн/
Д.Энхчимэг /ҮБХИС, ЭШИХ, мэргэжилтэн/
К.Гэрэлчимэг /ҮБХИС, ЭШИХ, мэргэжилтэн/
Д.Доржмягмар /ҮБХИС, ЭШИХ, мэргэжилтэн/
В.Нарантуяа /ҮБХИС, АБС, ИТ, дэд профессор/
Б.Баярмөнх /ҮБХИС, ЭС, АБСТ, дэд профессор/
Т.Уянга /ҮБХИС, АБС, ИТ, ахлах багш/
О.Энхтуяа /ҮБХИС, АБС, ИТ, багш/
С.Билэгт /ҮБХИС, ЦНС, КАБТ, тэнхимийн дарга/
Э.Мөнхцэцэг /ҮБХИС, ЦНС, КАБТ, ахлах багш/
Ч.Энхбаатар /ҮБХИС, ЦНС, КАБТ, ахлах багш/
О.Цогзолбоо /ҮБХИС, БУТ, багш/
Н.Норовсүрэн /ҮБХИС, БУТ, ахлах багш/
Д.Анхбаяр /ҮБХИС, БУТ, багш/
Э.Ариунзул /ҮБХИС, БУТ, ахлах багш/
Л.Одмаа /ҮБХИС, ЦНС, ЦХТ, ахлах багш, профессор

ММТ-2026 "СӨРӨН ТЭСВЭРЛЭХ ЧАДАВХ"
ЭРДЭМ ШИНЖИЛГЭЭНИЙ ХУРЛЫН ХӨТӨЛБӨР

Цаг	Хөтөлбөр		Байршил
	БҮРТГЭЛ		
08:30-09:00	Хуралд оролцогчдын бүртгэл		*ҮБХИС-ийн хичээлийн I байр, 1 давхарт*
	НЭЭЛТИЙН ҮЙЛ АЖИЛЛАГАА		
09:00-09:10	Хүндэт зочид, хуралд оролцогчдыг танилцуулах		*ҮБХИС-ийн хичээлийн I байр, I давхарт 212 тоот танхим /Сүбээдэй/*
09:10-09:30	Нээлтийн үг: 1. ҮБХИС-ийн захирал Ч.Буянбат 2. ЦХИХХ-ны сайд Ч.Номин 3. БХ-ын сайд Д.Батлут 4. Ерөнхийлөгчийн зөвлөх Ч.Лодойравсал 5. ММТК-ын ерөнхийлөгч А.Эрдэнэбаатар		
09:30-09:35	"ММТ-2026" ЭШХ-ын хөтөлбөр танилцуулах		
	ҮНДСЭН ХУРАЛДААН		
09:35-09:50	Уригдсан илтгэл 1. Д.Ганзориг		*ҮБХИС-ийн хичээлийн I байр, II давхарт 212 тоот танхим /Сүбээдэй/*
09:50-10:05	Уригдсан илтгэл 2. Sachin Dixit		
10:05-10:20	Уригдсан илтгэл 3. Г.Мөнхжаргал		
10:20-10:35	Дурсгалын зураг татуулах		
10:35-11:00	ЦАЙНЫ ЗАВСАРЛАГА		*ҮБХИС-ийн хичээлийн I байр, II давхарт 214 тоот танхим*
	САЛБАР ХУРАЛДААН		
11:00-12:20	**Салбар хуралдаан 1:** СӨРӨН ТЭСВЭРЛЭХ ЧАДАВХ *ҮБХИС-ийн хичээлийн I байр, 212 тоот танхим /Сүбээдэй/*	**Салбар хуралдаан 2:** ХОЛБОО, СҮЛЖЭЭ *ҮБХИС-ийн хичээлийн I байр, 306 тоот танхим /Их-Эе/*	**Салбар хуралдаан 3:** ЭХ ХЭЛ БОЛОВСРУУЛАЛТ *ҮБХИС-ийн төв номын сангийн хурлын танхим /4 давхарт/*
	ҮЗҮҮЛЭН ИЛТГЭЛ		
12:20-13:00	ҮЗҮҮЛЭН ИЛТГЭЛ		*ҮБХИС-ийн хичээлийн I байрны II, III давхарт*
	ӨДРИЙН ХООЛНЫ ЗАВСАРЛАГА		
13:00-14:00	ӨДРИЙН ХООЛ		*ҮБХИС-ийн Сонсогчийн байрны I давхар /2009/*
	ЦЭРГИЙН ЗӨВСӨГ, ТЕХНИКИЙН ҮЗЭСГЭЛЭН		
12:20-14:00	ЦЭРГИЙН ЗӨВСӨГЛӨЛ, ТЕХНИК, ИННОВАЦЫН БҮТЭЭЛИЙН ҮЗЭСГЭЛЭН */Нууцад хамаарах ангиллаас бусад/*		*ҮБХИС-ийн Жагсаалын талбайд*
	САЛБАР ХУРАЛДААН		
14:00-15:20	**Салбар хуралдаан 4:** КИБЕР АЮУЛГҮЙ БАЙДАЛ *ҮБХИС-ийн хичээлийн I байр, 212 тоот танхим /Сүбээдэй/*	**Салбар хуралдаан 5** ӨГӨГДЛИЙН ШИНЖИЛГЭЭ *ҮБХИС-ийн хичээлийн I байр, 306 тоот танхим /Их Эе/*	**Салбар хуралдаан 6:** ДОХИО, МЕДИА БОЛОВСРУУЛАЛТ, АВТОМАТЖУУЛАЛТ *ҮБХИС-ийн төв номын сангийн хурлын танхим /4 давхарт/*
15:20-15:40	ЗАВСАРЛАГА		*ҮБХИС-ийн хичээлийн I байр, II давхарт 214 тоот танхим*
	САЛБАР ХУРАЛДААН		
16:00-17:45	**Салбар хуралдаан 7:** ХИЙМЭЛ ОЮУН *ҮБХИС-ийн төв номын сангийн хурлын танхим /4 давхарт/*		**Салбар хуралдаан 8:** КОМПЬЮТЕРЫН ХАРАА *ҮБХИС-ийн хичээлийн I байр, 306 тоот танхим /Их Эе/*
	1. ХААЛТ, ХҮНДЭТГЭЛИЙН АРГА ХЭМЖЭЭ		
17:15-15:45	НЭГДСЭН ХЭЛЭЛЦҮҮЛЭГ, ХААЛТ		*ҮБХИС-ийн хичээлийн I байр, 212 тоот танхим /Сүбээдэй/*
18:00-20:00	ХҮНДЭТГЭЛИЙН АРГА ХЭМЖЭЭ		Тэмүүжин палас /Офицеруудын ордоны зүүн талд/

ҮНДСЭН БОЛОН САЛБАР ХУРАЛДААНЫ ХӨТӨЛБӨР	
ҮНДСЭН ХУРАЛДААН: ***ҮБХИС-ийн хичээлийн I байр, 212 тоот-Сүбээдэй танхим*** **Хуралдаан даргалагч: Д.Бат-Эрдэнэ /ҮБХИС/**	
09:30-09:50	Сэдэв: Үндэсний сөрөн тэсвэрлэх чадавх *Д.Ганзориг-доктор /Ph.D/, профессор, дэслэгч генерал*
09:50-10:10	Сэдэв: Artificial Intelligence in Cybersecurity *Sachin Dixit, a Cybersecurity Researcher from India, Lt. Colonel*
10:10-10:30	Сэдэв: Илтгэлийн сэдэв: TCubes: Non-Privacy Invasive Dataset for Activities of Daily Living as Thermal Cubes *Г.Мөнхжаргал Дэд профессор, Арабын нэгдсэн эмират улсын их сургууль*
Салбар хуралдаан 1: **СӨРӨН ТЭСВЭРЛЭХ ЧАДАВХ** ***ҮБХИС-ийн хичээлийн I байр, 212 тоот-Сүбээдэй танхим*** **Хуралдаан даргалагч: Б.Өсөхбаяр /МУИС/** **Б.Лхагвадорж /ҮБХИС/**	
11:00-11:10	Зэвсэгт хүчний кибер сөрөн тэсвэрлэх чадвар … *Э.Батзориг нар…*
11:10-11:20	Монгол Улсын онц чухал мэдээллийн дэд бүтцийн кибер тэсвэрлэх чадамжийг үнэлэх нь (IoT-д чиглэсэн таниур халдлагын шинжилгээ) *Erdenesuvd Unenbat, Byambadorj Dondogmegd and Densmaa Batbayar*
11:20-12:30	Хууль бус хуулбарлалтаас сэргийлэх *Bat-Orgil Munkh-Ochir*
11:30-11:40	Квант компьютерын сорилтыг сөрөн тэсвэрлэх хосолсон алгоритмыг ашуулгүй холболт болон өгөгдөл нууцлалд хэрэгжүүлэх судалгаа *Нямдэлгэр Баян-Очир, Мөнхбагана Мөнхтөр, Уянгаа Хүчит, Уганчимэг Ганчимэг, Дашдорж Ямхин, Баяр-Эрдэнэ Лхагвасүрэн*
11:40-11:50	Монгол Улсын дижитал хамаарлын индекс (DDI)-д үндэсний харилцаа холбооны хиймэл дагуулын үзүүлэх нөлөөллийн үнэлгээ *Золбаяр Чулууныцэцэг, Баяр-Эрдэнэ Лхагвасүрэн, Отгонбаяр Батаа, Баярмаа Рагчаа*
11:50-12:00	Өргөтгөсөн бөркэлей пакет фильтр, экспресс өгөгдлийн зам болон машин сургалт ашиглан халдлага илрүүлэх, хамгаалах систем хөгжүүлэх *Chinbold Arvintuya, Dashdorj Ya, Enkhtur Ts and Uyangaa Kh*
12:00-12:20	Салбар хуралдааны хэлэлцүүлэг
Салбар хуралдаан 2: **ХОЛБОО, СҮЛЖЭЭ** ***ҮБХИС-ийн хичээлийн I байр, 306 тоот, Их-Эе танхим*** **Хуралдаан даргалагч: Г.Отгонбаатар /МУИС/** **Ю.Отгонбаатар /Хүрээ ХТДС**	
11:00-11:10	Когнитив радио сүлжээний спектр мэдрэлтэнд анхдагч хэрэглэгчийн дараагийн слотын төлөвийг таамаглах нь *Erdenebayar Lamjav and Otgonbayar Bataa*
11:10-11:20	Харилцаа холбооны санал болгох системийг сайжруулах загвар, түүний туршилт *Solongo Tumurbaatar, Bilguun Dagvadorj and Uyanga Sambuu*
11:20-12:30	5G, WiFi6 сүлжээнд зориулсан давтамжийн дахин тохируулгатай антенны загварчлал *Batjargal Enkhbold, Saruulbuyan Azkhuu, Urangua Bayanbat, Баярсайхан Пүрэвцэрэн and Пүрэвдорж Уганбаяр*
11:30-11:40	FSO холбооны системийн бит алдааны магадлалын үнэлгээ *Ulzii-Orshikh Manjaa, Khurelbaatar Tseveenjav and Munkhbayar Adiya*
11:40-11:50	IPv6 шилжилтийн механизмын гүйцэтгэлийн харьцуулалт: Их сургуулийн кампус сүлжээнд *Enkhbayar Batjargal, Bayasgalan P, Davgasuren Ts, Zolzaya Kh and Battulga Davaasambuu*
11:50-12:00	Чөлөөт орон зайд ОӨМ бүхий дохиог дамжуулахад турбуленсийн нөлөөлөх нөлөөллийн судалгаа *Namuun G and Munkhbayar Adiya*
12:00-12:20	Салбар хуралдааны хэлэлцүүлэг

Салбар хуралдаан 3: **ЭХ ХЭЛ БОЛОВСРУУЛАЛТ** ***УБХИС-ийн төв номын сангийн хурлын танхим /4 давхарт/*** **Хуралдаан даргалагч: Ч.Алтангэрэл /МУИС/** **А.Алтангэрэл /ШУТИС/**	
11:00-11:10	Бага нөөцтэй нутгийн аялгын ярианы синтезийн судалгаа *Золзаяа Бямбадорж, Өлзийбаяр Соном-Очир, Батгэрэл Ариунзаяа*
11:10-11:20	Латин галиглалтаас Кирилл рүү: Нөөц багатай хэлэнд хэлний том загвар ашигласан шийдэл *Өлзийбаяр Соном-Очир, Бямбадорж Золзаяа, Батгэрэл Ариунзаяа*
11:20-12:30	Munkhnote: A Modular AI-Assisted Meeting Minutes System for Mongolian with Improved Workflow Efficiency *Magvan-Erdene Gantulga, Munkhtuya Uurtsaikh and Ganchimeg Ganbold*
11:30-11:40	Дуу хоолой хувиргалтыг сайжруулахад авианы постериограмм (PPG)-д суурилсан аргыг хэрэгжүүлэх нь *Lkhagvatseren Lkhagvasuren, Adiyabat Enkhjargal, Byambajav Dorj and Bayarpurev Mongol*
11:40-11:50	Монгол хэлний ярианаас бичвэр буулгах системийн нууцлалын алгоритмуудын гүйцэтгэлийн жишиг судалгаа *Дашренчин Эрдэнэболд, Эгшиглэн Наранбаатар, Нямжаргал Тэгшжаргал and Enkhtuya Tsogtbaatar*
11:50-12:00	Монголын уртын дууны акустик онцлогт суурилсан хэв маягийн анализ *Uranchimeg N, Amartuvshin Baatarkhuu, Adiyabat Enkhjargal, Oyunerdene O and Chuluuntsetseg Ts*
12:00-12:20	Салбар хуралдааны хэлэлцүүлэг
Салбар хуралдаан 4: **КИБЕР АЮУЛГҮЙ БАЙДАЛ** ***УБХИС-ийн хичээлийн I байр, 212 тоот-Сүбээдэй танхим*** **Хуралдаан даргалагч: М.Золзаяа /ЗХЖШ/** **Я.Дашдорж /ШУТИС/**	
14:00-14:10	Машин сургалтад суурилсан утасгүй сүлжээний халдлага илрүүлэх систем *Шүрэнцэцэг Н, Дашдорж Ямхин*
14:10-14:20	Хиймэл оюун ашигласан автоматжуулсан лог хяналт, шинжилгээний систем *Zolboot Tsoodol, Erhem Otgonbayar, Bilguun E and Munkhbayar Bat-Erdene*
14:20-14:30	Honeypot болон уламжлалт IDS-ийг хослуулсан халдлага илрүүлэх систем хөгжүүлэх *Munkhbayar Bat-Erdene, Dulguun Battur, Dulguun Batchuluun, Nominzul Bayantsatsral and Ijiltsetseg Dawaajargal*
14:30-14:40	IEEE 802.11 сүлжээ бүхий дроны сүргийн системийн auction-based даалгавар хуваарилах механизмын DDOS халдлагын нөлөөллийг симмуляциар үнэлэх нь *Tuvshinbayar Purev, Enkhtur Ts, Dashdorj Ya, Uyangaakh Kh and Erdenebat Ch*
14:40-14:50	Тайлбарлагдахуйц хиймэл оюуныг халдлага илрүүлэх системд хэрэглэх нь *Usukhbayar Baldangombo, Enkhbayar Batjargal and Ugtakhbayar Naidansuren*
14:50-15:00	Хиймэл оюунд суурилсан статик ба динамик шинжилгээг хослуулсан хортой программ илрүүлэх системийн загвар *Oyungerel Sukhbaatar*
15:00-15:20	Салбар хуралдааны хэлэлцүүлэг
Салбар хуралдаан 5: **ӨГӨГДЛИЙН ШИНЖИЛГЭЭ** ***УБХИС-ийн хичээлийн I байр, 306 тоот, Их-Эе танхим*** **Хуралдаан даргалагч: Б.Батзолбоо /ШУТИС/** **Ц.Жавзмаа /АШУҮИС/**	
14:00-14:10	Монгол хэл дээрх залиланги йн шинжтэй сурталчилгааны текстийн ангилалд машин сургалтын загваруудын харьцуулсан судалгаа *Ganchimeg Ganbold and Bayarjargal Byambasuren*
14:10-14:20	Цаг уур, хиймэл дагуулын болон цаг хугацааны өгөгдөлд тулгуурлан pm2.5 концентрацийг машин сургалтын аргаар таамаглах нь *Ganchimeg Ganbold and Khangay Arlungua*
14:20-14:30	Умайн хүзүүний эсийн өөрчлөлтийг бөөм, цитоплазмын сегментаци болон морфологийн тоон шинжид суурилан хиймэл оюуны аргаар ангилах нь *Түмэнзаяа Ганхуяг, Золбоо Төмөрболд, Долгорсүрэн Батжаргал*
14:30-14:40	Харилцагч-борлуулагчийг санал болгох SelfGNN загварын графын шинжийн нөлөөлөл *Altanshagai Erdenechimeg and Amarsanaa Ganbold*
14:40-14:50	Цаг хугацааны цуваа хэлбэрийн мэдрэгчийн өгөгдөл дээрх гажуудал тодорхойлох ба зүгшрүүлэх нь

	Gantumur Nyam and Batzolboo Bataa
14:50-15:00	PISA-2022 өгөгдөлд суурилсан Монгол сурагчдын сурлагын амжилтыг таамаглах загвар ба сургалтын орчин дахь квази-туршилтын судалгаа *Serchmaa Tserenbaljir, Battsetseg Erdenebat and Bayarbat Ulziit*
15:00-15:20	Салбар хуралдааны хэлэлцүүлэг
Салбар хуралдаан 6: **ДОХИО, МЕДИА БОЛОВСРУУЛАЛТ, АВТОМАТЖУУЛАЛТ** ***УБХИС-ийн төв номын сангийн хурлын танхим /4 давхарт/*** **Хуралдаан даргалагч: А.Одгэрэл /ШУТИС/** **С.Баттуяа /ХААИС/**	
14:00-14:10	Дөрвөн сэнст төхөөрөмжийн PD удирдлагын коэффициентуудыг хиймэл оюун ухаан ашиглан оновчлох нь *Namuunaa Dashtseren and Tengis Tserendondog*
14:10-14:20	Speech-Conditioned Image Generation via Direct Projection into a Diffusion Decoder *Magvan-Erdene Gantulga, Zoljargal Munkhnasan and Orgil Jargalsaikhan*
14:20-14:30	Монгол гэрийн хаалганы текстур үүсгэх арга *Orgil Ulzii-Undral and Amarsanaa Ganbold*
14:30-14:40	Assessment of Vibration Energy Harvesting Potential on Mongolian Freight Trains for IoT Applications *Batdalai Sukh, Nyamsuren Purevsuren, Myeirambyek Tashkyen, Battur Ganbat and Ulzii-Orshikh Manjaa*
14:40-14:50	Дээд мөчний нөхөн сэргээх эмчилгээг дэмжих VR тоглоомын хөгжүүлэлт *Oyunbat Otgonbayar and Oyundolgor Khorloo*
14:50-15:00	Modeling adaptive survival through somatic evolution: a unity-based agent simulation of mutation intensity in rabbit populations *Bilguun Baigalmaa*
15:00-15:20	Салбар хуралдааны хэлэлцүүлэг
Салбар хуралдаан 7: **ХИЙМЭЛ ОЮУН** ***УБХИС-ийн төв номын сангийн хурлын танхим /4 давхарт/*** **Хуралдаан даргалагч: С.Цолмон /МУИС/** **Д.Золзаяа /ШУТИС/**	
15:40-15:50	Хиймэл оюун ухаанд суурилсан жолоочийн аюулгүй байдлыг хянах системийн судалгаа ба хөгжүүлэлт *Amartuvshin Bukhbat, Nyam Gantumur and Byambadorj Zolzaya*
15:50-16:00	Өгөөж ба уналтын эрсдэлийн зөрчил: ертөнцийн загварт туйлын эрсдэлийн сохор цэгийг илрүүлэх нь (Return and Drawdown Risk Discrepancy: Detecting Blind Spots in Tail Risk of World Models *Сайханбаяр Х, Пүрэвцогт Н, Золжаргал М*
16:00-16:10	EEG дохионд суурилсан тархины нейропластик шинж чанарын судалгаа ба машин сургалтаар ялган таних нь *Баянбилэг Цэрэн-Ойдов, Тэнгис Цэрэндондог, Долгорсүрэн Батжаргал*
16:10-16:20	MongolDiff: A Masked Discrete Diffusion Language Model for Low-Resource Mongolian Tiny Story Generation *Orgil Jargalsaikhan, Gankhuyag Narantuya and Dolgorsuren Dulamjav*
16:20-16:30	Personal AI Assistant for Cross-Service Information Retrieval and Automation *Batgerel Puntsag*
16:30-16:40	ECG болон PPG, артерийн судасны даралтын дохионоос тархины дотоод даралт ихсэлтийг илрүүлэх машин сургалтын арга *Choijamts Namsraijav and Gerelmaa Byambatsogt*
16:40-16:50	Хиймэл оюун ухаанд суурилсан suvmax тооцооллоор умайн хүзүүний хорт хавдрын үе шат ба үсэрхийллийг үнэлэх нь *Tumen-Ayush Munkhbayar and Ulziibayar Sonom-Ochir*
16:50-17:15	Салбар хуралдааны хэлэлцүүлэг
Салбар хуралдаан 8: **КОМПЬЮТЕРЫН ХАРАА** ***УБХИС-ийн хичээлийн I байр, 306 тоот, Их-Эе танхим*** **Хуралдаан даргалагч: Б.Сувдаа /МУИС/** **Ж.Оргил /ШУТИС/**	
15:40-15:50	Үзэмчин бүжгийн дижитал өгөгдөл бэлтгэл ба хөдөлгөөний кинематик шинж чанарын шинжилгээ *Цэцэнбилэг Дааацэрэн, Cheli Muge, Саранчимэг Хандсүрэн, Оюун-Эрдэнэ Намсрай*

15:50-16:00	Монголын хадны зураг дахь амьтдын дүрслэлийг YOLOv8 загвараар илрүүлэх нь *Perenleilkhundev Gantuya, Rentsendorj Elchinbat and Batsuuri Suvdaa*
16:00-16:10	Компьютерийн хараанд суурилсан хувийн хамгаалах хэрэгслийн хяналтын системийн загварчлал, хэрэгжилт *Sumiyaragchaa Jandalsuren and Ulziibayar Sonom-Ochir*
16:10-16:20	Хүүхдийн аюулгүй байдлыг бодит цаг хугацаанд хянах систем *Оюунжаргал Очир, Азжаргал Сүрэнжав and Золзаяа Бямбадорж*
16:20-16:30	Веб камераар, нүүрний цэгүүдэд суурилан харааны чиглэл таамаглах машин сургалтын загвар сургалт, calibration хийх нь *Sergelen Lkhagvadorj, Altangerel Chagnaa and Bilguun Ochirbat*
16:30-16:40	Гүн сургалтад суурилсан цээжний рентген зургаас эмнэлзүйн тайлан үүсгэх аргын судалгаа *Энхболор Мэндбаяр, Золзаяа Бямбадорж*
16:40-16:50	Эрчимт эмчилгээний тасгийн нярайн амьсгалын давтамжийг хүрэлцэхгүй, тасралтгүй хянах дулааны зураглалд суурилсан системийн хөгжүүлэлт *Javkhlan Altansukh*
16:50-17:15	Салбар хуралдааны хэлэлцүүлэг

	ҮЗҮҮЛЭН ИЛТГЭЛ 1 ***ҮБХИС-ийн хичээлийн I байр, II, III давхар***
1	“Хиймэл оюун болон AR технологид суурилсан анхны тусламжийн системийн судалгаа, хөгжүүлэлт” *Munkhtuya Uurtsaikh and Gankhuyag Narantuya*
2	“Компьютерын харааnд суурилсан объект алдагдлыг илрүүлэх судалгаа *Азжаргал Сүрэнжав, Оюунжаргал Очир, Золзаяа Бямбадорж*
3	Нүүрсний богино ба урт хугацааны үнийн таамаглалыг сайжруулах машин сургалтын алгоритмуудын харьцуулалт *Дорждэрэм Д, Золжаргал М*
4	Эртний Рүни-Монгол бичээсийг орчин цагт “амилуулах” нь: Дүрсэд суурилсан систем *Өлзийбаяр Соном-Очир, Бямбадорж Золзаяа, Болдбаатар Юндэнбат*
5	Comparative Analysis of Building Video Intercom Systems in New Residential Buildings in Mongolia *Bayarmunkh Bulten and Bayarkhuu Purevsuren*
6	CNN ашиглан CAPTCHA системийн эмзэг байдлыг судлах: Энгийн ба нарийн төвөгтэй CAPTCHA-ийн харьцуулсан шинжилгээ *Баянбат Эрдэнэбат and Enkhtuya Tsogtbaatar*
7	Нөөц багатай нөхцөлд монгол-англи холимог бичвэрийн мэдрэмжийн шинжилгээний загвар боловсруулах *Tsolmonbayar Ariunbold and Zolzaya Byambadorj*
8	CNN архитектуруудын олон удаагийн туршилтад суурилсан нарийвчилсан статистик харьцуулалт *Enkhbayar Sundui and Purevtsogt Nugjgar*
9	The Technical Debt of Rapid Digitalization: A Quantitative Analysis of Software Maturity in Mongolia *Youngmi Kim*
10	Рансомвэр төрлийн халдлагаас урьдчилан сэргийлэх илрүүлэх програм хангамж хөгжүүлэх нь *Ганбаяр Ганзориг, Мөнх-Эрдэнэ Зоригтбаатар, Тэмүүлэн, Цэдэнбал, Сугарсүрэн Содномдамба and Мөнхбаяр Бат-Эрдэнэ*
11	Нарны хавтангийн гадаргуу цэвэрлэх алсын удирдлагатай робот системийн судалгаа, хөгжүүлэлт *Anujin Lkhagvasuren, Sumyabazar Enkhbaatar, Zorigtbaatar Baatar and Chuluunbandi Naimannaran*
12	Монгол улсад дронд суурилсан хүргэлтийн системийг хэрэгжүүлэх боломжийн судалгаа *Amarjargal Batzorig and Sonom-Ochir Ulziibayar*
13	Хувийн санхүүг оновчтой болгох AI-д суурилсан ухаалаг туслах мобайл апплик *Bat-Erdene Batbayar*
14	Чичиргээнд суурилсан автомашины хөдөлгүүрийн эвдрэл илрүүлэх: domain adaptation аргаар загварыг шинэ орчинд дасан зохицуулах *Sarantuya Zorigtbaatar and Naranbat Darkhanbaatar*
15	Үүлэн орчин дахь ашуулгүй байдлын удирдлагыг автоматжуулах нь *Erdenebayar Buyandelger, Sergelen Tumurbaatar, Bilguun, Dorjderem and Munkhbayar Bat-Erdene*
16	Гүн сургалтын аргаар ct зургаас гавал тархины гэмтлийг илрүүлэх нь *Sarangerel Dorjgochoo*
17	Хиймэл оюунд суурилсан имэйлийн фишинг ангиллын систем хөгжүүлэх *Сугархүү Ган-Эрдэнэ, Төгөлдөр Батнасан, Жаргал Батболд and Мөнхбаяр Бат-Эрдэнэ*
18	Бэхжүүлэх сургалтад суурилсан халдлага илрүүлэх систем хөгжүүлэх *Lhagwa-Ochir Chingaral*
19	Нам өндрийн агаарын хөдөлгөөний удирдлагын системийн өгөгдлийн сангийн загварчлал *Дүүриймаа Гансүх, Наранчимэг Болд*
20	Нисгэгчгүй нисэх төхөөрөмжийн оптик холбооны судалгаа *Dolgorsuren Dulamjav, Ganbold Shagdar and Buyankhishig Zundui*
21	Өсвөр насныхны сэтгэцийн эрүүл мэнд ба асран хамгаалагчийн хүн ам зүйн байдлын хоорондын хамаарлын загварчлал *Лувсан-Иш Ажнай, Цагаан Баатархүү, Цэнд Жавзмаа*, Лувсанцэрэн Пүрэвдолгор, Ахыт Тилеубай, Оюунбилэг Бат-Энх, Болдбаатар Одгэрэл, Энхжаргал Энх-Үрэл, Чилхаасүрэн Баасандорж*

22	Машин сургалт ашиглан макро эдийн засгийн эрсдэлийн индекс боловсруулах: Монгол Улсын инфляцийг урьдчилан таамаглах судалгаа *Solongoo Tsend-Ayush and Battsetseg Erdenebat*
23	Машин сургалтын аргад суурилсан робот сургалтын төхөөрөмжийн хөгжүүлэлт *Khurelbaatar Tseveenjav, Uuganchimeg Ganchimeg and Munkhbayar Adiya*
24	Хөдөлмөрийн аюулгүй байдлын VR сургалтын платформ *Namuun Bayarsaikhan and Oyundolgor Khorloo*
25	Монгол гэрийн 3D модель үүсгэх арга *Orgil Ulzii-Undral and Amarsanaa Ganbold*
26	Монгол адууны тамганы нийлмэл дүрсийг задлан танихад зориулсан бүтцэд суурилсан дүрмийн арга *Gantuya Perenleilkhundev and Suvdaa Batsuuri*
27	Машин сургалтад суурилсан тосны шинжилгээ ба зогсолтын өгөгдлөөр тоног төхөөрөмжийн эвдрэлийг таамаглах нь *Ganchimeg Ganbold and Jargakh Bolortungalag*
28	Эрүүл мэндийн мэдээлэл зүйн хөтөлбөрийн төгсөгчдийн сэтгэл ханамжийн шинжилгээ ба машин сургалтын загварууд *Odgerel Boldbaatar, Baatarkhuu Tsagaan, Purevdolger Luvsantseren, Javzmaa Tsend, Baasandorj Chilkhaasuren, Ajnai Luvsan and Akhit Tilyebai*
29	Фишинг халдлагыг илрүүлэх аргуудын харьцуулсан судалгаа *Oyuntsetseg Luvsandash, Amartaivan Jigjid, Enkhtsetseg Luvsannyam*
30	Монгол хэлний үсгийн авиа танилтад машин ба гүн сургалтын загваруудын гүйцэтгэлийг харьцуулсан нь / *Чигцалмаа Нянтайжанчив, Солонгоо Баярболд, Нарансолонго Буяндэлгэр, Энхтуяа Цогтбаатар*
31	“Aqua Leaf-Ai” Хиймэл оюунд суурилсан тасалгааны ургамал усалгааны системийн судалгаа, хөгжүүлэлт *Duurenbayar Baatarkhuu and Davaasuren Nyamaa*
32	Байгууллагын ажилтны зан төлөвт суурилсан хиймэл оюун болон технологийн шийдлийг турших нь *Иса Нугман, Хажмухан Жагаслаг, Хүрэлбаатар Батжаргал, Мөнхбаяр Бат-Эрдэнэ*
33	Speech Emotion Recognition via Attentive Pooling on Whisper Encoder Representations *Magvan-Erdene Gantulga, Zoljargal Munkhnasan and Orgil Jargalsaikhan*
34	LBPH болон Haar Cascade алгоритмд суурилсан Монгол хүний царай таних систем *Baatarkhuu Tsagaan, Javzmaa Tsend, Bat-Enkh Oyunbileg and Battugs Oyunbileg*
35	Машин орчуулга ашиглан Монгол хэлний өгүүлбэрийн утгыг жиших загвар сургах нь *Munkh-Orshikh Tuvshinbaatar*
36	Дөрвөн сэнстэй дэлгэгддэг дроны загвар болон хөгжүүлэлт *Bat-Orgil Erdenetsogt and Munkhtulga Byambaa*
37	Өөрийгөө хөгжүүлэх ухаалаг систем: Хиймэл оюуны шинжилгээ ба мэргэжилтний зөвлөгөөний хосолсон загвар *Sarangerel Dorjgochoo and Usukh-Ireedui Batkhuu*
38	Халдлага илрүүлэх системд ашиглах санг харьцуулан судлах нь *Ugtakhbayar Naidansuren, Bayasgalan Purevdorj, Byambajav Dorj and Usukhbayar Baldangombo*
39	YOLO-д суурилсан хог хаягдлын автомат ангиллын аргын судалгаа *Khulan Dashkhorloo and Uuganbayr Ganbold*
40	IoT төхөөрөмжүүдийн сүлжээний дүрэмд суурилсан халдлага илрүүлэх систем боловсруулах *Буянцог Пүрэвдорж, Алтансүх Тэмүүлэн, Бат-Эрдэнэ Мөнхбаяр*
41	Зургаас автомашины модел өнгөний ангилал *Darmaasuren Khaidav and Ganbold Uuganbayar*
42	Захын тооцооллыг ашиглан зүйлсийн интернэтийн өгөгдөл дамжуулалтын ачааллыг бууруулах *Баяраа Эрдэнэпүрэв, Отгонбаяр Баярмаа, Нороврenчин Сарангоо, Золбаяр Тэмүүлэн, Бат Эрдэнэ Мөнхбаяр*
43	Бодит цагийн халдлагын дүрслэлтэй үүлэн орчинд суурилсан honeypot систем хөгжүүлэх нь *Даваадэмбэрэл Пүрэвдорж, Дашням Батсайхан, Батзориг Энхболд, Мөнхбаяр Бат-Эрдэнэ*

44	Лимбэний дуугаралтын цахилгаан дохионы судалгаа *Baatarkhuu Tsagaan, Javzmaa Tsend, Battugs Oyunbileg and Bat-Enkh Oyunbileg*
45	Нисэх буудлын хөөрч, буух зурваст гаднын биетийг илрүүлэх mmWave ба X-band радарын системийн судалгаа *Нарантуяа Эрхэмбаатар, Сэлэнгэсүрэн Батмөнх*
46	Comparative Evaluation of Modern OCR Pipelines for Mongolian Banking Document Digitization *Munkhtsatsral Tumurbat*
47	Радарын хөндлөн огтлолын коэффициентыг бууруулах мета-гадаргуугийн загварчлал *Anudari Dashbaatar, Undarmaa Battulga, Sanjsuren Dashgyantsan, Purevtseren Bayarsaikhan and Uuganbayar Purevdorj*
48	Монгол үндэсний хээ угалзын зарим дүрсийг математик загварчлалаар вектор график болгон зурах *Bagabold Gendensuren, Gantuya Perenleilhundev and Byambasuren Ivanov*
49	5G - ийн дээд зурваст ажиллах тархинаас мэдрэлийн өгөгдлийг өндөр хурдаар дамжуулах суулгац антены загварчлал *Zagarsuren Batbaatar, Oyun-Erdene Bolormaa, Пүрэвцэрэн Баярсайхан, Уганбаяр Пүрэвдорж*
50	ЕБС-ийн багш, сурагчдын кибер аюулгүй байдлын мэдлэгийг тодорхойлох судалгаа *Ундрал Б*
51	AI red teaming аргачлалд суурилсан хиймэл оюуны суурь загварын эмзэг байдлын судалгаа *Д.Сарангэрэл*
52	Олон үүлэн орчны хандалтын бодлого дахь халдлагын замын графыг загварчлах нь *Yanjinlkham Baasankhuu, Sanchirmaa Erdenejargal and Munkhbayar Bat-Erdene*
53	Өсвөр насныхны сэтгэцийн эрүүл мэнд ба асран хамгаалагчийн хүн ам зүйн байдлын хоорондын хамаарлын загварчлал *Baatarkhuu Tsagaan, Ajnai Luvsan-Ish, Javzmaa Tsend, Purevdolgor Luvsantseren, Baasandorj Chilkhaasuren, Akhyt Tilyubai, Odgerel Boldbaatar, Enkh-Urel Enkhjargal and Oyunbileg Batenkh*

ҮНДСЭН ХУРАЛДААН

УРИЛГЫН ИЛТГЭЛ

ҮНДСЭН БОЛОН САЛБАР ХУРАЛДААНЫ ХӨТӨЛБӨР	
ҮНДСЭН ХУРАЛДААН: ***ҮБХИС-ийн хичээлийн I байр, 212 тоот-Сүбээдэй танхим*** **Хуралдаан даргалагч: Д.Бат-Эрдэнэ /ҮБХИС/**	
09:30-09:50	Сэдэв: Үндэсний сөрөн тэсвэрлэх чадавх *Д.Ганзориг, Доктор /Ph.D/, профессор, дэслэгч генерал*
09:50-10:10	Сэдэв: Convergence of AI and Cybersecurity: Emerging Threats, Defenses, and the Road to Resilient Systems *Sachin Dixit, a Cybersecurity Researcher from India, Lt. Colonel*
10:10-10:30	Сэдэв: TCubes: Non-Privacy Invasive Dataset for Activities of Daily Living as Thermal Cubes *Мөнхжаргал Гочоо. Доктор /Ph.D/, дэд профессор, Арабын нэгдсэн эмират улсын их сургууль*



Үндэсний сөрөн тэсвэрлэх чадавх

Д.Ганзориг

Доктор /Ph.D/, профессор,

дэслэгч генерал

Хураангуй

Үндэсний сөрөн тэсвэрлэх чадавхыг бэхжүүлэх үйл ажиллагааны зорилго нь Монгол Улсын хөгжил, Үндэсний аюулгүй байдлыг нэгдсэн бодлого, стратегийн хүрээнд уялдуулах; иргэн бүрийн оролцоо, эх оронч ухамсарт тулгуурласан үндэсний чадавхыг төлөвшүүлэх; орон нутгийн хөгжлийг төв–орон нутгийн нэгдмэл тогтолцоонд уялдуулахад чиглэнэ. Монгол Улсын Засгийн газрын 2024–2028 оны үйл ажиллагааны хөтөлбөрт “Үндэсний сөрөн тэсвэрлэх чадавх ба амар тайван нийгэм” гэсэн бүлэг тусгагдсан нь энэ ойлголтыг төрийн бодлогын түвшинд институцчилсэн чухал алхам гэж үзэж байна.

Товч танилцуулга

Монгол овогтой Довчинсүрэнгийн Ганзориг нь 1971 онд Дархан хотод төрсөн. Орос, англи хэлний мэдлэгтэй.

Боловсрол:

- 1980-1990 он, Зөвлөлтийн 8 дугаар дунд сургууль, Улаанбаатар хот, Монгол Улс;
- 1990-1995 он, Хими хамгаалалтын цэргийн дээд сургууль, командын офицер, Тамбов хот, ОХУ, Техникийн ухааны магистр;
- 1999-2000 он, Тусгай хүчний сургууль, командын офицер, Ши Жи Жуан, Хэбэй муж, БНХАУ, диплом;
- 2001 он, Явган цэргийн офицерын үндсэн дамжаа, Форт-Беннинг, Жеорж муж, АНУ, диплом;
- 2002 он, Рейнжерийн дамжаа, Форт-Беннинг, Жеорж муж, АНУ, диплом;
- 2007-2008 он, АНУ-ын Тэнгисийн явган цэргийн Команд-штабын коллеж, Квантико, Виржиния муж, АНУ, диплом, Цэргийн судлалын магистр;
- 2011-2012 он, АНУ-ын Армийн Дайны коллеж, Карл Айл, Пенсильвани муж, диплом, Стратеги судлалын магистр;
- 2013-2025 он, ҮБХИС, докторантур, Аюулгүй байдлын судлалын чиглэлээр доктор (PhD).

Сургалт дамжаа:

- 2000 он, Олон улсын энхийг сахиулахын чадавхын сургагч багшийн ахисан шатны дамжаа, Хонолулу, Хавай, АНУ;

- 2002 он, Ахлагчийн үүрэг, АНУ-ын Олон улсын эрх зүй судлалын батлан хамгаалах хүрээлэн, Улаанбаатар хот, Монгол Улс;
- 2008 он, Олон улсын батлан хамгаалах нөөцийн менежментийн дамжаа, Оттава, Канад;
- 2009 он, Цогц энхийн ажиллагааны удирдах бүрэлдэхүүний менежментийн дамжаа, Корнвалис Парк, Нова Скошиа, Канад;
- 2013 он, Ардчилсан нийгэмлэгийн Зэвсэгт хүчин семинар, Улаанбаатар хот, Монгол Улс;
- 2014 он, Номхон далайн бүс нутгийн ажлын хэсгийн уулзалт, Улаанбаатар хот, Монгол Улс.

Ажил эрхлэлт:

- 1995-1997 он, ЗХ-ний 119 дүгээр анги, салааны захирагч;
- 1997-2003 он, ЗХ-ний 150 дугаар анги, албаны дарга, сургагч-багш;
- 2003-2007 он, ЗХЖШ, мэргэжилтэн;
- 2008-2010 он, ЗХ-ний 311 дүгээр анги, захирагчийн орлогч, төвийн дарга;
- 2010-2011 он, ҮАБЗ-ийн Ажлын алба, референт;
- 2012-2015 он, ЗХЖШ, Стратегийн Бодлого-төлөвлөлтийн газар, орлогч дарга, хэлтсийн дарга;
- 2015-2019 он, ЗХЖШ, Ажиллагааны удирдлагын газрын дарга;
- 2019-2022 он, ЗХЖШ-ын дарга;
- 2022-2023 он, ҮАБЗ, шинжээч;
- 2023-оос Сайнс Ренж ХХК, гүйцэтгэх захирал;
- 2026 – БХЭШХ-нд эрдэм шинжилгээний ажилтан;
- 2026 – КАБЗ-ийн Ажлын албаны дарга;
- 2026 - Сайнс Ренж ХХК.



TCubes: Non-Privacy Invasive Dataset for Activities of Daily Living as Thermal Cubes using multi-LSTM

Доктор (PhD), Гочоо овогтой Мөнхжаргал
Дэд профессор,
Арабын нэгдсэн эмират улсын их сургууль

Хураангуй

This work introduces the first open multi–infrared (IR) thermal array sensor dataset for recognizing 19 activities of daily living (ADLs) performed by 74 subjects. Each activity sample forms a 32×32×32 thermal tensor cube, hence the name TCubes. The dataset surpasses existing resources in scale, featuring the highest number of activities, twice as many sensors, and three times as many participants as comparable datasets. Thermal patterns were captured using nine low-resolution (8×8) IR thermal sensors, providing a non-invasive and privacy-preserving means of activity recognition. A comprehensive benchmarking study evaluates both convolutional and transformer-based architectures—including C3D, R (2+1) D-18, MViTv2, and Swin-T—to assess their ability to learn spatiotemporal representations from coarse thermal imagery. Results are highly promising: R (2+1) D-18 achieves the most consistent performance with an F1-score of 0.900, while transformer models such as MViTv2 and Swin-T effectively capture subtle gestures and generalize well across 18 and 19 activity classes. The lightweight 3DCNN-Mixed model further demonstrates strong efficiency for resource-constrained applications, highlighting the trade-off between accuracy and computational cost. Analyses leveraging entropy, mutual confusion, and frequency-domain representations reveal how factors such as activity location, posture, temporal dynamics, and motion periodicity influence recognition accuracy. Overall, this dataset and benchmarking suite establish a robust foundation for future research in low-resolution, low-compute, non-invasive, and privacy-preserving human activity recognition, with broad implications for eldercare, healthcare monitoring, and smart environments

Товч танилцуулга

Доктор Г.Мөнхжаргал нь 2019 оноос хойш Арабын нэгдсэн эмират улсын их сургуулийн Мэдээллийн технологийн коллежийн Компьютерийн шинжлэх ухаан, программ хангамжийн инженерчлэлийн тэнхимд дэд профессороор ажиллаж байна. Тэрээр "Embedded AI" нэгжийг удирдаж, "AI and Robotics" лабораторт үндсэн судлаачаар ажилладаг. Түүний судалгааны үндсэн чиглэл нь гүн сургалт, хиймэл оюун, компьютерийн хараа, робот техник болон хувь хүний нууцлалыг хадгалсан үйлдэл таних технологи юм. Түүнчлэн доктор М. Гочоо цар тахлын үеэр тус их сургуулийн дэмжлэгтэйгээр "MetaTouch" гарааны бизнесийг хамтран байгуулж, COVID-19-ийн тархалтыг бууруулах зорилгоор лифтний зайнаас мэдрэгчтэй товчлуурын технологийг нэвтрүүлсэн нь Абу Даби Эмиратын олон зуун барилгад ашиглагдсан бөгөөд 2026 оны 5 сарын 15-ний өдрийн байдлаар түүний нийт ишлэл 4446, h-индекс 36 байна.



Convergence of AI and Cybersecurity: Emerging Threats, Defenses, and the Road to Resilient Systems

Sachin Dixit,

a Cybersecurity Researcher from India,

Lt. Colonel

Abstract

Artificial intelligence and cybersecurity, long treated as adjacent disciplines, have converged into a single strategic terrain. In this terrain, AI is simultaneously a weapon used by attackers, a target whose models and training data must be defended, and a shield deployed by defenders inside security operations centers. This invited lecture surveys this convergence and argues that classical security models — perimeter, signature, human-paced response — are no longer sufficient for the AI-saturated enterprise.

We begin by characterizing the new attack surface, in which probabilistic systems with emergent behavior, machine identities outnumbering human ones, and natural-language inputs that become executable instructions, jointly invalidate longstanding assumptions. We then survey the offensive landscape using the NIST AI 100-2 taxonomy — poisoning, evasion, privacy, and abuse — and devote particular attention to prompt injection, memory poisoning, and the emerging risks of autonomous AI agents that act through tool-using protocols. Empirical evidence from 2025 and 2026 is presented to demonstrate that an AI-enabled attack is no longer prospective but operational.

We next examine where defensive AI genuinely changes the defender's economics, distinguishing measurable gains in alert triage, anomaly detection, and automated response from areas where current capabilities remain limited. A four-layer reference architecture is presented, anchored on zero-trust identity, trusted AI pipelines, AI-augmented operations, and governance aligned with NIST, OWASP, ISO/IEC, and the EU AI Act.

Finally, we frame resilience — not defense — as the appropriate north star, articulating five principles for systems that absorb, adapt, and continue functioning under sustained adversarial pressure. The lecture concludes by identifying open research problems in certified robustness, provenance, agent security semantics, and cross-border AI norms, with an invitation to the Mongolian research community to contribute to this international agenda.

Keywords — Artificial intelligence; cybersecurity; adversarial machine learning; prompt injection; agentic AI; resilience; AI governance; NIST AI 100-2

САЛБАР ХУРАЛДААН - I

Сөрөн тэсвэрлэх чадавх

№	**ҮБХИС-ийн хичээлийн I байр, 212 тоот-Сүбээдэй танхим** ***Хуралдаан даргалагч:*** *Б.Өсөхбаяр /МУИС/, Х.Баасандорж /ҮБХИС/*
1	“Зэвсэгт хүчний кибер сөрөн тэсвэрлэх чадвар” *Э.Батзориг нар*
2	“Монгол Улсын онц чухал мэдээллийн дэд бүтцийн кибер тэсвэрлэх чадамжийг үнэлэх нь (IoT-д чиглэсэн таниур халдлагын шинжилгээ)” *Erdenesuvd Unenbat, Byambadorj Dondogmegd and Densmaa Batbayar*
3	“Динамик DRM Систем Ашиглан Видео Контентийн Хууль Бус Хуулбарлалтаас Сэргийлэх” *Bat-Orgil Munkh-Ochir, Suvd Soninbbayar, Arvin Tamir, Munkhbayar Bat-Erdene*
4	“Квант компьютерын сорилтыг сөрөн тэсвэрлэх хосолсон алгоритмыг аюулгүй холболт болон өгөгдөл нууцлалд хэрэгжүүлэх судалгаа” *Нямдэлгэр Баян-Очир, Мөнхбагана Мөнхтөр, Уянгаа Хүчит, Уганчимэг Ганчимэг, Дашдорж Ямхин, Баяр-Эрдэнэ Лхагвасүрэн*
5	“Монгол Улсын дижитал хамаарлын индекс (DDI)-д үндэсний харилцаа холбооны хиймэл дагуулын үзүүлэх нөлөөллийн үнэлгээ” *Золбаяр Чулуунцэцэг, Баяр-Эрдэнэ Лхагвасүрэн, Отгонбаяр Батаа, Баярмаа Рагчаа*
6	“Өргөтгөсөн бөркэлей пакет фильтр, экспресс өгөгдлийн зам болон машин сургалт ашиглан халдлага илрүүлэх, хамгаалах систем хөгжүүлэх” *Chinbold Arvintuya, Dashdorj Ya, Enkhtur Ts and Uyangaa Kh*

“ХИЙМЭЛ ОЮУН УХААН, ӨНДӨР ТЕХНОЛОГИЙН ШИЛЖИЛТ” БАТЛАН ХАМГААЛАХ САЛБАРТ

Батлан хамгаалахын “Хиймэл оюун ухаан, өндөр технологийн шилжилт” стратеги төлөвлөгөө боловсруулах ажлын хэсгийн судлаачид
ҮБХИС-ийн Эрдэм шинжилгээ, инновацын хэлтэс
rid@mndu.gov.mn

Хураангуй—Хиймэл оюун ухаан, өндөр технологийг улс орны бүхий л салбарт нэвтрүүлж, үр дүнгээр нийгэм эдийн засгийн хөгжилд хувь нэмэр оруулах үндсэн чиглэлтэй Монгол Улсын ерөнхий сайдын 2024 оны 10 дугаар сарын 30-ны өдрийн “Алсын хараа 2050 урт хугацааны хөгжлийн бодлого, шинэ сэргэлтийн бодлогын хэрэгжилтийг эрчимжүүлэх” албан даалгаварыг хэрэгжүүлэх ажлын хүрээнд “Батлан хамгаалахын хиймэл оюун ухаан, өндөр технологийн шилжилт” стратеги төлөвлөгөөний төслийг ҮБХИС-ийн ажлын хэсгийг түшиглэн боловсруулж, хэрэгжүүлэх арга хэмжээг тодорхойлов.

Түлхүүр үг—Хиймэл оюун ухаан, өндөр технологи, туршилт, судалгаа, шилжилт

Оршил: 2024 оны аравдугаар сарын байдлаар дэлхийн 87 улс хиймэл оюун ухаанд чиглэсэн үндэсний хэмжээний стратеги төлөвлөгөөг баталж, хиймэл оюун ухаанд суурилсан бүтээгдэхүүн, үйлчилгээг нэмэгдүүлэх чиг хандлагыг баримталж байна[1].

Дижитал өөрчлөлттэй холбоотой асуудлаар байгууллага, Засгийн газарт зөвлөгөө өгдөг “Oxford insights” нэртэй компанийн хийсэн “Дэлхийн 160 орны хиймэл оюун ухааныг төрийн үйлчилгээнд ашиглахад бэлэн байгаа эсэхийг эрэмблэсэн судалгаа”-аар[2], БНХАУ, БНЭУ тэргүүлж, АНУ, Австрали, Их Британи, Канад, Сингапур Улс удаалж, харин хиймэл оюун ухааныг төрийн үйлчилгээнд ашиглахад хамгийн бага оноотой бүс нутагт Сахарын цөлөөс өмнө зүгт орших Африк, Карибын тэнгис, Латин Америк зэрэг хөгжиж буй орнууд орсон байна. Түүнчлэн Төв болон өмнөд Азийн зарим орнууд багтжээ.

Хүн төрөлхтөн хиймэл оюун ухаан (artificial intelligence-AI)-ыг ашиглаж эхэлснээр дэлхийн эдийн засагт томоохон өөрчлөлт гарч байна. Голдман Сакс (Goldman Sachs) компанийн хийсэн судалгаагаар хиймэл оюун ухааны тусламжтайгаар дэлхийн ДНБ-ийг 7 хувь буюу 7 орчим их наяд ам.долларoор өсгөж, бүтээмжийг 10 жилийн хугацаанд 1.5 хувиар нэмэгдүүлэх боломжтой гэж тодорхойлсон байна. Мөн тус компанийн мэдээллээснээр 2030 он гэхэд дэлхийн бизнес эрхлэгч нарын 70-аас доошгүй хувь нь ядаж нэг төрлийн хиймэл оюун ухааны технологийг нэвтрүүлнэ гэж тодорхойлсон нь энэ технологийн эрчимтэй өсөлтийг илэрхийлж байна.

Стратеги төлөвлөгөөний төслийг боловсруулахдаа бодлогын дараах баримт бичгүүдийг эрх зүйн үндэслэл болгов[3].

Хиймэл оюун ухаан, өндөр технологийн шилжилтийг батлан хамгаалах салбар, Зэвсэгт хүчинд хэрэгжүүлэх стратеги төлөвлөгөөний ерөнхий чиглэлийг 3 зорилтын хүрээнд авч үзэн 50 ажил, арга хэмжээг хэрэгжүүлэхээр төлөвлөв.

[1] https://theconversation.com/developing-countries-are-being-left-behind-in-the-ai-race-and-thats-a-problem-for-all-of-us-180218

[2] https://www.oxfordinsights.com/government-ai-readiness-index2021

[3] “Алсын хараа-2050” Монгол Улсын урт хугацааны хөгжлийн бодлого,

- “Монгол Улсын батлан хамгаалах бодлогын үндэс”
- “Цахим үндэстэн” баримтлах чиглэл
- “Батлан хамгаалахын хүний нөөцийн бодлого”
- “Батлан хамгаалахын стратеги төлөвлөгөө”
- “Батлан хамгаалах яамны стратеги төлөвлөгөө”
- “Үндэсний батлан хамгаалахын их сургуулийн хөгжлийн стратеги”
- Засгийн газрын 2023 оны 07 дугаар сарын 05-ны өдрийн 249 дүгээр тогтоолоор батлагдсан “Шинжлэх ухаан, технологийн тэргүүлэх чиглэл, цөм технологийн жагсаалт” (2023-2026)
- Засгийн газрын 2020 оны 95 дугаар тогтоолоор батлагдсан “Инновацийн үйл ажиллагааны тэргүүлэх чиглэл”
- “Батлан хамгаалах судлал, цэргийн шинжлэх ухаан, технологи, инновацыг 2030 он хүртэл хөгжүүлэх бодлого”
- “Зэвсэгт хүчний холбоо, мэдээллийн техник, технологийг шинэчлэх бодлого” болон бусад төлөвлөгөө, эрх зүйн баримт бичгүүд

Мөн хэрэгжүүлэх арга хэмжээний үндсэн чиглэл, дүгнэлт, саналыг боловсруулав.

Стратеги төлөвлөгөөний үндсэн чиглэл

Батлан хамгаалах салбарт хиймэл оюун ухаан, өндөр технологийн шилжилтийн бодлого, төлөвлөлтийг хэрэгжүүлэх

Эрх зүйн орчин бүрдүүлэх зорилтын хүрээнд:

- Хиймэл оюун ухаан, өндөр технологийг цэргийн хэрэгт нэвтрүүлэх ажлыг хэрэгжүүлэх, судалгааны үр дүнг үйлдвэрлэлд нэвтрүүлэх эрх зүйн орчинг бүрдүүлэх;
- Батлан хамгаалахын өндөр технологи, хиймэл оюун ухааныг хөгжүүлэх чиглэлээр судалгаа явуулах боломжийг бүрдүүлэх;
- Зэвсэгт хүчний нэгтгэл, анги, байгууллагад сургалт-судалгааны орчинг бий болгож, цэргийн алба хаагчдыг судалгааны ажилд оролцох боломж, хүрээг нэмэгдүүлэх
- Өндөр технологийн хэрэгцээт импортын хүрээ, далайцыг тодорхойлж, гадаад орнуудаас худалдан авалт хийх эрх зүйн орчинг бүрдүүлэх;

Өндөр технологийг нутагшуулах, хиймэл оюун ухааныг хөгжүүлэх зорилтын хүрээнд:

- Цэргийн болон цэрэг-техникийн хамтын ажиллагаатай орнуудаас батлан хамгаалахын дэвшилтэт технологийг нутагшуулах;
- Гадаад орнуудын зэвсэгт хүчний хиймэл оюун, өндөр технологийн хэрэглээг судалж, зэвсэглэлд нэвтрүүлэх техник, хэрэгслийг зохион бүтээх, үйлдвэрлэх, сайжруулах, технологи нутагшуулах талаар судалгаа, туршилт явуулж, инновацыг тасралтгүй хөгжүүлэх;
- Батлан хамгаалах судлал, цэргийн шинжлэх ухаанд их өгөгдлийн судалгааг хөгжүүлэх;
- Батлан хамгаалах судлал, цэргийн шинжлэх ухаан, технологийн судалгаа-инновац-туршилт-үйлдвэрлэлийн тогтолцоог хөгжүүлэх;
- Цэргийн хамтын ажиллагааны хүрээнд гадаад орны хиймэл оюун ухаан хөгжүүлэлтийн туршлагыг судалж, үр дүнг Зэвсэгт хүчинд нэвтрүүлэх.

Инновацын үр ашигтай тогтолцоог бий болгох зорилтын хүрээнд:

- Хиймэл оюун ухаан, өндөр технологийн туршилт судалгааны төвийг Үндэсний батлан хамгаалахын их сургуулийг түшиглэн байгуулж, түүнийг дэс дараатай бэхжүүлэх, тоног төхөөрөмжөөр хангах зорилго бүхий төсөл, арга хэмжээг төлөвлөн хэрэгжүүлэх;
- Хиймэл оюун ухаан, өндөр технологийн зарим төрлийн бүтээгдэхүүнийг дотооддоо үйлдвэрлэх;
- Батлан хамгаалахын захиалгат ханган нийлүүлэлтэд дэмжлэг үзүүлэх үйлдвэрлэл, үйлчилгээний бааз суурийг бүрдүүлэх;

Батлан хамгаалах салбарт хиймэл оюун ухаан өндөр технологийг хэрэглэх бодлого, стратегийг тодорхойлж, хууль эрх зүй таатай орчин бүрдүүлэх, хэрэгжилтийг хангах зорилтын хүрээнд:

1. Батлан хамгаалах салбар, Зэвсэгт хүчинд хиймэл оюун ухааныг хэрэглэхтэй холбоотой хууль, эрх зүйн орчныг бүрдүүлэх;
2. Зэвсэгт хүчинд хиймэл оюун ухааныг хэрэглэх үндсэн чиглэлийг тодорхойлох;
3. Өндөр технологийн хэрэгцээт импортын хүрээ, далайцыг тодорхойлж, гадаад орнуудаас худалдан авалт хийх эрх зүйн орчинг бүрдүүлэх;
4. Өндөр технологи, хиймэл оюун ухааны хөгжүүлэлтийн судалгаа явуулах боломжийг бүрдүүлэх;
5. ҮБХИС-ийг түшиглэн хиймэл оюун ухаан, өндөр технологийн туршилт судалгааны нэгж байгуулах;
   - Цэргийн автоматжуулалт, хиймэл оюун судлалын төв байгуулах /БХЭШХ/
6. Судалгаа, туршилтын нэгжийг дэс дараатай бэхжүүлэх, тоног төхөөрөмжөөр хангах зорилго бүхий төсөл, арга хэмжээг төлөвлөн хэрэгжүүлэх;
7. Зэвсэгт хүчний нэгтгэл, анги, байгууллагад сургалт-судалгааны орчинг бий болгож, цэргийн алба хаагчдыг судалгааны ажилд оролцох боломж, хүрээг нэмэгдүүлэх;
8. Цэргийн хамтын ажиллагааны хүрээнд гадаад орны хиймэл оюун ухаан хөгжүүлэлтийн туршлагыг судалж, үр дүнг Зэвсэгт хүчинд нэвтрүүлэх;
9. Цэргийн хамтын ажиллагааны хүрээнд гадаад орны хиймэл оюун ухаан хөгжүүлэлтийн туршлагыг судалж, үр дүнг Зэвсэгт хүчинд нэвтрүүлэх;

Хүний нөөцийг бэлтгэх, хиймэл оюун ухаан хөгжүүлэх, өндөр технологийг нутагшуулах, түүнд суурилсан бүтээгдэхүүн үйлдвэрлэх, инновацын тогтолцоог бий болгох зорилтын хүрээнд:

1. Зэвсэгт хүчинд ХОУ-ны чиглэлээр мэргэшсэн инженер, судлаачдыг бэлтгэх;
2. ХОУ-ыг хөгжүүлэлтийн чиглэлээр дотоод, гадаадын мэргэжлийн багш, мэргэжилтнүүдтэй хамтарсан сургалт, зөвлөгөөн, хэлэлцүүлэг зохион байгуулах;

3. Хиймэл оюун ухаан, өндөр технологийн судалгаа-сургалт-инновац-туршилт-үйлдвэрлэлийн тогтолцоог хөгжүүлэх;

Өндөр технологийг нутагшуулах, хөгжүүлэх зорилтын хүрээнд:

1. Их өгөгдөлийг санг Батлан хамгаалах салбар, ЗХ-нд бий болгох;
   - Үүлэн технологи ашиглах
   - Блокчейн технологийг ашиглах
   - Өгөгдлийн нөөцлөлт, сэргээх журам боловсруулах
2. Машин сургалт (machine-learning)-ын алгоритмд суурилсан технологийг Батлан хамгаалах салбар, Зэвсэгт хүчинд нэвтрүүлэх;
3. Хиймэл оюун ухаан (AI)-д суурилсан робот технологийг хөгжүүлэх хүрээнд;
   - Роботын дизайн, бүтэц, задаргаа болон ашиглалтыг өргөжүүлэх
   - Хүний нөөцийг чадавхжуулах
   - Хэрэглээний хүрээг тэлэх
   - Роботын дизайн, бүтэц, задаргаа болон ашиглалтыг өргөжүүлэх
   - Хүний нөөцийг чадавхжуулах
4. Хэрэглээний хүрээг тэлэх
5. Нисгэгчгүй нисэх хэрэгсэл /Дрон/ хөгжүүлэх замаар агаарын орон зайг хянах, тандах, ХОУ-нд суурилсан системүүдийг хөгжүүлэх;
   - Нисгэгчгүй нисэх хэрэгслийг ЗХ-нд ашиглахтай холбоотой хууль, эрх зүйн орчин бүрдүүлэх, аюулгүй, найдвартай ашиглах стандарт, журмыг боловсруулах
   - Нисгэгчгүй нисэх хэрэгслийг эрэн хайх ажиллагаа, тээвэрлэлт, хилийн хяналт, аврах, кибер аюулгүй байдлын чиглэлээр хөгжүүлж ашиглах
   - Нисгэгчгүй нисэх хэрэгслийн судалгаа, туршилтыг цэргийн сургалт, бэлтгэлийн үед хэрэглэх
   - Нисгэгчгүй нисэх хэрэгслийг ашиглах, удирдах, хөгжүүлэх ур чадвартай ЦАХ-дыг бэлтгэх сургалтыг ҮБХИС-ийн ЦНДС-ийг түшиглэн явуулах
   - Энэ чиглэлээр олон улсын хамтарсан дадлага, сургуулилтыг тогтмол явуулах, дрон үйлдвэрэлэлийн чиглэлд хамтын ажиллагааг өргөжүүлэх
6. Сансрын технологийг Батлан хамгаалах салбар, ЗХ-нд хэрэглэх;
   - Нам тойрог замын цэрэг, тагнуулын зориулалттай хиймэл дагуул хөөргөх
   - Үндэсний хиймэл дагуулын 1 сувгийг /транспондер/ цэргийн зориулалтаар ашиглах
   - ЗХ-ний алслагдсан анги, салбаруудад сансрын холбоог нэвтрүүлэх
7. Биотехнологи, Нанотехнологийг нутагшуулах, шинэ материалын судалгаа, туршилт хийх;
   - Нано технологийг тусгай хүчний цэрэг, агаарын цэрэг, кибер цэрэгт хэрэглэх
   - Биотехнологи, шинэ материалын судалгаа, туршилтыг ЗХ-ний химийн лабораторуудыг түшиглэн явуулах
   - Дээрх чиглэлүүдээр мэргэжилтэн бэлтгэх
8. IoT /юмсын интернет/-д суурилсан автоматжуулалтыг хөгжүүлэх;
   - Цэргийн зориулалттай техник хэрэгслүүдийг IoT сүлжээнд холбох ажлыг эхлүүлэх
   - “IoT технологи” бакалаврын сургалтын хөтөлбөрийг ҮБХИС-д хэрэгжүүлэх
   - IoT технологийн чиглэлээр багш, судлаач, инженерийг бэлтгэх
9. AR/VR болон Metaverse технологийг нутагшуулах;
   - БХЯ-ны дуулга студийн харьяанд AR/VR болон Metaverse технологид суурилсан мультимедиа студи байгуулах
   - AR/VR, Metaverse технологийн дагнасан судалгаа, туршилтын төв байгуулах /ҮБХИС, БХЭШХ/
   - AR/VR болон Metaverse технологийг цэргийн сургалт, бэлтгэлд ашиглах
   - Тусгай хүчний салаа ротын түвшинд AR/VR, Metaverse технологийн тусгай хэрэглэлийг хэрэглэх
10. Салбарын мэдээллийн болон кибер аюулгүй байдлын журмыг боловсронгуй болгох;
    - Батлан хамгаалах салбарын хандах эрхийн хяналт болон удирдлагыг бий болгох
    - Гэмтэл сааталтай ажиллах журмыг боловсронгуй болгох
    - Кибер халдлагын үед ажиллах журмыг боловсронгуй болгох
11. Мэдээллийн аюулгүй байдал болон кибер аюулгүй байдлын хүний нөөцийг үе шаттайгаар бэлтгэх тогтолцоог боловсронгуй болгох;
    - ҮБХИС-ийн “Кибер аюулгүй байдал” хөтөлбөрийг үндэсний болон олон улсын магадлан итгэмжлэлд хамруулах
    - ҮБХИС-ийн КАБСТөвийг түшиглэн явуулах сургалтын хамрах хүрээг өргөжүүлэх

- КАБ-ын чиглэлээр мэргэжлийн багш, судлаач, инженерүүдийг гадаадад бэлтгэх

12. Монгол Улсын Ерөнхийлөгчийн санаачилсан “Үндэсний сөрөн тэсвэрлэх стратеги” боловсруулах ажлын хүрээнд салбарын “Кибер аюулгүй байдлын сөрөн тэсвэрлэх бодлого” боловсруулах, хэрэгжүүлэх;
13. Мэдээллийн нууцлал хамгаалах (Цэргийн албан хаагчдын хувийн мэдээлэл, зэвсэг техникийн аюулгүй байдал) технологийг боловсронгуй болгох;
14. Олон улсын мэдээллийн аюулгүй байдлын стандартыг хэрэгжүүлэх, Мэдээллийн аюулгүй байдлын аудит, эрсдлийн үнэлгээг тогтмол хийх;
    - ISO 270001 стандартыг нэвтрүүлэх
    - NIST кибер аюулгүй байдлын CSF2.0-ыг нэвтрүүлэх

Дүгнэлт

Хиймэл оюун ухаан, өндөр технологийн шилжилтийг батлан хамгаалах салбар, Зэвсэгт хүчинд үр дүнтэй хэрэгжүүлэхийн тулд дараах хэд хэдэн чухал алхам хийх шаардлагатайг онцолж байна. Үүнд:

1. Хиймэл оюун ухааны талаар баримтлах чиглэлийг тодорхойлсон баримт бичгийг боловсруулах, батлах
2. Хууль, эрх зүйн зохицуулалтыг тодорхой болгох: Хиймэл оюун ухааныг хариуцлагатай ашиглах, оюуны өмч, зохиогчийн эрхийн зөрчлөөс урьдчилан сэргийлэх, ёс зүй, дэг журмыг сайжруулах талаар бусад улс орны хууль, эрх зүйн зохицуулалтыг судалж, салбарын онцлогт тохируулан нутагшуулахад анхаарах
3. Шаардлагатай технологийн дэд бүтэц, мэдээллийн удирдлагын системд хөрөнгө оруулалт хийх
4. Боловсон хүчин бэлтгэх: Өндөр технологийн инженер, хиймэл оюун ухааны технологийн мэргэжилтнүүдийг бэлтгэх, чадварлаг боловсон хүчнийг бий болгох
5. Хамтын ажиллагаа: Хиймэл оюун ухааны шилжилт, хэрэгжилтийг дэмжихийн тулд хамтын ажиллагааг сайжруулах, хиймэл оюун ухаан хөгжүүлэх туршилт, судалгааны төв бий болгох, үйл ажиллагааны хамрах хүрээг өргөжүүлэх

Батлан хамгаалах салбар, Зэвсэгт хүчинд хиймэл оюун ухаан, өндөр технологийн судалгаа, туршилт, үйлдвэрлэлийн нэгдмэл үйл ажиллагааг хангахад баримтлах чиглэлийг тодорхойлох зорилготой стратеги төлөвлөгөөний хэрэгжилт нь улс орны төсөв, санхүү, салбарын хөрөнгө оруулалтын хэмжээ, хамтын ажиллагаа, нийгмийн түншлэл зэрэг олон хүчин зүйлээс шалтгаалах боловч технологийн шилжилтийн эрчимтэй хөгжлөөс хоцрохгүйн тулд зайлшгүй хэрэгжүүлэх шаардлагатай байгааг анхаарах нь зүйтэй.

Ашигласан эх сурвалж, ном зүй:

[1] “Алсын хараа-2050” Монгол Улсын урт хугацааны хөгжлийн бодлого,
[2] “Монгол Улсын батлан хамгаалах бодлогын үндэс”
[3] “Цахим үндэстэн” баримтлах чиглэл
[4] “Батлан хамгаалахын хүний нөөцийн бодлого”
[5] “Батлан хамгаалахын стратеги төлөвлөгөө”
[6] “Батлан хамгаалах яамны стратеги төлөвлөгөө”
[7] “Үндэсний батлан хамгаалахын их сургуулийн хөгжлийн стратеги”
[8] Засгийн газрын 2023 оны 07 дугаар сарын 05-ны өдрийн 249 дүгээр тогтоолоор батлагдсан “Шинжлэх ухаан, технологийн тэргүүлэх чиглэл, цөм технологийн жагсаалт” (2023-2026)
[9] Засгийн газрын 2020 оны 95 дугаар тогтоолоор батлагдсан “Инновацийн үйл ажиллагааны тэргүүлэх чиглэл”
[10] “Батлан хамгаалах судлал, цэргийн шинжлэх ухаан, технологи, инновацыг 2030 он хүртэл хөгжүүлэх бодлого ”Рябов К. Как противодействовать беспилотнику /Военное обозрение/.
[11] Y. Liu, H.-N. Dai, Q. Wang, M.K. Shukla, M. Imran, Unmanned aerial vehicle for internet of everything: Opportunities and challenges, Computer Communications (2020).
[12] S. Vashisht, S. Jain, G.S. Aujla, Mac protocols for unmanned aerial vehicle ecosystems: Review and challenges, Computer Communications (2020).
[13] J.P. Skrinjar, ˇ P. Skorput, ˇ M. Furdi´c, Application of unmanned aerial vehicles in logistic processes, in: I. Karabegovi´c (Ed.), New Technologies, Development and Application, Springer International Publishing, Cham, 2019, pp. 359–366.
[14] V. Hassija, V. Chamola, V. Gupta, S. Jain, N. Guizani, A survey on supply chain security: Application areas, security threats, and solution architectures, IEEE Internet of Things Journal (2020).
[15] Y. Unpaprom, N. Dussadeeb, R. Ramaraj, Modern Agriculture Drones: The Development of Smart Farmers, 2018, Maejo University, pp. 13–19.
[16] H. Shakhatreh, A.H. Sawalmeh, A. Al-Fuqaha, Z. Dou, E. Almaita, I. Khalil, N. S. Othman, A. Khreishah, M. Guizani, Unmanned aerial vehicles (uavs): A survey on civil applications and key research challenges, IEEE Access 7 (2019) 48572–48634.
[17] S.A. Hoseini, J. Hassan, A. Bokani, S.S. Kanhere, Trajectory optimization of flying energy sources using Q-learning to recharge hotspot UAVs, arXiv preprint arXiv: 2003.12258 (2020). 9. P. Garg, A. Srinivasan, M. Mandal, P. Narang, V. Chamola, M. Guizani, Isdnet: Aienabled instance segmentation of aerial scenes for smart cities, ACM Transactions on Internet Technology.
[18] Restas, Drone applications for supporting disaster management, World Journal of Engineering and Technology 03 (2015) 316–321, https://doi.org/10.4236/ wjet. 2015.33C047

Монгол Улсын онц чухал мэдээллийн дэд бүтцийн кибер тэсвэрлэх чадамжийг үнэлэх нь (IoT-д чиглэсэн таниур халдлагын шинжилгээ)

Үнэнбатын Эрдэнэсувд
Аюулгүй байдал судлалын докторант, ҮБХИС
Мэдээллийн технологийн тэнхим, Идэр их сургууль
Улаанбаатар хот, Монгол Улс
erdenesuvd@ider.edu.mn

Дондогмэгдийн Бямбадорж
Кибер аюулгүй байдлын тэнхим, МХТС, ШУТИС
Улаанбаатар хот, Монгол Улс
byambadorj.d@must.edu.mn

Батбаярын Дэнсмаа
Кибер аюулгүй байдлын тэнхим, МХТС, ШУТИС
Улаанбаатар хот, Монгол Улс
densmaa.b@must.edu.mn

Хураангуй—Сүүлийн жилүүдэд мэдээлэл, харилцаа холбооны технологийн хурдацтай хөгжил нь кибер халдлага, заналхийллийн цар хүрээ, нарийвчлалыг нэмэгдүүлж, эрчим хүч, санхүү, харилцаа холбоо, төрийн үйлчилгээ зэрэг онц чухал мэдээллийн дэд бүтцэд ноцтой эрсдэл учруулах болсон. Иймд нөхцөлд кибер аюулгүй байдлыг зөвхөн урьдчилан сэргийлэх хамгаалалтын хүрээнд авч үзэх нь хангалтгүй бөгөөд халдлагын үеийн тасралтгүй ажиллагаа, сэргэлт, дасан зохицох чадварыг хамарсан кибер тэсвэрлэх чадамжийн хүрээнд судлах шаардлага тулгарч байна. Энэхүү өгүүлэл нь Монгол Улсын онц чухал мэдээллийн дэд бүтцийн кибер тэсвэрлэх чадамжийг IoT орчны эмзэг байдал болон IoT-д чиглэсэн таниур халдлагын шинжтэй сүлжээний урсгалтай уялдуулан шинжилнэ. Судалгаанд баримт бичгийн шинжилгээ, харьцуулсан тойм, аналитик тайлбарын аргыг хослуулж, UB-19 халдлагын өгөгдөлд тэмдэглэгдсэн ARP spoofing, SSDP/UPnP, WS-Discovery урсгалуудыг тайлбарлав. Судалгааны үр дүнгээс эдгээр урсгал нь энгийн сүлжээний солилцоо бус, харин төхөөрөмж илрүүлэх, гарц тодорхойлох, цаашдын нэвтрэлтийн бэлтгэл хангах таниур халдлагын эхний үе шатны чухал индикатор болох нь тогтоогдов. Иймээс Монгол Улсад онц чухал мэдээллийн дэд бүтцийн кибер тэсвэрлэх чадамжийг бэхжүүлэхэд бодлого, засаглал, техникийн хяналт, хөрөнгийн харагдац, сэргээн босголтын төлөвлөлт, мэргэшсэн хүний нөөц, салбар хоорондын хамтын ажиллагааг цогцоор нь хөгжүүлэх шаардлагатай.

Түлхүүр үг—кибер тэсвэрлэх чадамж, онц чухал мэдээллийн дэд бүтэц, IoT аюулгүй байдал, ARP spoofing, SSDP, WS-Discovery, UB-19

I. УДИРТГАЛ

Дижитал шилжилт эрчимжихийн хэрээр төрийн удирдлага, эдийн засаг, санхүү, эрчим хүч, харилцаа холбоо, тээвэр, аж үйлдвэр зэрэг салбарууд мэдээллийн систем, сүлжээ, ухаалаг төхөөрөмжөөс улам бүр хамааралтай болж байна. Энэхүү хамаарал өсөхийн зэрэгцээ кибер халдлага нь зөвхөн мэдээллийн нууцлал, бүрэн бүтэн байдалд бус, үйл ажиллагааны тасралтгүй байдал, олон нийтийн аюулгүй байдал, үндэсний тогтвортой байдалд шууд нөлөөлөх болсон [1], [2].

Монгол Улсад төрийн үйлчилгээ, санхүүгийн систем, видео хяналт, сүлжээний төхөөрөмж, үйлдвэрлэлийн автоматжуулалт, харилцаа холбооны орчинд IoT суурьтай шийдлүүдийн хэрэглээ нэмэгдэж байгаа нь дижитал үр ашиг авчирч буй хэдий ч нөгөө талаас халдлагын гадаргууг өргөжүүлж байна. Ялангуяа шинэчлэлт дутмаг, анхдагч тохиргоо хэвээр үлдсэн, баталгаажуулалт сул төхөөрөмжүүд нь онц чухал мэдээллийн дэд бүтцийн эмзэг байдлыг нэмэгдүүлэх эрсдэлтэй [3]. Иймээс кибер аюулгүй байдлын асуудлыг зөвхөн халдлагыг зогсоох техникийн арга хэрэгслийн хүрээнд бус, халдлагын үед үйл ажиллагааг хэвийн хадгалах, хязгаарлагдсан үйлчилгээний горимд үргэлжлүүлэх, богино хугацаанд сэргээх, цаашид дасан зохицох кибер тэсвэрлэх чадамжийн хүрээнд авч үзэх шаардлагатай байна [1]. Энэхүү судалгааны зорилго нь Монгол Улсын онц чухал мэдээллийн дэд бүтцийн кибер тэсвэрлэх чадамжийг IoT-д чиглэсэн таниур халдлагын шинжтэй уялдуулан шинжлэх, бодлогын болон техникийн зөвлөмж боловсруулахад оршино. Үүний хүрээнд дараах зорилтуудыг дэвшүүлэв. Үүнд:

- кибер аюулгүй байдал ба кибер тэсвэрлэх чадамжийн онолын хамаарлыг тодорхойлох;
- онц чухал мэдээллийн дэд бүтцийн IoT орчны эмзэг байдлыг шинжлэх;
- бодит халдлагын өгөгдөлд тулгуурлан discovery шинжтэй урсгалын эрсдэлийг тайлбарлах;
- Монгол Улсын нөхцөлд тохирсон бодлогын зөвлөмж дэвшүүлэх.

II. СУДЛАГДСАН БАЙДАЛ

A. Кибер аюулгүй байдал ба кибер тэсвэрлэх чадамжийн хамаарал

Кибер аюулгүй байдал нь мэдээлэл, систем, сүлжээний нууцлал (confidentiality), бүрэн бүтэн

байдал (integrity), хүртээмж (availability)-ийг хамгаалахад чиглэдэг уламжлалт ойлголт юм. Харин кибер тэсвэрлэх чадамж нь энэхүү хүрээг өргөжүүлж, сөрөг нөхцөл, халдлага, эвдрэл, компромисс үүссэн үед системийн үндсэн чиг үүргийг хадгалах, сэргээх, дасан зохицох чадварыг хамардаг [1]. Иймээс кибер тэсвэрлэх чадамжийг кибер аюулгүй байдлын орлуулах ойлголт бус, харин үйл ажиллагааны тасралтгүй байдал, системийн найдвартай байдал, сэргэлтийн чадвартай уялдуулдаг. Энэ хамаарлыг дараах байдлаар томьёолж болно.

$$CR = P + D + R + Rc + A \qquad (1)$$

энд: CR = Cyber Resilience; P = Prevention; D = Detection; R = Response; Rc = Recovery; A = Adaptation.

Энэхүү томьёог практик үнэлгээнд ашиглахдаа бүрэлдэхүүн тус бүрийг бодит хэмжигдэхүүнтэй холбож болно. Тухайлбал, P нь эмзэг байдлын нөхөлтийн хувь, сүлжээний сегментчиллийн хамрах хүрээ, төхөөрөмжийн хатууруулалтын түвшин; D нь ARP, SSDP/UPnP, WS-Discovery аномали илрүүлэх хугацаа; R нь инцидентэд хариу өгөх хугацаа; Rc нь сэргээх хугацааны зорилт (RTO/RPO)-ын биелэлт; A нь сургамжид суурилсан сайжруулалтын хэрэгжилтийн хувиар хэмжигдэнэ.

B. Онц чухал мэдээллийн дэд бүтцийн тэсвэрлэх шаардлага

Онц чухал мэдээллийн дэд бүтцийн хүрээнд эрчим хүч, санхүү, харилцаа холбоо, дижитал дэд бүтэц, тээвэр, төрийн удирдлагын систем зэрэг салбарууд стратегийн ач холбогдолтой байр суурь эзэлдэг. Эдгээр салбаруудад кибер эрсдэлийг зөвхөн техникийн эмзэг байдлын түвшинд авч үзэх нь хангалтгүй бөгөөд доголдлын үед үйлчилгээг үргэлжлүүлэх, хязгаарлагдсан үйлчилгээний горимд ажиллуулах, богино хугацаанд сэргээх чадамжийг хамтад нь авч үзэх шаардлагатай [1], [2]. Ийм учраас онц чухал мэдээллийн дэд бүтцийн хамгаалалт нь зөвхөн довтолгооноос сэргийлэх асуудал биш, харин тогтвортой ажиллагааг хадгалах үндэсний чадамж юм.

C. IoT орчин дахь эрсдэл ба өртөмтгий байдал

IoT төхөөрөмжүүд нь байгууллагын болон салбарын дижитал орчны салшгүй бүрэлдэхүүн болсон хэдий ч тэдгээрийн аюулгүй байдал, нууцлалын эрсдэлийг амьдралын мөчлөгийн турш удирдах шаардлагатай [3]. IP камер, ухаалаг мэдрэгч, gateway, сүлжээний хэвлэгч, smart TV зэрэг төхөөрөмжүүд нь хязгаарлагдмал тооцоолох нөөц, шинэчлэлтийн дутагдал, анхдагч нууц үг, баталгаажуулалтын сул байдал зэргээс шалтгаалан халдлагад өртөмтгий байдаг [11]–[15]. Иймээс IoT орчинд ажиглагдах discovery, broadcast, spoofing шинжтэй урсгалуудыг зөвхөн техникийн үзэгдэл бус, харин системийн түвшний эрсдэлийн үзүүлэлт гэж үзэх шаардлагатай юм.

D. ARP, SSDP, WS-Discovery протоколын аюулгүй байдлын ач холбогдол

ARP нь локал сүлжээнд хаягийн хөрвүүлэлт хийх үндсэн механизм бөгөөд ARP spoofing нь дотоод сүлжээний итгэлцлийн загварыг ашиглан урсгалд хөндлөнгөөс нөлөөлөх, төхөөрөмжүүдийн хоорондын харилцааг төөрөгдүүлэх боломж олгодог [4]. SSDP болон UPnP нь multicast суурьтай төхөөрөмж, үйлчилгээ илрүүлэх механизм бөгөөд M-SEARCH хүсэлтээр сүлжээн дэх гарц болон бусад үйлчилгээний төхөөрөмжийг илрүүлдэг [5], [6]. WS-Discovery нь мөн multicast probe ашиглан сүлжээнд буй үйлчилгээ, төхөөрөмжүүдийг илрүүлэх протокол бөгөөд баталгаажуулалтгүй орчинд unauthorized enumeration, amplification, lateral movement зэрэг эрсдэлийг нэмэгдүүлдэг [7]. Иймээс ARP spoofing, SSDP, WS-Discovery төрлийн урсгалууд нь IoT болон кибер-физикийн орчин дахь төхөөрөмжийн харагдац, дотоод итгэлцэл, үйлчилгээний бүтцийг тандах эхний шатны гол индикаторууд болж чадна [3]–[7].

E. Олон улсын судалгааны чиг хандлага

Олон улсын судалгаанд кибер тэсвэрлэх чадамжийг уламжлалт хамгаалалтаас илүү өргөн хүрээнд авч үзсэн байдаг. Sterbenz нар сүлжээний тэсвэрлэх чадварыг алдаа, халдлага, сорилтын үед зөвшөөрөгдөх түвшний үйлчилгээг хадгалах чадвартай холбон тайлбарласан [8]. Cárdenas нар кибер-физикийн системийн хүрээнд secure control болон survivability асуудлыг дэвшүүлж, үйлдвэрлэлийн болон үйл ажиллагааны системүүд нь уламжлалт IT орчноос өөр эрсдэлийн шинжтэйг онцолсон [9]. Cherdantseva нарын тойм судалгаа нь SCADA орчинд эрсдэлийг бүтэцтэй үнэлэх аргын хэрэгцээг тодорхой харуулсан [10].

IoT орчны судалгаанд Roman нар тархмал орчин, төрөлжилт, нөөцийн хязгаарлалт зэргийг гол сорилт гэж үзсэн [11]. Sicari нар аюулгүй байдал, нууцлал, итгэлцлийн асуудлыг амьдралын мөчлөгийн хүрээнд авч үзсэн [12]. Mosenia ба Jha нар архитектурын болон хэрэгжилтийн сул талууд нь жижиг төхөөрөмжүүдийг ч стратегийн эрсдэлийн эх үүсвэр болгож болохыг онцолсон [13]. Humayed нар кибер-физикийн системийн кибер үйл явдлын физик үр дагаврыг тодотгож, тэсвэрлэх чадвар яагаад зайлшгүй хэрэгтэйг харуулсан [14]. Kolias нар Mirai ботнетийн жишээгээр хамгаалалт сул IoT төхөөрөмжүүд өргөн хүрээний DDoS халдлагын эх үүсвэр болж болохыг нотолсон [15].

III. СУДАЛГААНЫ АРГА ЗҮЙ

Судалгаанд чанарын тайлбар шинжилгээний аргыг ашиглав. Нэгдүгээрт, кибер тэсвэрлэх чадамж, онц чухал мэдээллийн дэд бүтэц, IoT аюулгүй байдалтай холбоотой онолын эх сурвалж, стандарт, судалгааны бүтээлүүдэд баримт бичгийн шинжилгээ хийв. Хоёрдугаарт, Монгол Улсын дижитал орчны онцлог, бодлогын нөхцөл байдлыг харгалзан харьцуулсан

тайлбарлал хийв. Гуравдугаарт, UB-19 халдлагын өгөгдөлд тэмдэглэгдсэн пакет түвшний ажиглалтуудыг таниур халдлагын шинж, төхөөрөмжийн илрүүлэлт, гарц тодорхойлолт, хажуугийн хөдөлгөөний боломж гэсэн аналитик ангиллаар үнэлэв. Энэхүү өгүүлэлд дараах гурван төрлийн урсгалд төвлөрөв.

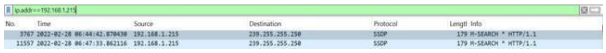
- ARP spoofing ашигласан төхөөрөмжийн илрүүлэлт;
- SSDP/UPnP ашигласан интернэт гарцын хайлт;
- WS-Discovery протокол ашигласан давталттай тандалт.

Дээрх сүлжээний урсгалыг кибер халдлагын таниур үе шат, эмзэг байдлын ашиглалтын өмнөх бэлтгэл, дэд бүтцийн түвшний эрсдэлийн дохио гэсэн дүн шинжилгээний хүрээнд тайлбарласан. Ажиглалтыг IP хаяг, протокол, порт, хүсэлтийн төрөл, давталтын шинж, эрсдэлийн тайлбар гэсэн үзүүлэлтээр нэгтгэн баталгаажуулсан нь чанарын тайлбарыг бодит пакетийн түвшний индикатортой холбох боломж бүрдүүлэв.

IV. СУДАЛГААНЫ ҮР ДҮН

A. ARP spoofing ашигласан төхөөрөмж илрүүлэлт

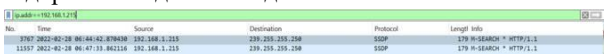
Судалгааны өгөгдөлд IP 192.168.1.6 хаягтай төхөөрөмжөөс ARP spoofing шинжтэй урсгал ажиглагдсан. Энэ нь байгууллагын дотоод сүлжээнд холбогдсон IP хяналтын камерын төхөөрөмжийг илрүүлэх зорилготой байж болохыг харуулж байна. ARP spoofing нь зөвхөн пакет дамжуулалтад хөндлөнгөөс оролцох арга бус, харин сүлжээний бүтэц, зангилааны байршил, төхөөрөмжийн холбоосыг тогтоох боломж олгодог. Иймд энэ төрлийн урсгал нь дараагийн шатны халдлагын бэлтгэлийн нэг хэсэг гэж тайлбарлагдана.



Зураг 1. IP 192.168.1.6 хаягтай төхөөрөмжөөр ARP spoofing халдлага хийж байгаа байдал

B. SSDP/UPnP ашигласан интернэт гарцын илрүүлэлт

Мөн IP 192.168.1.215 хаягтай төхөөрөмжөөс 239.255.255.250:1900 multicast хаяг руу SSDP M-SEARCH хүсэлт илгээж, Internet Gateway Device төрлийн төхөөрөмжийг хайж буй шинж илэрсэн. Энэ нь халдагч зөвхөн эцсийн төхөөрөмжүүдийг бус, дотоод сүлжээний гарц, маршрутын зангилаа, удирдлагын цэгүүдийг илрүүлэх зорилго агуулж болзошгүйг харуулна. UPnP/SSDP орчин дахь ийм discovery хүсэлт нь сүлжээний топологи, үйлчилгээний бүтэц, боломжит нэвтрэх цэгүүдийг тодорхойлоход ашиглагдаж болно.



Зураг 2. IP 192.168.1.215 хаягтай төхөөрөмжөөс SSDP протокол руу хандалт хийж байгаа байдал

Зураг 3-т SSDP M-SEARCH хүсэлт нь Internet Gateway Device төрлийн гарц төхөөрөмжийг хайж буй пакетийн агуулгыг харуулав.

```
M-SEARCH * HTTP/1.1
HOST: 239.255.255.250:1900
ST: urn:schemas-upnp-org:device:InternetGatewayDevice:1
MAN: "ssdp:discover"
MX: 1
```

Зураг 3. IP 239.255.255.250:1900 хаягийн интернэт гарцыг хайж байгаа байдал

Хүснэгт 1-д SSDP M-SEARCH хүсэлтийн үндсэн талбарууд болон тэдгээрийн утгыг нэгтгэн харуулав.

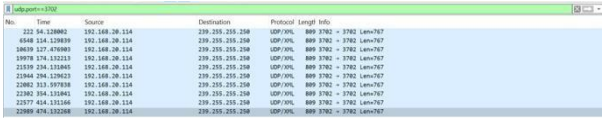
Хүснэгт 1. Хяналтын камерын төхөөрөмжөөс SSDP протокол руу HTTP хүсэлт илгээж байгаа байдал

Талбар	Тайлбар
1. M-SEARCH *	HTTP M-SEARCH үйлдэл бөгөөд сүлжээн дэх UPnP төхөөрөмжүүдийг хайх зориулалттай
2. HOST: 239.255.255.250:1900	SSDP multicast хаяг ба порт
3. ST (Search Target)	urn:schemas-upnp-org:device:InternetGatewayDevice:1; өөрөөр хэлбэл интернэт гарц төхөөрөмжийг хайж байгааг илтгэнэ
4. MAN: "ssdp:discover"	UPnP discovery хүсэлт болохыг заана
5. MX: 1	Хариу өгөх дээд хүлээлгийн хугацаа

Хүснэгт 1-ээс харахад HTTP хүсэлт нь сүлжээнд буй UPnP төхөөрөмж болон интернэт гарцын төхөөрөмжийг тогтоох зорилготой байгааг харуулж байна. Иймээс энэ урсгал нь сүлжээний таниур үйл ажиллагааны дараагийн шатны нэг хэлбэр гэж үзэж болно.

C. WS-Discovery протоколын давталттай тандалт

Судалгааны өгөгдөлд UDP 3702 порт руу олон давталттай probe шинжтэй урсгал илэрсэн. UDP 3702 нь WS-Discovery протоколын үйл ажиллагаанд ашиглагддаг бөгөөд дотоод сүлжээнд байгаа төхөөрөмж, үйлчилгээг автоматаар илрүүлэх боломж олгодог. Гэвч энэхүү протокол нь баталгаажуулалт, шифрлэлгүй орчинд ашиглагдах үед unauthorized device enumeration, lateral movement, reflection/amplification эрсдэлийг нэмэгдүүлэх боломжтой. Энэ нь онц чухал мэдээллийн дэд бүтцийн орчинд онцгой анхаарах шаардлагатай эмзэг шинж тэмдэг юм.



Зураг 4. IP 192.168.20.114 хаягтай төхөөрөмжөөс UDP 3702 порт руу олон давталттай тандалт хийж буй байдал

D. Үр дүнгийн нэгтгэл

Дээрх ажиглалтууд нь ARP spoofing, SSDP/UPnP, WS-Discovery урсгалууд тус тусдаа салангид үзэгдэл бус, харин дараалсан таниур логиктой байгааг харуулж байна. Уг дарааллыг дараах байдлаар нэгтгэн үзэж болно.



Зураг 5. Таниур халдлагыг илрүүлэх аргачлал

Зураг 5-д харуулсан халдлагыг техникийн лог биш, онц чухал мэдээллийн дэд бүтцийн эсрэг халдлагын эрт үеийн индикатор гэж үзэх үндэслэлтэй юм.

V. ХЭЛЭЛЦҮҮЛЭГ

Уг судалгааны үр дүн нь Монгол Улсад кибер аюулгүй байдлын асуудлыг зөвхөн хамгаалалтын хэрэгсэл суурилуулах түвшинд авч үзэх нь хангалтгүйг харуулж байна. Ялангуяа онц чухал мэдээллийн дэд бүтцийн хүрээнд ашиглагдаж буй IoT төхөөрөмжүүд нь аюулгүй байдлын удирдлага сул үед халдагчийн таниур, нэвтрэлт, өргөтгөсөн халдлагын суурь цэг болж хувирч байна. Иймээс үндэсний хэмжээний кибер тэсвэрлэх чадамжийг дараах чиглэлүүдэд сайжруулах шаардлагатай. Тухайлбал:

Нэгдүгээрт, бодлого, эрх зүйн орчны түвшинд онц чухал мэдээллийн дэд бүтцийн ангилал, хамгаалалтын доод шаардлага, эрсдэлийн үнэлгээ, инцидентын тайлагналын механизмыг салбар бүрт илүү тодорхой болгох шаардлагатай.

Хоёрдугаарт, техникийн түвшинд сүлжээний сегментчилэл, asset inventory, IoT device hardening, firmware шинэчлэлтийн удирдлага, multicast discovery урсгалын хяналт, аномали илрүүлэлт, zero-trust зарчмыг нэвтрүүлэх хэрэгтэй.

Гуравдугаарт, үйл ажиллагааны түвшинд халдлагын үед хязгаарлагдсан үйлчилгээний горимд ажиллах, түргэн сэргээх, нөөц горимд шилжих төлөвлөгөө, сургуулилалт тогтмол хэрэгжүүлэх шаардлагатай.

Дөрөвдүгээрт, хүний нөөцийн түвшинд IoT, OT, ICS орчны аюулгүй байдлын чиглэлээр мэргэшсэн боловсон хүчин системтэй бэлтгэх нь нэн чухал.

Тавдугаарт, салбар хоорондын хамтын ажиллагааны түвшинд төр, хувийн хэвшил, судалгааны байгууллага, үндэсний CERT хооронд аюулын мэдээлэл солилцох, хамтарсан дасгал сургуулилт хийх, сургамж хуваалцах тогтолцоо шаардлагатай.

VI. ДҮГНЭЛТ

Энэхүү өгүүлэлд Монгол Улсын онц чухал мэдээллийн дэд бүтцийн кибер тэсвэрлэх чадамжийг IoT-д чиглэсэн таниур шинжтэй халдлагын урсгалтай уялдуулан шинжилж, UB-19 өгөгдөлд илэрсэн ARP spoofing, SSDP/UPnP, WS-Discovery урсгалуудыг таниур халдлагын үе шаттай холбон үнэлэв. Судалгааны үр дүнгээр дараах дүгнэлт гарлаа.

Нэгдүгээрт, CR = P + D + R + Rc + A томьёо нь урьдчилан сэргийлэлт, илрүүлэлт, хариу арга хэмжээ, сэргэлт, дасан зохицох чадварыг нэгтгэн үнэлэх боломжтой бөгөөд эдгээр бүрэлдэхүүнийг бодит шалгуур үзүүлэлтээр хэмжих шаардлагатайг харууллаа.

Хоёрдугаарт, 192.168.1.6 хаягаас ажиглагдсан ARP spoofing урсгал нь дотоод сүлжээний төхөөрөмж, ялангуяа IP камерын зангилааг илрүүлэх, хандалтын цэгийг тодорхойлох таниур шинжтэй индикатор болохыг харууллаа.

Гуравдугаарт, 192.168.1.215 хаягаас 239.255.255.250:1900 multicast хаяг руу илгээсэн SSDP M-SEARCH хүсэлт нь Internet Gateway Device төрлийн гарц төхөөрөмжийг тодорхойлох оролдлого байж болох тул сүлжээний топологийг тандах эрсдэлийн дохио гэж үзэв.

Дөрөвдүгээрт, 192.168.20.114 хаягтай төхөөрөмжөөс UDP 3702 порт руу ажиглагдсан WS-Discovery probe урсгал нь unauthorized enumeration, lateral movement, reflection/amplification эрсдэлийн өмнөх шатны дохио байж болохыг тогтоов.

Иймээс Монгол Улсад онц чухал мэдээллийн дэд бүтцийн хамгаалалтыг уламжлалт хамгаалалтын логикоос кибер тэсвэрлэх чадамжид суурилсан, бодит сүлжээний индикатороор баталгааждаг эрсдэлийн удирдлагын тогтолцоо руу үе шаттайгаар шилжүүлэх нь стратегийн тэргүүлэх зорилт байх ёстой.

ТАЛАРХАЛ

Судалгааны ажилд үнэтэй санал, зөвлөгөө өгч дэмжлэг үзүүлсэн ШУТИС-ийн МХТС-ийн Кибер аюулгүй байдлын судалгааны багийн хамт олонд талархал илэрхийлье.

НОМ ЗҮЙ

[1] NIST, Cyber Resiliency Engineering Framework.
[2] ENISA, Cybersecurity and Resilience of Critical Sectors.
[3] NIST IR 8228, Considerations for Managing Internet of Things (IoT) Cybersecurity and Privacy Risks.
[4] D. C. Plummer, “An Ethernet Address Resolution Protocol,” RFC 826, 1982.
[5] Y. Goland et al., Simple Service Discovery Protocol/1.0.
[6] UPnP Forum, UPnP Device Architecture 1.1.
[7] OASIS, Web Services Dynamic Discovery (WS-Discovery) Version 1.1.
[8] J. P. G. Sterbenz et al., “Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines,” Computer Networks, 2010.
[9] A. A. Cárdenas, S. Amin, and S. Sastry, “Research challenges for the security of control systems,” HotSec, 2008.
[10] Y. Cherdantseva et al., “A review of cyber security risk assessment methods for SCADA systems,” Computers & Security, 2016.
[11] R. Roman, P. Najera, and J. Lopez, “Securing the Internet of Things,” Computer, 2011.
[12] S. Sicari, A. Rizzardi, L. A. Grieco, and A. Coen-Porisini, “Security, privacy and trust in Internet of Things: The road ahead,” Computer Networks, 2015.

[13] A. Mosenia and N. K. Jha, “A comprehensive study of security of Internet-of-Things,” IEEE Transactions on Emerging Topics in Computing, 2017.
[14] A. Humayed, J. Lin, F. Li, and B. Luo, “Cyber-Physical Systems Security—A Survey,” IEEE Internet of Things Journal, 2017.
[15] C. Kolias, G. Kambourakis, A. Stavrou, and J. Voas, “DDoS in the IoT: Mirai and other botnets,” Computer, 2017.
[16] W. Knowles et al., “A survey of cyber security management in industrial control systems,” International Journal of Critical Infrastructure Protection, 2015.
[17] C.-W. Ten, C.-C. Liu, and G. Manimaran, “Cybersecurity for critical infrastructures: Attack and defense modeling,” IEEE Transactions on Systems, Man, and Cybernetics, 2008.

Динамик DRM Систем Ашиглан Видео Контентийн Хууль Бус Хуулбарлалтаас Сэргийлэх

Мөнх-Очирийн БАТ-ОРГИЛ
Кибер аюулгүй байдлын судалгааны баг,
Кибер аюулгүй байдлын тэнхимийн оюутан,
ШУТИС, Мэдээлэл, холбооны технологийн сургууль,
Улаанбаатар, Монгол улс
origoo116@gmail.com

Сонинбаярын СУВД
Кибер аюулгүй байдлын судалгааны баг,
Кибер аюулгүй байдлын тэнхимийн оюутан,
ШУТИС, Мэдээлэл, холбооны технологийн сургууль,
Улаанбаатар, Монгол улс
suvdaaomg@gmail.com

Тамирын АРВИН
Кибер аюулгүй байдлын судалгааны баг,
Кибер аюулгүй байдлын тэнхимийн оюутан,
ШУТИС, Мэдээлэл, холбооны технологийн сургууль,
Улаанбаатар, Монгол улс
arwintamirs@gmail.com

Бат-Эрдэнийн МӨНХБАЯР
Кибер аюулгүй байдлын судалгааны баг,
Кибер аюулгүй байдлын тэнхимийн эрхлэгч, доктор(Ph.D.), дэд профессор
ШУТИС, Мэдээлэл, холбооны технологийн сургууль,
Улаанбаатар, Монгол улс
munkhbayar.b@must.edu.mn

Хураангуй: Дижитал контентыг хууль бусаар хуулбарлах, түгээх асуудал нь контент бүтээгчид болон түгээгч байгууллагуудын хувьд эдийн засаг, аюулгүй байдлын томоохон эрсдэл болж байна. Уламжлалт Digital Rights Management (DRM) системүүд нь ихэвчлэн шифрлэлт, лицензийн шалгалтад тулгуурласан статик хамгаалалтын механизм ашигладаг боловч хэрэглэгчийн орчны нөхцөл, сэжигтэй үйлдлийг бодит цагт хангалттай үнэлэх боломж хязгаарлагдмал хэвээр байна. Иймээс зөвхөн түлхүүр болон лицензийн түвшний хамгаалалтаас илүү уян хатан, нөхцөлд суурилсан шийдвэр гаргах шаардлагатай байна. Энэхүү судалгаанд хэрэглэгчийн орчны үзүүлэлтүүдэд суурилсан динамик DRM хамгаалалтын механизмыг санал болгосон. Уг механизм нь хандалтын токены хүчинтэй байдал, дэлгэц бичигч, дебаггер ашиглалт, клиент талын зөвшөөрөлгүй өөрчлөлт, төхөөрөмжийн өөрчлөлт, давтамжтай алдаатай хүсэлт зэрэг үзүүлэлтүүдийг нэгтгэн эрсдэлийн үнэлгээ хийж, хандалтыг ALLOW, SUSPICIOUS, BLOCK гэсэн түвшинд ангилан шийдвэр гаргадаг. Эрсдэлийн үнэлгээний загвар нь сэжигтэй үйлдэл бүрд харгалзан оноо өгч, токен хүчингүй үед шууд хандалтыг хориглох нөхцөлийг агуулсан болно. Санал болгож буй аргыг симуляцийн орчинд хэрэгжүүлэн, таван төрлийн туршилтын нөхцөлөөр үнэлэв. Туршилтын үр дүнд хэрэглэгчийн нөхцөл хэвийн үед систем хандалтыг зөвшөөрсөн бол дэлгэц бичигч болон дебаггер илэрсэн нөхцөлд сэжигтэй ангилалд хамааруулж, клиент талын зөвшөөрөлгүй өөрчлөлт болон хүчингүй токен илэрсэн тохиолдолд хандалтыг шууд хориглосон байна. Судалгааны үр дүн нь DRM хамгаалалтыг статик түлхүүрийн түвшнээс давуулан, хэрэглэгчийн орчныг үнэлж динамик шийдвэр гаргадаг илүү уян хатан аргачлал руу өргөтгөх боломжтойг харуулж байна.

Түлхүүр үг: Эрсдэлийн үнэлгээ, Хандалтын шийдвэр, Сэжигтэй нөхцөл, Токен шалгалт, Клиент талын зөвшөөрөлгүй өөрчлөлт

I. Удиртгал

Дижитал технологийн хөгжил, өндөр хурдны интернэт болон стриминг үйлчилгээний өргөн хэрэглээ нь мультимедиа контентыг дэлхий даяар хурдан түгээх боломжийг бүрдүүлсэн. Гэвч энэхүү хөгжлийн зэрэгцээ дижитал хууль бус хуулбарлалт эрчимтэй нэмэгдэж, контент бүтээгчид болон түгээгч байгууллагуудад эдийн засгийн ноцтой хохирол учруулах болсон. Судалгаагаар 2023 онд дижитал хулгайн сайтуудад хандсан нийт хандалтын тоо 229 тэрбумд хүрч, өмнөх оноос 6.7%-иар өссөн байна. Мөн кино, телевизийн салбар зөвхөн 2022 онд дижитал хуулбарлалтын улмаас 50 тэрбум ам.долларын алдагдал хүлээсэн нь уг асуудлын ноцтой байдлыг харуулж байна [1].

Энэхүү асуудлыг бууруулах зорилгоор Digital Rights Management системүүдийг өргөн хэрэглэж байна. DRM системүүд нь криптографын аргууд, ялангуяа Advanced Encryption Standard зэрэг симметрик шифрлэлтийг ашиглан контентыг хамгаалж, лицензийн сервер болон баталгаажуулалтын механизмаар дамжуулан зөвшөөрөгдсөн хэрэглэгчдэд хандалт олгодог [3]. Widevine, PlayReady, FairPlay зэрэг орчин үеийн DRM шийдлүүд нь веб болон мобайл орчинд хамгаалалттай контент түгээхэд чухал үүрэг гүйцэтгэж байна [2], [4].

Гэсэн хэдий ч уламжлалт DRM системүүд нь ихэвчлэн статик лицензийн шалгалт болон түлхүүрийн удирдлагад тулгуурладаг бөгөөд хэрэглэгчийн орчны өөрчлөлт, сэжигтэй үйлдлийг динамикаар үнэлэх боломж хязгаарлагдмал хэвээр байна. Жишээлбэл, дэлгэц бичигч ашиглах, дебаггер ажиллуулах, клиент талд өөрчлөлт хийх, токеныг буруугаар ашиглах зэрэг халдлагын нөхцөлүүдэд хамгаалалтын механизм хангалтгүй ажиллах тохиолдол ажиглагддаг. Мөн DRM системүүд нь төвлөрсөн лицензийн серверээс хэт хамааралтай байдаг нь системийн найдвартай байдал болон хэрэглэгчийн туршлагад сөргөөр нөлөөлж болзошгүй [6].

Иймээс DRM хамгаалалтыг зөвхөн шифрлэлт болон лицензийн хяналтаар хязгаарлах бус, хэрэглэгчийн хандалтын токены хүчинтэй байдал, сэшний төлөв, төхөөрөмжийн өөрчлөлт, дэлгэц бичлэг болон дебаггер ашиглалт зэрэг орчны болон үйлдлийн тодорхой үзүүлэлтүүдэд суурилсан динамик эрсдэлийн үнэлгээний механизмтай хослуулах шаардлага үүсэж байна.

Энэхүү судалгааны ажлын зорилго нь криптограф болон баталгаажуулалтын үндсэн зарчмыг хадгалсан хэвээр, дээрх үзүүлэлтүүдэд тулгуурлан эрсдэлийг тооцоолох, улмаар хандалтыг **ALLOW, SUSPICIOUS, BLOCK** гэсэн түвшинд ангилан шийдвэр гаргах динамик DRM механизмыг санал болгоход оршино.

II. Судалгааны Ажлууд

Дижитал контентыг хамгаалах зорилгоор Digital Rights Management системүүдийг өргөн хэрэглэж ирсэн бөгөөд эдгээр системүүд нь криптографын аргууд, лицензийн удирдлага болон баталгаажуулалтын механизмыг хослуулан контентын хандалтыг хянах үндсэн үүрэг гүйцэтгэдэг [2]. Ялангуяа Advanced Encryption Standard зэрэг симметрик шифрлэлтийн алгоритмууд нь өндөр аюулгүй байдал, гүйцэтгэлээрээ DRM системүүдэд түгээмэл ашиглагддаг [3].

Орчин үеийн мультимедиа үйлчилгээ болон стриминг платформууд нь Widevine, PlayReady, FairPlay зэрэг DRM технологиудыг ашиглан хамгаалалттай контент түгээж байна. Widevine DRM нь Android, веб браузер болон Smart TV орчинд өргөн хэрэглэгддэг бөгөөд Encrypted Media Extensions болон Content Decryption Module-д суурилсан архитектур ашигладаг. Харин PlayReady DRM нь Windows болон Xbox платформд лицензийн удирдлага болон төхөөрөмжийн баталгаажуулалтыг хэрэгжүүлдэг бол FairPlay DRM нь Apple төхөөрөмжүүдэд HLS стриминг орчинд хамгаалалт үзүүлдэг [4].

Гэсэн хэдий ч эдгээр DRM системүүдэд тодорхой сул талууд оршсоор байна. Судалгаагаар Widevine системд криптограф түлхүүр задлах, дэлгэц бичлэгийг тойрч гарах, hardware acceleration-тай холбоотой эмзэг байдал зэрэг асуудлууд илэрсэн байна. PlayReady системд түлхүүр хулгайлах болон лицензийн сервертэй холбоотой эмзэг байдал ажиглагдсан бол FairPlay системд дэлгэц бичлэг болон төхөөрөмжийн нийцтэй байдлын хязгаарлалт байсаар байна [4]. Эдгээр сул талууд нь DRM хамгаалалтыг тойрон гарах боломжийг нэмэгдүүлж, дижитал хуулбарлалтыг бүрэн зогсоох боломжийг хязгаарлаж байна.

Нөгөө талаас, баталгаажуулалтын механизм нь DRM системийн чухал бүрэлдэхүүн хэсэг боловч ихэвчлэн токены хүчинтэй байдлыг шалгах түвшинд хязгаарлагддаг. Жишээлбэл, OAuth 2.0 токен шалгах механизм нь хэрэглэгчийн эрхийг баталгаажуулахад ашиглагддаг боловч хэрэглэгчийн орчны аюулгүй байдал, сэжигтэй үйлдлийг үнэлэх боломж хангалтгүй байдаг [5]. Үүнээс гадна дижитал хуулбарлалтыг судалсан ажлуудад дижитал хулгайн хэрэглээ өндөр хэвээр байгааг онцолж, уламжлалт хамгаалалтын аргууд дангаараа хангалтгүй байгааг тэмдэглэсэн байна [1].

Иймээс орчин үеийн судалгаанд статик хамгаалалтаас илүү динамик, орчны мэдээлэлд суурилсан хандлага руу шилжих шаардлага байгааг онцолж байна. Хэрэглэгчийн хандалтын токены төлөв, сэшний хэв шинж, төхөөрөмжийн өөрчлөлт, мөн дэлгэц бичлэг болон дебаггер ашиглалт зэрэг үзүүлэлтүүдийг бодит цагт үнэлэх замаар илүү уян хатан хамгаалалтын механизм хэрэгжүүлэх боломжтой гэж үзэж байна.

Дээрх судалгаануудаас харахад DRM системүүд нь криптограф болон лицензийн удирдлагын түвшинд хангалттай хөгжсөн боловч хэрэглэгчийн орчны нөхцөл болон сэжигтэй үйлдлийг динамикаар үнэлэх, шийдвэр гаргах механизмыг бүрэн хэрэгжүүлээгүй хэвээр байна. Энэ нь DRM хамгаалалтыг илүү уян хатан, орчны мэдээлэлд суурилсан байдлаар өргөтгөх шаардлагатай байгааг харуулж байна.

DRM системүүдийн үндсэн механизм болон сул талуудын харьцуулалтыг “TABLE I” дээр харуулав.

TABLE I. DRM СИСТЕМҮҮДИЙН ХАРЬЦУУЛАЛТ

DRM систем	**Платформ**	**Аюулгүй байдлын механизм**	**Сул тал**
Widewine	Android, Browser, Smart TV	AES, EME/CDM, License server	Key extraction, screen bypass
PlayReady	Windows, Xbox	AES, License server, Authentication	Key theft, server vulnerability
FairPlay	Apple төхөөрөмжүүд	AES, HLS, License server	Screen recording, compatibility

III. Асуудлын Тодорхойлолт

Орчин үеийн DRM системүүд нь дижитал контентыг хамгаалахад өргөн хэрэглэгдэж байгаа боловч бодит хэрэглээний орчинд тодорхой хязгаарлалтуудтай хэвээр байна. Эдгээр системүүд нь ихэвчлэн криптограф болон лицензийн хяналтад тулгуурласан статик хамгаалалтын механизм ашигладаг бөгөөд хэрэглэгчийн орчны өөрчлөлт, сэжигтэй үйлдлийг бодит цагт үнэлэх боломж хангалтгүй байдаг.

DRM системүүд нь төхөөрөмжийн нийцтэй байдалд ихээхэн хамааралтай байдаг. Зарим тохиолдолд зөвхөн баталгаажсан төхөөрөмжүүдэд өндөр чанартай контент дамжуулах боломжийг олгодог бөгөөд энэ нь хэрэглэгчийн хүртээмжийг хязгаарлах сул талтай. Үүнээс гадна дэлгэц бичлэг болон hardware acceleration-ийг өөрчлөх зэрэг аргуудыг ашиглан DRM хамгаалалтыг тойрон гарах боломж оршдог [4].

Мөн DRM системүүд нь ихэвчлэн төвлөрсөн лицензийн серверээс хамааралтай байдаг бөгөөд серверийн доголдол, сүлжээний саатал үүссэн тохиолдолд контентод хандах боломжийг хязгаарлах эрсдэлтэй. Энэ нь системийн найдвартай байдал болон хэрэглэгчийн туршлагад сөргөөр нөлөөлж болзошгүй.

Нөгөө талаас баталгаажуулалтын механизм нь DRM хамгаалалтын чухал бүрэлдэхүүн хэсэг боловч ихэнх тохиолдолд зөвхөн токены хүчинтэй байдлыг шалгах түвшинд хязгаарлагддаг. Жишээлбэл, OAuth 2.0 токен шалгах арга нь хэрэглэгчийн хандалтыг баталгаажуулахад ашиглагддаг боловч хэрэглэгчийн орчны аюулгүй байдал, тухайлбал сэшний хэвийн бус байдал, төхөөрөмжийн өөрчлөлт, клиент талын өөрчлөлт зэрэг нөхцөлүүдийг хангалттай үнэлдэггүй [5].

Дээрх хязгаарлалтууд нь DRM хамгаалалтыг тойрон гарах боломжийг нэмэгдүүлж, дижитал хуулбарлалтыг бүрэн зогсоох боломжийг бууруулж байна. Иймээс зөвхөн статик хамгаалалтад тулгуурлах

бус, хэрэглэгчийн орчны үзүүлэлтүүдийг бодит цагт үнэлж, нөхцөлд суурилсан шийдвэр гаргах боломжтой динамик DRM механизм шаардлагатай байна.

IV. Санал Болгосон Арга

Энэхүү судалгаанд уламжлалт DRM системүүдийн хязгаарлалтыг даван туулах зорилгоор хэрэглэгчийн орчны үзүүлэлтүүдэд суурилсан динамик DRM хамгаалалтын механизмыг санал болгож байна. Уг арга нь криптограф болон баталгаажуулалтын үндсэн зарчмыг хадгалсан хэвээр, хэрэглэгчийн хандалтын нөхцөл болон системийн орчны төлөвийг бодит цагт үнэлэх замаар хандалтыг хянах боломжийг бүрдүүлдэг.

Санал болгож буй механизм нь клиент болон серверийн орчноос цуглуулсан өгөгдөлд тулгуурлан эрсдэлийн үнэлгээг тооцоолдог. Үүнд хандалтын токены хүчинтэй байдал, сэшний төлөв, төхөөрөмжийн өөрчлөлт, дэлгэц бичлэг, дебаггер ашиглалт, мөн клиент талын өөрчлөлт зэрэг үзүүлэлтүүдийг хамруулсан. Эдгээр үзүүлэлтүүд нь хэрэглэгчийн хэвийн бус үйл ажиллагааг илрүүлэх, түүнчлэн хамгаалалтыг тойрон гарах оролдлогыг тодорхойлох зорилготой.

Эрсдэлийн үнэлгээнд ашигласан үзүүлэлтүүд болон тэдгээрийн оноог “TABLE II” дээр үзүүлэв.

TABLE II. Эрсдэлийн Үнэлгээний Үзүүлэлтүүд

Үзүүлэлт	**Оноо**
Токен хүчингүй	100
Клиент талын зөвшөөрөлгүй өөрчлөлт илэрсэн	40
Дэлгэц бичигч илэрсэн	35
Дебаггер илэрсэн	25
Давтамжтай алдаатай хүсэлт	25
Төхөөрөмжийн өөрчлөлт	15

Эдгээр үзүүлэлтүүдэд үндэслэн нийт эрсдэлийн оноог тооцоолж, урьдчилан тогтоосон босго утгуудтай харьцуулан хандалтын түвшнийг тодорхойлно.

Хүчингүй токен илэрсэн тохиолдолд шууд 100 оноо өгч контентийг өгөхгүй. Клиент талын зөвшөөрөлгүй өөрчлөлт нь систем хамгийн их хохирол учруулах боломжтой, контентийг бүрэн хулгайлах боломж бүрдүүлдэг тул хамгийн өндөр оноо буюу 40 оноо өгсөн. Дэлгэц бичлэг нь бусад DRM хамгаалалтаас тойрох хамгийн түгээмэл арга тул 35 оноо өгсөн. Эдгээр 2 үзүүлэлт илэрсэн тохиолдолд шууд сэжигтэй гэж үзнэ.

Дебаггерийг ихэвчлэн техникийн мэдлэгтэй хэрэглэгчид ашиглах магадлалтай боловч дангаараа хохирол үзүүлэх боломжгүй тул 25 оноо өгч зөвшөөрөх хариу үйлдэл хийнэ. Давтамжтай алдаатай хүсэлт ирүүлэх нь автомат сканерын шинж ба дангаараа хохирол үзүүлэхгүй тул мөн 25 оноо өгсөн. Эдгээр 2 үзүүлэлт нэг удаа илэрэхийг зөвшөөрнө.

Төхөөрөмжийн өөрчлөлт нь энгийн хэрэглэгчддэд тохиолдох үзүүлэлт ба хамгийн бага 15 оноо өгсөн. Төхөөрөмжийн өөрчлөлт 2-оос доош удаа илэрэхийг зөвшөөрч, 2-оос 4-н удаа илэрэхэд сэжигтэй гэж үзнэ.

Үзүүлэлтүүдийн оноонуудыг харгалзан дараах байдлаар ангилна:

- Оноо ≥ 70 тохиолдолд хандалтыг хориглох (BLOCK)
- Оноо ≥ 35 тохиолдолд сэжигтэй (SUSPICIOUS) гэж ангилах
- Оноо < 35 тохиолдолд хэвийн (ALLOW) гэж үзэх

Дээрх 7-н үзүүлэлтээс дебаггер илэрсэн, давтамжтай алдаатай хүсэлт ирүүлсэн болон төхөөрөмжийн өөрчлөлт илэрсэн тохиолдолд зөвшөөрөх ба харин клиент талын зөвшөөрөлгүй өөрчлөлт илэрсэн мөн дэлгэц бичигч илэрсэн тохиолдолд сэжигтэй гэж үзнэ. Хэрэв токены хүчингүй байвал систем шууд хандалтыг хориглох шийдвэр гаргана. Эдгээр үзүүлэлтүүд давхардаж илэрсэн бол үзүүлэлтүүдийн нийлбэр оноонд үндэслэн хандалтыг ангилна.

Санал болгож буй DRM хамгаалалтын механизмыг “Fig. 1.” дээр үзүүлэв.

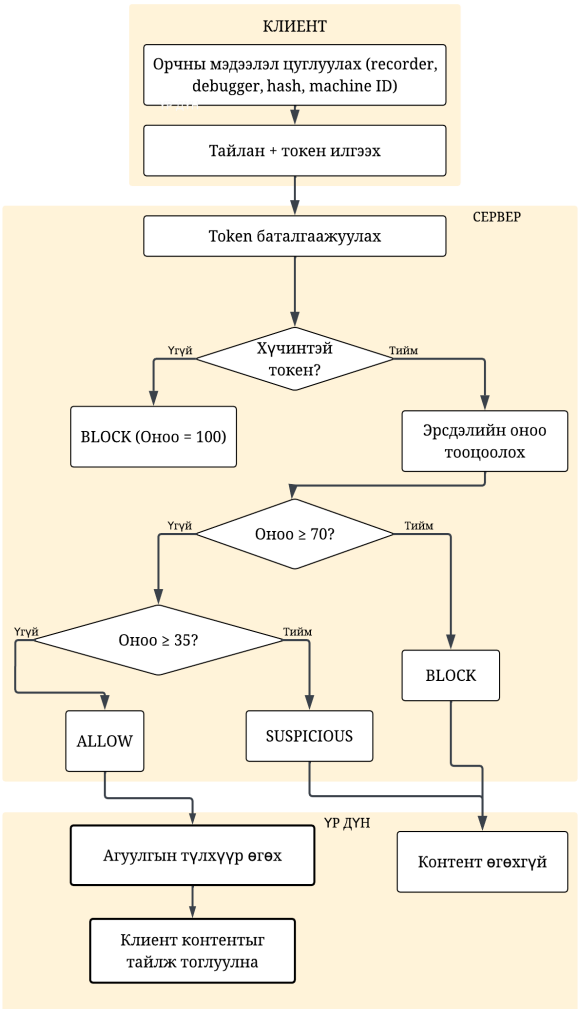


Fig. 1. DRM хамгаалалтын механизм

Одоогийн байдлаар сэжигтэй буюу SUSPICIOUS гэсэн тохиолдолд контентийг өгөхгүй хариу үйлдэл үзүүлнэ. Цаашид SUSPICIOUS тохиолдолд үзүүлсэн үйлдлүүдээс хамааруулж machine learning data driven аргын ашиглан үйлдэл тус бүр дээр харгалзан тодорхой нэг хариу үйлдэл хийхээр нарийвчлах боломжтой.

Системийн ажиллагаа нь дараах дарааллаар хэрэгжинэ. Эхлээд клиент талаас хэрэглэгчийн үйл ажиллагаатай холбоотой өгөгдлийг цуглуулж сервер рүү илгээнэ. Дараа нь сервер тухайн өгөгдлийг боловсруулж, эрсдэлийн оноог тооцоолно. Улмаар тооцоолсон оноог босго утгуудтай харьцуулж, хандалтын түвшнийг тодорхойлно. Ингэснээр хэрэглэгчийн орчны нөхцөлд суурилсан динамик, нөхцөлд дасан зохицох чадвартай хамгаалалтын механизм хэрэгжиж байна.

Энэхүү арга нь уламжлалт статик DRM хамгаалалтаас ялгаатайгаар хэрэглэгчийн орчны үзүүлэлтүүдийг бодит цагт үнэлэх, эрсдэлийн түвшинд үндэслэн нөхцөлд тохирсон шийдвэр гаргах боломжийг бүрдүүлж байна. Ингэснээр хамгаалалтыг тойрон гарах оролдлогыг илүү үр дүнтэй илрүүлэх, дижитал контентыг хамгаалах түвшнийг сайжруулах давуу талтай. Харуулав

V. Туршилтын өгөгдлийн бүтэц

Туршилтыг симуляцийн орчинд хэрэгжүүлж, системийн логикийг Python программчлалын хэл ашиглан загварчилсан бөгөөд өгөгдлийг JSON бүтэцтэйгээр боловсруулсан.

Туршилтад ашигласан параметрүүдийг III-р хүснэгтэд үзүүлэв.

TABLE III. Ашигласан параметрүүд

Параметр	Тодорхойлолт
Эх үүсвэр	Клиент талын бодит цагийн тайлан болон серверийн лог файл (event_log.json)
Бүтэц	Туршилтын таван нөхцөл (T1–T5)-д харгалзах хандалтын хүсэлтүүдийн бүртгэл
Цуглуулсан өгөгдөл	Хэрэглэгчийн нэр, machine_id, client_hash, recorder_on, debugger_on, токен, оноо, төлөв, шалтгаан, цагийн бүртгэл зэрэг талбарууд
Шошго	Туршилтын нөхцөлөөс хамааран “хэвийн”, “сэжигтэй”, “хориглох” гэж ангилсан
Хэмжээ	Нийт таван туршилтын бүлэг, бүлэг тус бүр 10–20 хандалтын бичлэг
Ашигласан зорилго	Эрсдэлийн үнэлгээний загварын гүйцэтгэл болон шийдвэрийн босго утгуудын үр нөлөөг үнэлэх

Туршилтын нөхцлүүдийг IV-р хүснэгтэд үзүүлэв.

TABLE IV. Туршилтын тодорхойлолт

Тест	Тодорхойлолт
T1	Хэвийн хэрэглэгч
T2	Дэлгэцийн бичлэг асаалттай
T3	Дебаггер ON
T4	Клиент талын зөвшөөрөлгүй өөрчлөлт
T5	Хүчингүй токен

A. Туршилтын өгөгдлийн хязгаарлалтууд.

Туршилтын өгөгдлийн хязгаарлалт 1.

Туршилтын өгөгдөл хязгаарлагдмал буюу 5 нөхцлийг 10-20 удаагийн давтамжтайгаар туршсан. Ингэж хийсэн нь судалгааны үзэл баримтлалыг батлахад хангалттай боловч бодит системд нэвтрүүлэхэд илүү их өгөгдөл шаардлагатай.

Туршилтын өгөгдлийн хязгаарлалт 2.

Манай систем хяналттай орчинд туршилт хийсэн ба бодит стриминг платформ эсвэл бодит хэрэглэгчийн өгөгдөл ашиглах боломжгүй байсан. Цаашид бодит орчинд турших шаардлагатай.

VI. Үр Дүн

Санал болгож буй динамик DRM хамгаалалтын механизмыг үнэлэх зорилгоор таван төрлийн туршилтын нөхцөл боловсруулж, системийн хариу үйлдэл, эрсдэлийн үнэлгээ болон шийдвэр гаргалтын үр дүнг шинжилсэн. Туршилтын нөхцөлүүдэд хэвийн хэрэглэгч, дэлгэцийн бичлэг асаалттай орчин, дебаггер идэвхтэй байдал, хүчингүй токен, мөн клиент талын өөрчлөлт зэрэг тохиолдлуудыг хамруулсан. Эдгээр нөхцөл нь T1–T5 туршилтын бүтэцтэй нийцэж байна.

A. Туршилтын үр дүн

Эрсдэлийн үнэлгээнд ашигласан үзүүлэлтүүд болон тэдгээрийн оноог “TABLE V” дээр үзүүлэв.

TABLE V. Туршилтын Үр Дүн

Тест	Нөхцөл	Илэрсэн үзүүлэлт	Оноо	Шийдвэр
T1	Хэвийн хэрэглэгч	Илрээгүй	0	ALLOW
T2	Дэлгэцийн бичлэг асаалттай	Дэлгэц бичигч	35	SUSPICIOUS
T3	Дебаггер ON	Дебаггер	25	SUSPICIOUS
T4	Клиент талын зөвшөөрөлгүй өөрчлөлт	Хэш зөрүү / клиент талын зөвшөөрөлгүй өөрчлөлт	40	BLOCK
T5	Хүчингүй токен	Токен баталгаажуулалт амжилтгүй	BLOCK	BLOCK

Хэвийн хэрэглэгчийн нөхцөлд систем ямар нэг сэжигтэй үйлдэл илрүүлээгүй бөгөөд эрсдэлийн оноо 0 байж, хандалтыг зөвшөөрсөн. Энэ нь санал болгож буй механизм суурь нөхцөлд зөв ажиллаж байгааг харуулсан. Хэвийн орчинд систем контентыг амжилттай нээх боломж олгосон байна.

Дэлгэцийн бичлэг асаалттай үед систем тухайн үйлдлийг амжилттай илрүүлж, эрсдэлийн оноог 35 болгон тооцоолж, хэрэглэгчийг SUSPICIOUS ангилалд оруулсан. Энэ нь DRM хамгаалалтыг тойрон гарах түгээмэл оролдлогыг илрүүлэх боломжтойг харуулж байна.

Дебаггер идэвхтэй байсан үед систем тухайн үйлдлийг илрүүлсэн боловч дангааpаа хамгийн өндөр эрсдэл гэж үзээгүй тул SUSPICIOUS ангилалд оруулсан. Өөрөөр хэлбэл, систем эрсдэлийн түвшнийг шаталсан байдлаар үнэлж байгааг энэ үр дүн илтгэнэ.

Клиент талын өөрчлөлт илэрсэн үед систем хэш зөрүүг илрүүлж, хандалтыг хориглосон. Энэ нь хамгаалалтыг тойрон гарах оролдлогыг өндөр эрсдэлтэй гэж үнэлэн шууд зогсоож байгааг харуулж байна. Клиент программд өөрчлөлт оруулах үед хэш зөрүү илэрч, хандалтыг шууд хаасан байна.

Харин хүчингүй токен ашигласан үед систем хэрэглэгчийг баталгаажуулж чадаагүй тул хандалтыг шууд хориглосон. Энэ нөхцөл нь баталгаажуулалтын түвшинд шууд таслан зогсоох шаардлагатай хамгийн өндөр эрсдэлтэй тохиолдол болохыг харуулж байна. Хүчингүй токен ашигласан үед систем шууд BLOCK шийдвэр гаргасан.

Эдгээр үр дүнгээс харахад санал болгож буй механизм нь хэрэглэгчийн орчны нөхцөл болон үйл ажиллагааны шинжүүдэд үндэслэн эрсдэлийг ялгамжтай үнэлж, түүнд тохирсон хандалтын шийдвэр гаргаж байгааг харуулж байна. Ялангуяа өндөр эрсдэлтэй нөхцөлүүдэд хандалтыг шууд хориглох, харин дунд эрсдэлтэй нөхцөлүүдэд сэжигтэй гэж ангилан хязгаарлах боломжтой байгааг туршилтаар баталсан.

B. Уламжлалт DRM системтэй харьцуулсан үзүүлэлтүүд

Уламжлалт DRM систем нь статик шифрлэлт, лиценз сервэр хослуулан ажилладаг. Манай санал болгож буй систем нь динамик эрсдэлийн оноо, зан төлөвт суурилсан шийдвэр нэмэж уламжлалт DRM системийг сайжруулсан.

TABLE VI. DRM СИСТЕМҮҮДИЙГ ХАРЬЦУУЛСАН ҮЗҮҮЛЭЛТҮҮД

Үзүүлэлт	**Уламжлалт DRM (Widevine, PlayReady, FairPlay)**	**Санал болгож буй DRM систем**
Хамгаалалтын зарчим	Статик шифрлэлт + лиценз сервер	Динамик эрсдэлийн оноо + зан төлөвт суурилсан шийдвэр
Дэлгэц бичлэг илрүүлэх	Хязгаарлагдмал (зөвхөн тодорхой төхөөрөмж дээр)	Илрүүлсэн (T2 тест)
Debugger илрүүлэх	Илрүүлдэггүй	Илрүүлсэн (T3 тест)
Client tampering илрүүлэх	Илрүүлдэггүй	Илрүүлсэн (T4 тест)
Invalid token илрүүлэх	Токен шалгадаг	Илрүүлсэн (T5 тест)
Шийдвэрийн түвшин	ALLOW / DENY (2 түвшин)	ALLOW / SUSPICIOUS / BLOCK (3 түвшин)
Нууцлалд нөлөөлөх	Төхөөрөмжийн мэдээлэл цуглуулдаг	Зөвхөн орчны мэдээлэл (recorder, debugger гэх мэт)
Detection rate (бидний тестээр)	Мэдэгдээгүй (хаалттай систем)	100% (T2–T5 бүх сэжигтэй үйлдэл илэрсэн)
False positive (бидний тестээр)	Мэдэгдээгүй	0% (T1 хэвийн хэрэглэгч зөвшөөрөгдсөн)

VII. ДҮГНЭЛТ

Энэхүү судалгааны ажлаар уламжлалт DRM хамгаалалтын хязгаарлалтыг даван туулах зорилгоор хэрэглэгчийн орчны мэдээлэлд суурилсан динамик DRM хамгаалалтын механизмыг санал болголоо. Уг механизм нь дэлгэц бичигч, дебаггер, клиент талын зөвшөөрөлгүй өөрчлөлт, хүчингүй токен, төхөөрөмжийн өөрчлөлт зэрэг үзүүлэлтүүдийг нэгтгэн үнэлж, ALLOW, SUSPICIOUS, BLOCK гэсэн гурван түвшний шийдвэр гаргах зарчмаар хэрэгжсэн. Энэ бүтэц нь санал болгосон механизм, эрсдэлийн онооны загвар, шийдвэрийн босготой нийцэж байна.

Туршилтын үр дүнгээс харахад санал болгож буй арга нь хэвийн хэрэглэгчдэд саад учруулахгүйгээр ажиллахын зэрэгцээ сэжигтэй болон хориглох шаардлагатай нөхцөлүүдийг амжилттай илрүүлж чадсан. Ялангуяа клиент талын зөвшөөрөлгүй өөрчлөлт болон хүчингүй токен зэрэг өндөр эрсдэлтэй нөхцөлд шууд BLOCK, харин дэлгэц бичигч болон дебаггер зэрэг нөхцөлд SUSPICIOUS ангилал хэрэглэсэн нь системийн уян хатан, шаталсан хамгаалалтын давуу талыг харуулж байна.

Иймээс энэхүү ажил нь DRM хамгаалалтыг зөвхөн статик түлхүүр болон лицензийн хяналтаар хязгаарлах бус, хэрэглэгчийн орчныг үнэлж динамик шийдвэр гаргадаг илүү уян хатан аргачлал руу өргөтгөх боломжтойг харуулж байна. Цаашид уг механизмыг илүү өргөн хэмжээний бодит хэрэглээний өгөгдөл дээр турших, эрсдэлийн үзүүлэлтүүдийн жинг оновчлох, мөн машин сургалтын аргатай хослуулах замаар илүү боловсронгуй болгох боломжтой.

Тухайн санал болгож буй систем нь дэлгэц бичигч, дебаггер, клиент талын зөвшөөрөлгүй өөрчлөлт, хүчингүй токен, төхөөрөмжийн өөрчлөлтийг илрүүлдэг. Санах ойн засварчлал, API барих (API hooking), виртуалчлагдсан орчин дахь халдлага зэрэг орчин үеийн нарийн төвөгтэй халдлагуудыг илрүүлэхгүй.

Илрүүлэх боломжгүй халдлагуудын төрлүүд болон санал болгож буй сайжруулалтыг VII-р хүснэгтэд харууллаа.

TABLE VII. АЮУЛГҮЙ БАЙДЛЫН НАРИЙВЧИЛСАН АНАЛИЗ

Халдлагын төрөл	**Илрүүлэхгүй шалтгаан**	**Сайжруулалтын санал**
Memory patching	Клиент талд memory access шалгадаггүй	Анти-тампер библиотек ашиглах
API hooking	Системийн дуудлагыг хянадаггүй	Хэрэглэгчихэд хүнд
Virtual machine / sandbox	VM илрүүлэх логик байхгүй	VM илрүүлэлт нэмэх
Network replay attack	Token хугацаа дуустал дахин	Nonce механизм нэмэх

Халдлагын төрөл	**Илрүүлэхгүй шалтгаан**	**Сайжруулалтын санал**
	илгээж болно	
Hardware-level recording	Программын түвшинд илрүүлэх боломжгүй	Forensic watermarking нэмэх

Цаашид уг механизмыг илүү өргөн хэмжээний бодит хэрэглээний өгөгдөл дээр турших, эрсдэлийн үзүүлэлтүүдийн жинг оновчлох, мөн машин сургалтын аргатай хослуулах замаар илүү боловсронгуй болгох боломжтой.

Ашигласан Материал, Ном Зүй

[1] D'Souza, J. (2025, September 24). Piracy statistics, trends and facts (2025). ElectroIQ.

[2] S. R. Subramanya and B. K. Yi, "Digital rights management," *IEEE Potentials*, vol. 25, no. 2, pp. 31–34, 2006.

[3] M. Dworkin, *Advanced encryption standard (AES)*, NIST FIPS 197, 2001.

[4] Cincopa, "Digital rights management: Widevine, PlayReady and FairPlay explained," 2024.

[5] IETF, *OAuth 2.0 token introspection*, RFC 7662, 2015.

[6] OWASP, *MASVS / MASTG: Anti-debugging and resilience against reverse engineering and tampering*, 2023.

[7] J. Poort, J. P. Quintais, M. A. van der Ende, A. Yagafarova, and M. Hageraats, "Global online piracy study," *Amsterdam Law School Research Paper*, no. 2018-21, 2018.

Цахим эх сурвалжууд

[8] https://en.wikipedia.org/wiki/Digital_rights_management

[9] https://electroiq.com/stats/piracy-statistics/

[10] https://doverunner.com/blogs/widevine-playready-fairplay-drm-comparison/

[11] https://medium.com/@technedigitale/drm-101-introduction-to-drm-systems-66ba88f74244

Квант компьютерын сорилтыг сөрөн тэсвэрлэх хосолсон алгоритмыг ашиглах холболт болон өгөгдөл нууцлалд хэрэгжүүлэх судалгаа

Б. Нямдэлгэр[1], М. Мөнхбагана[1], Х. Уянгаа[1*], Г. Ууганчимэг[2], Я. Дашдорж[1], Л. Баяр-Эрдэнэ[3]
[1]*Кибер аюулгүй байдлын тэнхим*, [2]*Электроникийн тэнхим*, [3]*Захиргаа*
ШУТИС, МХТС
Монгол улс, Улаанбаатар хот
{nyamdelgernaymka, baganaa237}@gmail.com, [*]uyangaakh@must.edu.mn

Хураангуй **— Квант компьютерын хурдацтай хөгжил нь RSA, ECC зэрэг өнөөгийн нийтийн түлхүүрт криптографт ноцтой эрсдэл учруулж байна. Шорын алгоритм нь квант орчинд бүхэл тоог үржигдэхүүнд задлах болон дискрет логарифмын асуудлыг богино хугацаанд шийдвэрлэх боломж олгодог тул асимметрик криптографийн аюулгүй байдал алдагдах эрсдэлтэй. Үүнийг бууруулах зорилгоор квант компьютерт тэсвэртэй алгоритмууд хурдацтай хөгжиж, NIST 2024 онд FIPS 203 стандартаар ML-KEM түлхүүр солилцооны алгоритмыг баталсан. Энэхүү судалгаанд хост компьютер дээр өгөгдөл нууцлах туршилт хийж, TLS 1.3 протокол дээр x25519 болон x25519:ML-KEM hybrid алгоритмуудын гүйцэтгэлийг харьцуулсан. Туршилтын үр дүнгээс hybrid TLS нь өгөгдөл нууцлал болон түлхүүр солилцоонд квантын эсрэг хамгаалалтыг нэмэгдүүлдэг ч сүлжээний ачаалал болон нөөцийн хэрэглээг ихэсгэдэг. Бодит орчинд өргөн нэвтрүүлэхийн тулд илүү үр ашигтай, нөөцийн хэрэглээг бууруулсан шийдэл боловсруулах шаардлагатай байна.**

Түлхүүр үг — PQC, ML-KEM алгоритм, квантын дараах үеийн криптограф, сүлжээний аюулгүй байдал, пакет шинжилгээ

I. УДИРТГАЛ

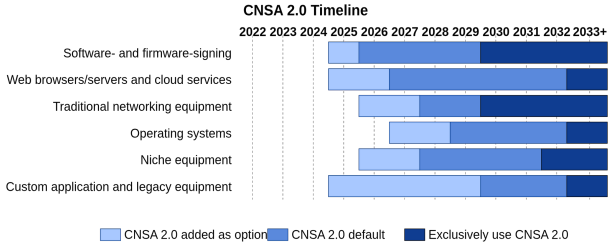
Криптограф нь мэдээллийн нууцлал, бүрэн бүтэн байдал, баталгаажуулалт зэрэг аюулгүй байдлын чухал шаардлагуудыг хангаж, цахим орчны найдвартай ажиллагааг дэмждэг. Банк, цахим худалдаа, харилцаа холбоо, аж үйлдвэрийн автоматжуулалт зэрэг олон чухал салбарт өргөнөөр ашиглагддаг бөгөөд мэдээллийн технологи болон кибер аюулгүй байдал дахь үндсэн технологийн нэг юм. 2015 онд Үндэсний аюулгүй байдлын агентлаг (NSA) нь CNSA 1.0 криптограф стандартыг Үндэсний аюулгүй байдлын системүүд (NSS)-ийг хамгаалах зорилгоор баталсан. АНУ-ын Үндэсний Стандарт ба Технологийн Хүрээлэн (NIST)-ийн тодорхойлсноор, энэхүү стандарт нь квант тооцооллоос хамгаалагдсан шинэ алгоритмууд боловсруулагдаж, батлагдах хүртэл ашиглагдах болно [1].

NIST нь мэдээллийн аюулгүй байдалд заналхийлж буй шинэ төрлийн аюулыг таслан зогсоохын тулд 2016 оноос эхлэн квант компьютерын халдлагаас хамгаалах шинэ нийтийн түлхүүрт криптографийн стандарт боловсруулах ажлыг эхлүүлэн нээлттэй сонгон шалгаруулалт явуулж эхэлсэн [2].

Сонгон шалгаруулалт дараах үе шатуудаас бүрдэнэ:

- Анхны шалгаруулалт (2016–2017): Шинэ алгоритмуудын саналыг хүлээн авч эхэлсэн.
- 1-р шат (2017–2019): Олон улсын Криптографийн судлаачид нийт 69 алгоритм санал болгосон.
- 2-р шат (2019–2020): 26 алгоритм сонгогдож, аюулгүй байдал болон гүйцэтгэлийн үнэлгээ хийгдсэн.
- 3-р шат (2020–2022): Сүүлийн шатанд 7 алгоритм шалгарсан.
- Албан ёсны сонголт (2022–2025): Олон улсын стандарт руу шилжүүлэх ажил хийгдсэн.

NSA нь 2024 оны 12-р сард NIST-ээс стандартчилахаар сонгосон алгоритмуудаас бүрдсэн CNSA 2.0 стандартыг баталсан. Үүнтэй холбоотойгоор, Зураг 1-т заасан хугацаанд Suite B болон CNSA 1.0 стандартын алгоритмуудыг ашиглаж буй бүх системүүд CNSA 2.0 уруу шилжих шаардлагатайг тодотгосон [3].



Зураг 1. CNSA 2.0 технологиудын шилжилтийн хугацааны төлөвлөгөө, оноор [3].

Зураг 1-т харуулснаар CNSA 2.0 криптографийн стандарт руу ялгаатай системүүд 2022–2033 оны хооронд шат дараатайгаар шилжих төлөвлөгөөг харуулж байна. Зарим системүүд хурдан шинэчлэгдэх бол хуучин болон

тусгай төхөөрөмжүүдийн шилжилтэд багагүй хугацаа зарцуулах шаардлагатайг харуулж байна.

II. АРГА ЗҮЙ БА ОНОЛЫН ҮНДЭС

A. Квант компьютер

Квант компьютер нь бөөмсийн харилцан үйлчлэл, квант механикийн зарчмуудыг ашиглан мэдээлэл боловсруулах өндөр хүчин чадалтай тооцоолох төхөөрөмж юм. Энэ нь физикийн онолын нэг салбар болох квант механикт тулгуурладаг [4]. Уламжлалт компьютероос ялгаатай нь мэдээллийн нэгж болох битийн оронд кубитийг ашигладаг. Кубит нь нэгэн зэрэг 0 болон 1-ийн төлөвт орших боломжтой ба суперпозици (Superposition), квант ороолцолдоо (Entanglement), квант интерференц (Quantum Interference), квант хэмжилт (Quantum Measurement) зэрэг квант механикийн шинж чанаруудыг ашиглан маш хурдан, үр ашигтай тооцоолол хийх боломжийг олгодог.

Квант компьютерын гол онцлогууд:

- Суперпозици – Кубитууд нь superposition зарчмаар ажилладаг бөгөөд нэгэн зэрэг 0 ба 1-ийн төлөвт байж болно.
- Ороолцолдоо – Хоёр ба түүнээс дээш кубитууд хоорондоо холбогдож, нэг кубитын төлөв нь нөгөө кубитын төлөвөөс шууд хамааралтай болох үзэгдэл юм.
- Квант интерференц – Кубитуудын төлөвүүд хоорондоо харилцан үйлчлэлцэж, зөвхөн зөв хариултуудыг нэмэгдүүлж, буруу хариултуудыг багасгадаг.
- Квант хэмжилт – Кубитийн төлөвийг классик 0 эсвэл 1 төлөв болгон хөрвүүлэх явц [4].

Квант компьютерын хөгжил нь криптографийн стандарт алгоритмуудад аюул учруулж эхэлсэн ба 2035 он гэхэд RSA-2048 шифрлэлтийн алгоритмыг бүрэн эвдэлнэ гэж урьдчилан таамаглаж байна [4].

B. Шорын алгоритм

1994 онд Белл Лабс-ын судлаач Питер Шор квант компьютер дээр олон оронтой тоог олон гишүүнт хугацаанд задлах алгоритмыг боловсруулсан ба Шорын алгоритм (Shor's Algorithm) хэмээн нэрлэсэн. Уламжлалт компьютер олон оронтой тоог (жишээлбэл, 2048 битийн урттай тоо) олон гишүүнт хугацаанд задалж чадахгүй. Одоогоор хамгийн хурдан нийтийн алгоритм нь NFS (Number Field Sieve) ба экспоненциал хугацаанд ажилладаг. Шорын алгоритмын үндсэн санаа нь квант Фурьегийн хувиргалт (Quantum Fourier Transform, QFT) ашиглан үе олох процесс юм [8]. Шорын алгоритм нь нийт 11 алхамтай ба үндсэндээ QFT ашиглан r болон x тоог дараах томьёогоор тооцоолно:

$$\frac{1}{\sqrt{|A|}} \sum_{a' \in A} \sum_{c=0}^{q-1} |c, k\rangle \cdot e^{2\pi i \, a^c c/q} \tag{1}$$

Сүүлчийн алхам нь дараах томьёогоор хийгддэг бөгөөд үүний үр дүнд том тооны анхны хуваагчийг олох боломжтой. Хамгийн их ерөнхий хуваагч (Greatest Common Divisor) тооцоолол нь уламжлалт компьютер дээр маш хурдан хийгддэг [8].

$$\gcd\Big(x^{r/2}+1,\, n\Big) \text{ and } \gcd\Big(x^{r/2}-1,\, n\Big) \tag{2}$$

2015 онд батлагдсан CNSA 1.0 стандартад ашиглагдсан криптографийн алгоритмуудыг Хүснэгт I-т харуулав. Энэхүү стандарт нь асимметрик болон симметрик криптографийн хамгийн найдвартай аюулгүй байдлыг хангах зорилготой [1].

Хүснэгт I. CNSA 1.0 буюу өнөөгийн бидний өргөн ашигладаг стандарт

Алгоритм	**Функц**	**Стандарт**	**Параметр**
Advanced Encryption Standard (AES)	Блок шифрлэлт	FIPS Pub 197 [12]	256 битийн түлхүүр
Elliptic Curve Diffie-Hellman (ECDH) Key Exchange	Түлхүүр солилцох нийтийн түлхүүрийн алгоритм	NIST SP 800-56A [13] [14]	P-384 муруй
Elliptical Curve Digital Signature Algorithm (ECDSA)	Цахим гарын үсгийн нийтийн түлхүүрийн алгоритм	FIPS Pub 186-4 [14] [15]	P-384 муруй
Secure Hash Algorithm (SHA)	Хэш алгоритм	FIPS Pub 180-4 [16]	SHA-384 хэш
Diffie-Hellman (DH) түлхүүр солилцоо	Түлхүүр солилцооны алгоритм	IETF RFC 3526 [17]	Хамгийн багадаа 3072-битийн модуль
Rivest-Shamir Adleman (RSA)	Түлхүүр үүсгэх алгоритм	NIST SP 800-56B rev 1 [18]	Хамгийн багадаа 3072-битийн модуль
Rivest-Shamir Adleman (RSA)	Цахим гарын үсэгт хэрэглэгддэг нийтийн түлхүүрийн алгоритм	FIPS Pub 186-4 [15]	Хамгийн багадаа 3072-битийн модуль

Асимметрик криптографийн хувьд:

- CNSA 1.0 стандарт нь RSA болон ECC алгоритмуудыг ашигладаг.
- Эдгээр алгоритмууд анхны тооны задрал болон дискрет логарифмын математикийн бодлогод тулгуурладаг.

Квант компьютерын Шорын алгоритм нь RSA болон ECC-г хурдан хугацаанд эвдэх чадвартай учраас эдгээр алгоритмууд найдвартай бус гэж тооцогдох болсон. Иймд, 2024 оны 12 сард CNSA 2.0 стандартад квант компьютерын дараах криптограф (PQC) буюу квантын халдлагад тэсвэртэй алгоритмууд багтааж, аюулгүй байдлын шинэ шаардлагад нийцүүлэн сайжруулсан [10].

Энэхүү судалгааны зорилго нь TLS 1.3 протоколд квантын дараах үеийн криптографийн алгоритм болох ML-KEM-ийг ашиглах үед сүлжээний холболтын гүйцэтгэлд ямар нөлөө үзүүлэхийг туршилтаар тодорхойлох явдал

Хүснэгт II. CNSA 2.0 буюу шинээр батлагдсан квант компьютерт тэсвэртэй стандарт

Алгоритм	Функц	Стандарт	Параметр
Advanced Encryption Standard (AES)	Симметрик блок шифрлэлтийн алгоритм	FIPS PUB 197 [19]	256 бит түлхүүр
ML-KEM (CRYSTALS-Kyber)	Түлхүүр солилцох асимметрик алгоритм	FIPS PUB 203 [20]	ML-KEM-1024
ML-DSA (CRYSTALS-Dilithium)	Цахим гарын үсгийн асимметрик алгоритм	FIPS PUB 204 [21]	ML-DSA-87
Secure Hash Algorithm (SHA)	Хэш алгоритм	FIPS PUB 180-4 [22]	SHA-384, SHA-512
Leighton-Micali Signature (LMS)	Цахим гарын үсгийн асимметрик алгоритм	NIST SP 800-208 [23]	LMS SHA-256/192
Xtended Merkle Signature Scheme (XMSS)	Цахим гарын үсгийн асимметрик алгоритм	NIST SP 800-208 [24]	–
Secure Hash Algorithm 3 (SHA3)	Хэш алгоритм	FIPS PUB 202 [25]	SHA3-388, SHA3-512

юм. Үүний хүрээнд уламжлалт TLS 1.3 холболтыг ML-KEM алгоритмын өөр өөр параметрийн багцтай харьцуулан судалсан. ML-KEM (Module-Lattice-Based Key-Encapsulation Mechanism) нь NIST-ээс FIPS 203 стандартаар батлагдсан, модуль-лэттис суурьтай, квант компьютерт тэсвэртэй алгоритм юм. Уг алгоритм нь өмнө нь CRYSTALS-Kyber нэрээр танигдаж байсан бөгөөд стандартчиллын эцсийн шатанд техникийн зарим өөрчлөлт орсны үндсэн дээр ML-KEM нэршилтэй болсон. Түүний аюулгүй байдал нь Module Learning With Errors (Module-LWE) зэрэг тооцооллын хувьд хүндрэлтэй математик бодлогууд дээр тулгуурладаг тул уламжлалт болон квант халдагчийн аль алинд тэсвэртэй гэж үздэг.

III. ТУРШИЛТЫН ОРЧИН

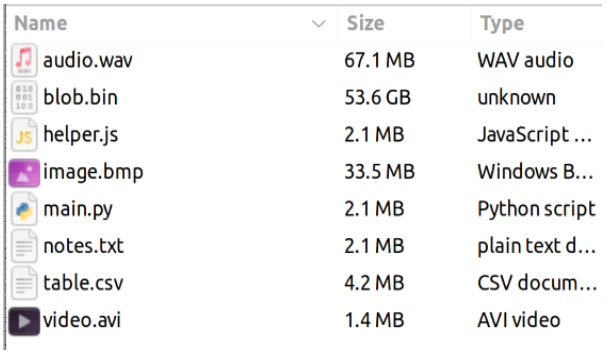
A. Хост компьютер дээрх өгөгдөл нууцлах туршилт

Энэхүү туршилт нь хост компьютер дээрх өгөгдлийг уламжлалт болон квант компьютерын дараах криптографийн алгоритмуудыг хослуулан ашиглан мэдээллийг өндөр түвшний нууцлалтай болгоход шаардлагатай түлхүүр үүсгэх, өгөгдөл нууцлах, тайлах хугацаа болон нууцалсан файлын хэмжээний өсөлт зэрэг хүчин зүйлсийг судалсан.

Туршилтын орчин:

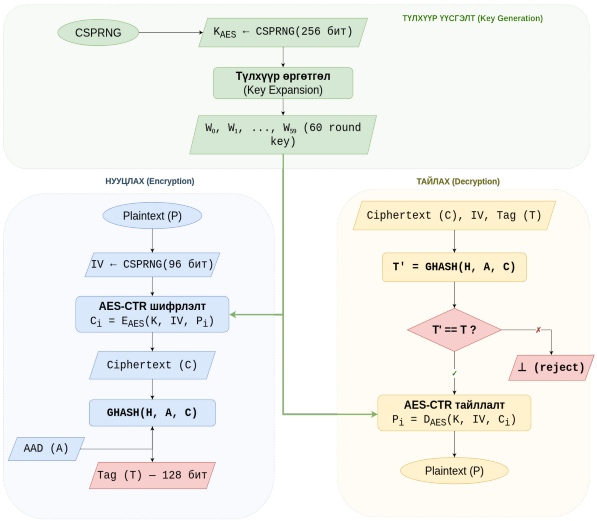
- Үйлдлийн систем: Ubuntu 22.04 LTS
- Ашигласан сан: OpenSSL 3.5.5
- Урьдчилан үүсгэсэн файлууд

Зураг 2-т харуулснаар туршилтын хүрээнд 1 МБ, 10 МБ, 100 МБ, 1 ГБ гэсэн хэмжээтэй файлуудыг урьдчилан бэлтгэн хэрэглэв.



Зураг 2. Туршилтад ашиглаж буй файлын агуулах мэдээллүүд.

Уламжлалт алгоритмаар The Advanced Encryption Standard (AES)-ийн GCM (Galois/Counter Mode) төрлийг сонгож ашигласан бөгөөд энэхүү алгоритм нь одоогийн стандарт өгөгдөл нууцлалтад хамгийн түгээмэл ашиглагддаг.

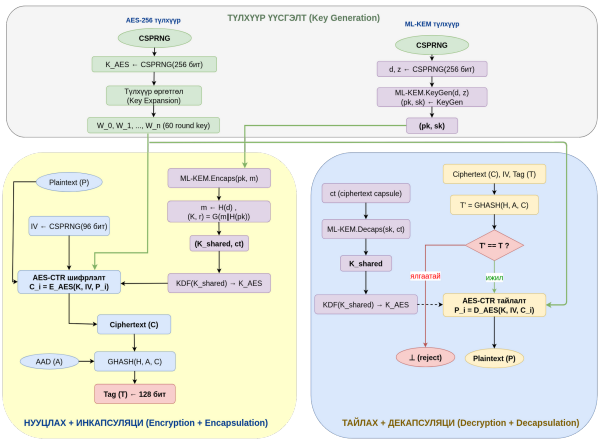


Зураг 3. AES алгоритмын ажиллагаа.

Зураг 3-т мэдээллийг нууцлах болон зөв эсэхийг шалгаж, тайлах үйл явцыг харуулж байна. Эхлээд нууц түлхүүр үүсгээд, түүгээр мэдээллийг шифрлэнэ. Дараа нь мэдээлэл өөрчлөгдсөн эсэхийг шалгах тусгай tag үүсгэнэ. Тайлах үед tag зөв байвал мэдээллийг буцааж уншина, буруу байвал шууд татгалзана.

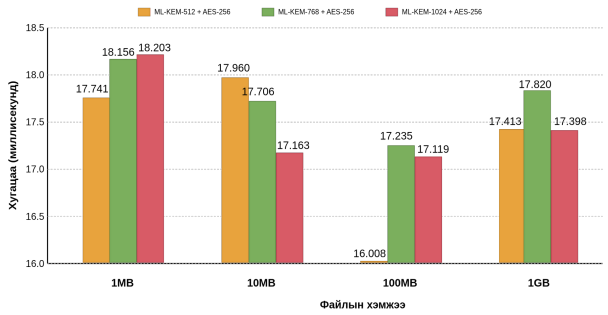
Квант компьютерын дараах үеийн алгоритмаар ML-KEM 512, 768, 1024 параметрүүдийн утгууд дээр тус тус хослуулан гүйцэтгэн туршилтад ашигласан.

Зураг 4-т AES-256+ML-KEM уламжлалт ба шинэ үеийн алгоритмуудыг хослуулан ажиллах зарчмыг харуулав. Эхлээд ML-KEM ашиглан хоёр талд ижил нууц түлхүүрийг аюулгүйгээр тохируулна. Дараа нь тэр түлхүүрээс AES түлхүүр гаргаж аваад, жинхэнэ мэдээллийг AES-CTR аргаар шифрлэж ciphertext болгоно.



Зураг 4. Хослуулсан алгоритмын ажиллагаа.

Мөн мэдээлэл замдаа өөрчлөгдсөн эсэхийг шалгахын тулд “tag” үүсгэнэ. Хүлээн авагч тал ML-KEM-ээр ижил түлхүүрээ дахин гаргаж аваад, эхлээд “tag”-ыг шалгана. Хэрэв зөв байвал шифрлэгдсэн текст (ciphertext)-ийг тайлж эх текст (plaintext)-ийг сэргээнэ, буруу байвал татгалзана.



Зураг 5. ML-KEM алгоритмуудын параметрүүдийн түлхүүр үүсгэх хугацаа.

Зураг 5-д туршилтын үеийн ML-KEM хувилбаруудын түлхүүр үүсгэх хугацааг харьцуулсан ба ойролцоогоор 16–18 мс хугацаанд ажиллаж байна. Үүнээс ML-KEM-512 + AES-256 арга нь хамгийн хурдан, харин ML-KEM-768 ба ML-KEM-1024 нь өндөр хугацаа үзүүлсэн ч зөрүү нь их биш байна. Мөн файлын хэмжээ өөрчлөгдөхөд хугацаа тогтмол өсөөгүй нь түлхүүр үүсгэх үйлдэл файлын хэмжээнээс шууд хамааралгүйг харуулж байна.

Хосолсон аргуудын хувьд файлын хэмжээ өсөхийн хэрээр нууцлах болон тайлах хугацаа нэмэгдэж байна. Бүх тохиолдолд тайлах хугацаа нь нууцлах хугацаанаас бага байсан нь ажиглагдсан. AES-256 дангаараа хамгийн хурдан ажилласан бол ML-KEM + AES-256 хосолсон хувилбарууд харьцангуй өндөр хугацаа үзүүлсэн.

Туршилтын дүнгээс харахад ML-KEM + AES-256 хосолсон арга нь аюулгүй байдлыг нэмэгдүүлэхийн зэрэгцээ хурдны алдагдал их биш байна. AES-256 дангаараа хамгийн хурдан ажилласан ч хосолсон хувилбарууд ч практикт ашиглахад хангалттай хурдан байсан. Мөн ML-KEM-512, 768, 1024 хосолсон аргуудын хоорондын ялгаа бага байсан тул хосолсон хамгаалалт нь үр ашигтай шийдэл гэж дүгнэж байна.

Хүснэгт III. Туршилтын тоон мэдээлэл

Туршилтын гүйцэтгэлийн хэмжилтийн метрикүүд	**Туршилтын дундаж гүйцэтгэлийн харьцуулалт**			
	AES-256-GCM	***ML-KEM 512+AES -256-GCM***	***ML-KEM 768+AES -256-GCM***	***ML-KEM 1024+AES -256-GCM***
Шифрлэсний дараах дундаж хэмжээ (МБ)	283.75	283.75	283.75	283.75
ML-KEM key generation (мс)	–	17.281	17.729	17.471
ML-KEM encap-sulation (мс)	–	7.226	8.246	7.690
ML-KEM decap-sulation (мс)	–	5.929	6.236	5.839
AES encryption (мс)	902.472	903.567	906.094	906.075
AES decryption (мс)	656.218	654.153	657.290	655.834
encryption (мс)	928.767	966.267	970.125	967.916
decryption (мс)	656.218	675.517	680.087	677.160

B. TLS 1.3 handshake-ийн гүйцэтгэлийн туршилт

Судалгаанд TLS 1.3 протоколын дараах тохиргоонуудыг ашигласан:

- x25519 (суурь хувилбар)
- x25519 + ML-KEM512
- x25519 + ML-KEM768
- x25519 + ML-KEM1024

Гүйцэтгэлийг дараах үзүүлэлтүүдээр хэмжсэн:

- Handshake latency (ms)
- Throughput (connection/sec)
- CPU хэрэглээ (%)
- RAM хэрэглээ (MB)
- Packet size (bytes)

Мөн single handshake болон 1000 удаагийн давталтын туршилт хийж системийн тогтвортой байдлыг үнэлсэн.

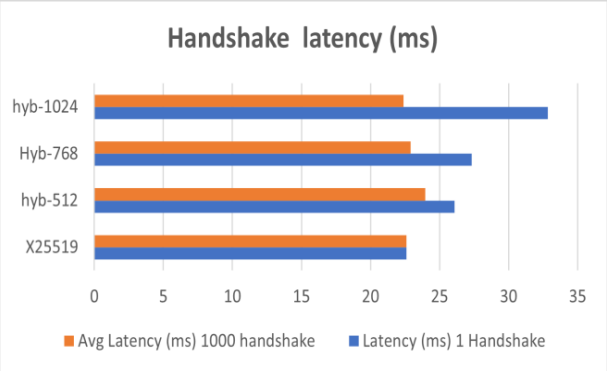
Хүснэгт IV. TLS 1.3 болон post-quantum key exchange ашиглан 1 удаагийн холболтын харьцуулсан үзүүлэлтүүд

Туршилтын гүйцэтгэлийн хэмжилтийн метрикүүд	**TLS 1.3 + post-quantum key exchange 1 handshake**			
	TLS 1.3 (X25519)	***TLS 1.3 + Hyb-512***	***TLS 1.3 + Hyb-768***	***TLS 1.3 + Hyb-1024***
Packet size (МБ)	0.83	5.08	10.01	14.60
Latency (мс)	22.592	26.081	27.317	32.317
RAM (МБ)	1373.95	1380.93	2816.62	3624.61

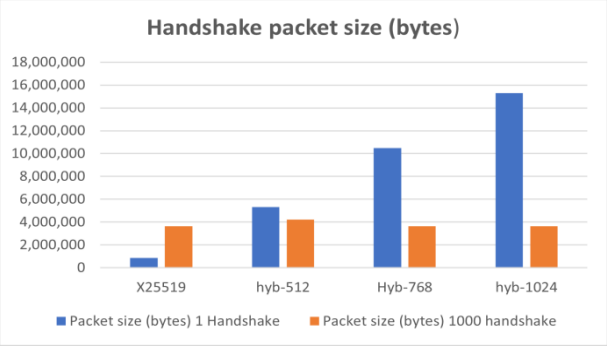
Энэхүү туршилт нь TLS 1.3 протокол дээрх x25519 (class) болон x25519:ML-KEM (hybrid) алгоритмуудын гүйцэтгэлийг харьцуулсан.

Хүснэгт V. TLS 1.3 болон post-quantum key exchange ашиглан 1000 удаагийн холболтын харьцуулсан үзүүлэлтүүд

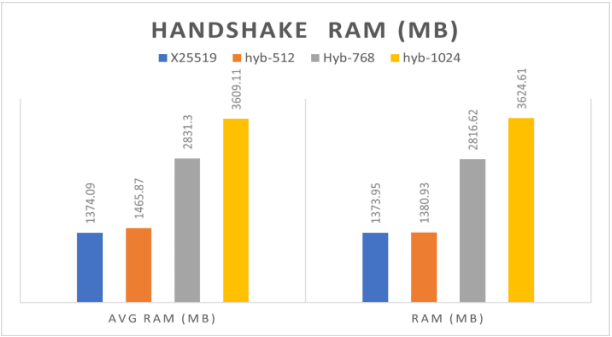
Туршилтын гүйцэтгэлийн хэмжилтийн метрикүүд	**TLS 1.3 + post-quantum key exchange 1000 handshake**			
	TLS 1.3 (X25519)	***TLS 1.3 + Hyb-512***	***TLS 1.3 + Hyb-768***	***TLS 1.3 + Hyb-1024***
Avg Latency (мс)	22.588	23.956	22.909	22.378
Throughput (conn/sec)	13.12	12.9	13.06	13.16
Packet size (МБ)	3.46	4.01	3.45	0.35
Avg CPU (%)	0.81	2.40	5.95	6.69
Avg RAM (МБ)	1374.09	1465.87	2831.30	3609.11
Total time (сек)	76.2476	77.532	76.5827	76.0027



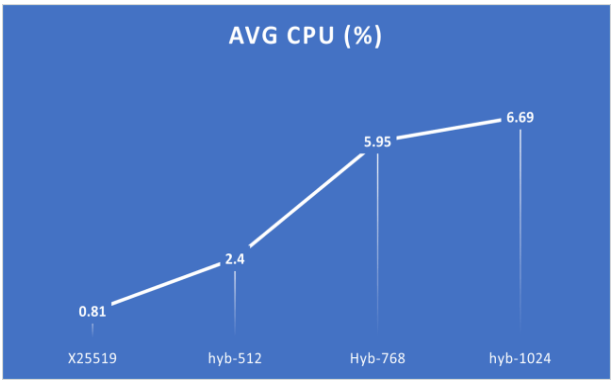
Зураг 6. Нэг удаагийн ба 1000 удаагийн TLS 1.3 handshake-ийн latency үр дүн.



Зураг 7. TLS 1.3 handshake-ийн гүйцэтгэлийн хэмжилт: нэг удаагийн болон 1000 удаагийн handshake-ийн packet size харьцуулсан үр дүн (МБ-аар).



Зураг 8. TLS 1.3 протоколын X25519 болон X25519 + PQC алгоритмын handshake-ийн RAM хэрэглээний харьцуулалт.



Зураг 9. TLS 1.3 протоколын 1000 удаагийн handshake-ийн CPU хэрэглээний харьцуулалт.

Нэг удаагийн handshake-ийн үр дүнгээс харахад hybrid алгоритмуудын latency (22.59мс-ээс 32.84мс) болон packet size (0.87 МБ-ээс 15.3 МБ) болж мэдэгдэхүйц өссөн. Мөн RAM хэрэглээ ML-KEM1024 дээр ~2.6 дахин нэмэгдсэн байна.

Харин 1000 удаагийн handshake-ийн дундаж үзүүлэлтүүдэд latency (~22–24 мс) болон throughput (~13 conn/sec) бараг өөрчлөгдөөгүй боловч CPU (0.81%-иас 6.69%) болон RAM хэрэглээ hybrid алгоритмууд дээр тус тус өссөн үзүүлэлт гарсан байна.

Иймээс hybrid TLS нь квантын эсрэг хамгаалалтыг нэмэгдүүлдэг ч сүлжээний ачаалал болон нөөцийн хэрэглээг ихэсгэдэг гэсэн дүгнэлтэд хүрч байна. Цаашид бодит орчинд хэрэглээнд нэвтрүүлэхийн тулд нөөцийн хэрэглээг бууруулах илүү үр ашигтай шийдэл боловсруулах шаардлагатай.

IV. ҮР ДҮН БА ХЭЛЭЛЦҮҮЛЭГ

Энэхүү судалгаанд практик орчинд квантын дараах криптографийн 2 өөр туршилтыг гүйцэтгэж үр дүнг үнэлсэн.

Хост компьютер дээрх өгөгдөл нууцлалтын туршилтаас харахад ML-KEM + AES-256-GCM хосолсон арга нь AES-256-GCM дангаараа ашиглахтай харьцуулахад нэмэлт

∼30–40 мс overhead үүсгэж байна. Энэ нэмэлт хугацаа нь голчлон ML-KEM-ийн түлхүүр үүсгэх (keygen ∼17 мс), капсулчлах (encapsulation ∼7–8 мс), задлах (decapsulation ∼6 мс) үйлдлүүдээс шалтгаалж байна. Харин AES шифрлэлт, тайлалтын хугацаа нь бүх хувилбарт бараг ижил байсан тул AES-ийн гүйцэтгэлд ML-KEM нөлөөлөөгүй. ML-KEM-512, 768, 1024 хувилбаруудын хоорондын гүйцэтгэлийн ялгаа маш бага (< 3 мс) байсан нь аюулгүй байдлын хамгийн өндөр түвшин болох ML-KEM-1024-ийг ашиглахад гүйцэтгэлийн алдагдалгүйгээр ашиглах боломжтойг харуулж байна.

TLS 1.3 handshake-ийн туршилтаас харахад нэг удаагийн handshake-д hybrid алгоритмуудын latency 22.59 мс-ээс 32.32 мс хүртэл (∼43%) өссөн бол packet size 0.87 МБ-аас 15.3 МБ хүртэл (∼17.5 дахин) нэмэгдсэн. RAM хэрэглээ ML-KEM-1024 дээр ∼2.6 дахин өсч 3624 МБ болсон. Гэвч 1000 удаагийн давтагдсан handshake-ийн дүнд дундаж latency (∼22–24 мс) болон throughput (∼13 холболт/сек) бараг өөрчлөгдөөгүй нь системийн тогтвортой нөхцөлд hybrid аргын гүйцэтгэл хүлээн зөвшөөрөгдөхүйц түвшинд байгааг харуулж байна. Харин CPU хэрэглээ 0.81%-аас 6.69% хүртэл, RAM хэрэглээ тогтмол өндөр хэвээр байсан нь нөөцийн шаардлага ихэсч буйг баталж байна.

Halak нар [9] эмбэддэд төхөөрөмж дээр PQC алгоритмуудын гүйцэтгэлийг судалсан бол бидний судалгаа нь хост компьютер дээрх TLS түвшний гүйцэтгэлд анхаарлаа хандуулсан. Хоёр судалгааны үр дүн нэгэн адил hybrid арга нь бодитой overhead нэмдэг ч практикт хэрэглэх боломжтой гэсэн дүгнэлтэд хүрч байна.

Судалгааны хязгаарлалт нь дотоод сүлжээний localhost орчинд туршилт хийсэн тул бодит интернэт орчны latency, packet loss зэрэг хүчин зүйлсийг харгалзаагүй юм. Мөн зөвхөн нэг хост компьютер дээр туршилт явуулсан тул олон тооны зэрэг холболт үүсэх үеийн нөлөөг үнэлээгүй.

V. ДҮГНЭЛТ

Энэхүү судалгаагаар квант компьютерт тэсвэртэй ML-KEM алгоритмыг AES-256-GCM болон TLS 1.3 протоколтой хослуулан ашиглах боломжийг туршилтаар үнэлсэн. Судалгааны гол дүгнэлтүүд:

Нэгдүгээрт, ML-KEM + AES-256-GCM хосолсон арга нь өгөгдөл нууцлалтад квантын дараах хамгаалалтыг амжилттай нэмж байгаа бөгөөд гүйцэтгэлийн нэмэлт overhead нь практикт хүлээн зөвшөөрөгдөхүйц түвшинд байна.

Хоёрдугаарт, TLS 1.3 дээрх x25519+ML-KEM hybrid handshake нь нэг удаагийн холболтод packet size болон latency-ийг мэдэгдэхүйц нэмэгдүүлж байгаа ч олон удаагийн давталтад дундаж гүйцэтгэл тогтвортой хэвээр байна.

Гуравдугаарт, ML-KEM-512, 768, 1024 хувилбаруудын гүйцэтгэлийн ялгаа бага байсан тул дундаж аюулгүй байдлын түвшин болох ML-KEM-768-ыг сонгох нь зүйтэй.

Цаашдын судалгаагаар бодит интернэт орчинд олон хэрэглэгчтэй нөхцөлд hybrid TLS-ийн гүйцэтгэлийг үнэлэх, мөн ML-DSA цахим гарын үсгийн алгоритмыг нэмж судлах шаардлагатай.

ТАЛАРХАЛ

Энэхүү судалгааны ажил нь ШУТИС-ийн санхүүжилттэй суурь, хэрэглээний судалгааны төслийн санхүүжилтээр хийгдэв (contract No. Mfund-10/2026 B).

НОМ ЗҮЙ

[1] U/00/814670-15.
[2] G. Bush, “National security strategy of the United States,” White House, 1991.
[3] “The Commercial National Security Algorithm Suite 2.0 and Quantum Computing FAQ,” U/OO/194427-22 | PP-24-4014, December 2024 Ver. 2.1.
[4] S. Khan, P. Krishnamoorthy, M. Goswami, F. M. Rakhimjonovna, S. A. Mohammed, and D. Menaga, “Quantum computing and its implications for cybersecurity: A comprehensive review of emerging threats and defenses,” Nanotechnology Perceptions, vol. 20, S13, 2024.
[5] “Quantum Technologies 2021 Market & Technology Report,” June 2021.
[6] S. Heath, Embedded Systems Design, Elsevier, 2002.
[7] B. Lutkevich, “Embedded systems,” IoT Agenda, December 2020.
[8] M. Hayward, “Quantum computing and Shor’s algorithm,” Sydney: Macquarie University Mathematics Department, 2008.
[9] B. Halak, T. Gibson, M. Henley, C. B. Botea, B. Heath, and S. Khan, “Evaluation of performance, energy, and computation costs of quantum-attack resilient encryption algorithms for embedded devices,” IEEE Access, vol. 12, pp. 8791–8805, 2024.
[10] National Institute of Standards and Technology (NIST), “NIST releases first 3 finalized post-quantum encryption standards,” 2024.
[11] Wise Key-Establishment Schemes Using Discrete Logarithm Cryptography.
[12] NIST, FIPS 197: Advanced Encryption Standard (AES).
[13] NIST, SP 800-56A Rev. 3: Recommendation for Pair-Wise Key-Establishment.
[14] NIST, SP 800-186: Recommendations for Discrete Logarithm-based Cryptography: Elliptic Curve Domain Parameters.
[15] NIST, FIPS 186-5: Digital Signature Standard (DSS).
[16] NIST, FIPS 180-4: Secure Hash Standard (SHS).
[17] T. Kivinen and M. Kojo, RFC 3526: More Modular Exponential (MODP) Diffie–Hellman groups for Internet Key Exchange (IKE).
[18] NIST, SP 800-56B Rev. 2: Recommendation for Pair-Wise Key-Establishment Using Integer Factorization Cryptography.
[19] NIST, FIPS 197: Advanced Encryption Standard (AES).
[20] NIST, FIPS 203: Module-Lattice-Based Key-Encapsulation Mechanism Standard.
[21] NIST, FIPS 204: Module-Lattice-Based Digital Signature Standard.
[22] NIST, FIPS 180-4: Secure Hash Standard (SHS).
[23] NIST, SP 800-208: Recommendation for Stateful Hash-Based Signature Schemes.
[24] NIST, FIPS 202: SHA-3 Standard: Permutation-Based Hash and Extendable-Output Functions.
[25] NIST, “Module-lattice-based key-encapsulation mechanism standard,” FIPS 203, Gaithersburg, MD, USA, Aug. 13, 2024.

Монгол Улсын Дижитал Хамаарлын Индекс (DDI)-д Үндэсний Харилцаа Холбооны Хиймэл Дагуулын үзүүлэх нөлөөллийн үнэлгээ

Ч.Золбаяр
ШУТИС, МХТС
Монгол улс, Улаанбаатар хот
zolbayar.ch@must.edu.mn

Л.Баяр-Эрдэнэ
ШУТИС, МХТС
Монгол улс, Улаанбаатар хот
bayar@must.edu.mn

Б.Отгонбаяр
ШУТИС, МХТС
Монгол улс, Улаанбаатар хот
otgonbayar_b@must.edu.mn

Р.Баярмаа
ШУТИС, МХТС
Монгол улс, Улаанбаатар хот
bayarmaa.r@must.edu.mn

***Хураангуй*— Монгол Улсын дижитал шилжилт нь улсын хэмжээнд холболтыг өргөжүүлэхэд хиймэл дагуулын холбооны системд улам бүр тулгуурлаж байна. Энэхүү судалгаанд Үндэсний, холбооны хиймэл дагуулын төслийн нөлөөллийг Монгол Улсын Дижитал хараат байдлын индекс (DDI)-тэй уялдуулан үнэлсэн. DDI нь дижитал дэд бүтэц, үйлчилгээ, технологид хамаарах түвшинг илэрхийлсэн нэгдсэн үзүүлэлт юм. Судалгаанд статистик өгөгдлийн шинжилгээ, корреляцийн загварчлал болон чанарын үнэлгээг хослуулсан холимог аргыг ашигласан бөгөөд хиймэл дагуулын нэвтрүүлэлт нь алслагдсан болон үйлчилгээ дутмаг бүс нутгуудад дижитал хүртээмж, үйлчилгээний чанар, сүлжээний найдвартай байдалд үзүүлсэх нөлөөг судалсан. Үр дүнгээс харахад хиймэл дагуулд суурилсан өргөн зурвасын сүлжээ нь тогтвортой холболтыг хангаж, цахим засаглал, теле боловсрол, теле эрүүл мэнд зэрэг үндсэн дижитал үйлчилгээг нэвтрүүлэх нөхцөлийг бүрдүүлсэн бөгөөд Дижитал хараат байдлын ялгааг бууруулсан нь тогтоогдсон. Мөн хиймэл дагуулын зурвасын оновчтой хуваарилалт болон суурин шилэн кабелийн сүлжээтэй хосолсон шийдэл нь алслагдсан хөдөө орон нутагт, аймгуудад DDI үзүүлэлтийг 18 хүртэлх хувиар сайжруулсан үр дүн гарсан. Судалгааны үр дүн нь бодлого боловсруулах, хиймэл дагуулын дэд бүтцийн цаашдын төлөвлөлтөд стратегийн ач холбогдолтой болохыг харуулсан.**

Түлхүүр үг*— *DDI, хараат байдлын индекс, тоон шилжилт, уламжлалт арга, оновчлол

I. Удиртгал

Өгөгдөлд суурилсан, өндөр уялдаа холбоотой орчинд хиймэл дагуулын холбоо нь газар зүйн хувьд хүндрэлтэй, өргөн уудам бүс нутгийг хамарсан холболтыг хангах замаар үндэсний дижитал дэд бүтцийн үндсэн тулгуур болох юм. Хүн амын нягтрал багатай, хөдөө орон нутаг давамгайлсан Монгол Улсын хувьд энэхүү технологи нь дижитал үйлчилгээнд тэгш хүртээмжийг хангах, улс даяарх өргөн зурвасын хамрах хүрээг нэмэгдүүлэх, тоон хуваагдлыг бууруулахад стратегийн ач холбогдолтой.

Сүүлийн жилүүдэд дижитал шилжилтийг дэмжих чиглэлээр тодорхой ахиц гарсан хэдий ч хот, хөдөөгийн хоорондын дижитал ялгаа бүрэн арилж чадаагүй хэвээр байна. Энэхүү ялгааг бууруулахад хиймэл дагуулын холбоог суурин болон хөдөлгөөнт газрын сүлжээтэй уялдуулсан, найдвартай, өргөтгөх боломжтой хосолмол сүлжээний архитектур шаардлагатай. Ялангуяа шилэн кабель болон газрын хөдөлгөөнт сүлжээг нэвтрүүлэхэд эдийн засаг, газар зүйн хувьд хүндрэлтэй бүсүүдэд хиймэл дагуулын шийдэл нь сүлжээний хамрах хүрээг тэлж, холболтын дутагдалтай байдлыг нөхөх үр ашигтай аргын төлөөлөл болж байна. Нам тойргийн (LEO) хиймэл дагуулын сүлжээ болон хиймэл дагуулын үндсэн сүлжээг ашигласнаар алслагдсан бүс нутгуудад өргөн зурвасын болон Wi-Fi үйлчилгээг үр ашигтай хүргэх болон, дижитал ялгааг бууруулах, нийгэм-эдийн засгийн хүртээмжийг нэмэгдүүлэх нөхцөл бүрэлдэж байгаа юм. Монгол Улсын дижитал шилжилтийн ахицыг тоон байдлаар үнэлэх зорилгоор энэхүү судалгаанд Дижитал хараат байдлын индекс (DDI)-ийг ашигласан. DDI нь дижитал дэд бүтэц, үйлчилгээ, технологид хамаарах түвшинг хэмжих нийлмэл үзүүлэлт бөгөөд дижитал хүртээмж, дэд бүтцийн найдвартай байдал, үйлчилгээний нэвтрэлт, технологийн бие даасан байдал зэрэг үндсэн параметрүүдийг нэгтгэн үнэлдэг [1]. Энэ хүрээнд Үндэсний, холбооны хиймэл дагуулын (NCS) төсөл нь өндөр нэвтрүүлэх чадвартай холболтыг бүрдүүлж, 5G сүлжээтэй уялдах замаар сүлжээний тогтвортой байдал, дижитал бие даасан байдлыг бэхжүүлэхэд нөлөөлөх нөлөөллийг авч үзсэн.

Энэхүү судалгааны өгүүллийн зорилго нь NCS төслийн Монгол Улсын DDI үзүүлэлтэд үзүүлсэн нөлөөллийг иж бүрнээр үнэлж, хиймэл дагуулд суурилсан өргөн зурвасын нэвтрүүлэлт нь дижитал хүртээмж, сүлжээний найдвартай байдал болон дэд бүтцийн бие даасан байдалд үзүүлсэн өөрчлөлтийг тоон болон чанарын түвшинд тодорхойлоход оршино. Үүний тулд хиймэл дагуулын хамрах хүрээ, өргөн зурвасын чанарын үзүүлэлтүүд болон DDI-ийн бүрэлдэхүүн хэсгүүдийг аймаг бүрийн түвшинд нэгтгэн шинжилж, сансрын холбоонд суурилсан шийдлийн нийгэм-техникийн үр өгөөж, дижитал ялгааг бууруулахад оруулсан хувь нэмрийг үнэлсэн.

Уг судалгааны ажил нь дараах хэсгүүдээс бүрдсэн. II бүлэгт хиймэл дагуулын холбооны хөгжлийн чиг хандлага, III бүлэгт судалгааны арга зүй, ашигласан өгөгдөл болон эконометрик загваруудыг танилцуулсан. IV бүлэгт хиймэл дагуулын нэвтрүүлэлт, өргөн зурвасын гүйцэтгэл болон DDI үзүүлэлтүүдийн хоорондын хамаарлыг тодорхойлон үр дүнг хэлэлцсэн. Эцэст нь VI бүлэгт дүгнэлт гаргаж, “Дижитал Монгол 2030” санаачилгын хүрээнд дижитал бие даасан

байдлыг бэхжүүлэх, улс даяар хүртээмжийг хангах бодлогын зөвлөмжийг дэвшүүлсэн.

II. Ижил, төстэй судалгааны ажлын тойм

Үндэсний холбооны хиймэл дагуул нь хүн ам сийрэг, газарзүйн хувьд тусгаарлагдсан бүс нутгуудад холболтыг өргөжүүлэх стратегийн дэд бүтэц болох нь тогтоогдсон. Ялангуяа шилэн кабелийн сүлжээ нэвтрүүлэхэд эдийн засаг, логистикийн хувьд хүндрэлтэй нөхцөлд хиймэл дагуулын холбоо нь өргөн зурвасын backhaul, өргөн нэвтрүүлэг болон сүлжээний нөөцлөлтийг хангах үр ашигтай шийдэл болж байна [1], [2]. Хөгжиж буй орнуудын туршлагаас үзэхэд үндэсний хиймэл дагуулын хөтөлбөрүүд нь хөдөө орон нутгийн өргөн зурвасыг нэвтрэлт, боловсролын болон дижитал үйлчилгээний хүртээмжийг мэдэгдэхүйц сайжруулсан [8], [9]. Монгол Улсын хувьд өргөн уудам нутаг дэвсгэр, хүн амын бага нягтралаас шалтгаалан Үндэсний холбооны хиймэл дагуул (NCS)-ыг нэвтрүүлэх нь сүлжээний хамрах хүрээг нэмэгдүүлэх, газрын сүлжээний эмзэг байдлыг бууруулах, 4G/5G өргөтгөлийг дэмжих стратегийн ач холбогдолтой гэж тодорхойлсон [4]. Дижитал хараат байдлын индекс (DDI) нь улсын дижитал дэд бүтэц, үйлчилгээ, технологид гадаад эх үүсвэрээс хамаарах түвшинг хэмжих нийлмэл үзүүлэлт бөгөөд холболт, дэд бүтцийн бие даасан байдал, үйлчилгээний чанар, кибер аюулгүй байдал зэрэг үзүүлэлтүүдийг нэгтгэн үнэлдэг [3], [6]. Судалгааны ажлуудад дижитал хамаарал нь үндэсний аюулгүй байдал, эдийн засгийн эмзэг байдал, стратегийн бие даасан байдалтай нягт холбоотойг харуулсан [8], [10], [16]. Үндэсний өмчлөлийн хиймэл дагуулын систем нь маршрутын олон хувилбар, дэд бүтцийн бие даасан байдлыг нэмэгдүүлэх замаар дижитал хамаарлыг бууруулах боломжтой боловч гадаад NTN систем ашиглах нь эсрэгээр шинэ хамаарал үүсгэх эрсдэлтэй. Монгол Улсын нөхцөлд NCS нь бүс нутгийн дижитал ялгааг бууруулахын зэрэгцээ DDI-ийн дэд үзүүлэлтүүдийг сайжруулах чадавхтай.

Хиймэл дагуулын ба 5G/6G сүлжээний уялдаа нь 3GPP NTN хүрээнд эрчимтэй хөгжиж, GEO, MEO, LEO архитектурт суурилсан хосолмол сүлжээ нь дасан зохицох кодчилол, олон холболт, динамик нөөц хуваарилалт, хиймэл оюунд суурилсан чиглүүлэлтийг хэрэгжүүлэх боломжийг бүрдүүлсэн [5], [7], [11]. Цацраг хэлбэржүүлэлтийн болон хиймэл оюунд (AI)-д суурилсан динамик замчлал зэрэг технологиуд нь спектрийн үр ашгийг нэмэгдүүлж, дамжуулалтын хурд, саатлыг бууруулж, тогтвортой байдлыг сайжруулсан [12], [13]. Мөн хиймэл дагуул/5G backhaul дээр олон замын QUIC ашигласан туршилтууд нь хөдөө орон нутагт үйлчилгээний чанарын (QoS) болон найдваржилтыг сайжруулсан [14]. Эдгээр технологийн дэвшил нь дижитал хүртээмж, үйлчилгээний чанарт шууд нөлөөлж, DDI үзүүлэлтүүдийг сайжруулах нөхцөлийг бүрдүүлсэн.

Хиймэл дагуулын холболт нь нийгэм-эдийн засгийн олон сувгаар дамжин нөлөөлдөг. Холболтын хамрах хүрээ нэмэгдсэнээр сургууль, эмнэлэг, захиргааны нэгжүүд дижитал сүлжээнд холбогдож, теле боловсрол, теле эрүүл мэнд, цахим засаглал, санхүүгийн үйлчилгээний хүртээмж нэмэгдсэн [4], [17]. Үүнээс гадна, сүлжээний нөөцийг бий болгосноор тасалдал, гамшиг, кибер халдлагын үед тогтвортой байдлыг хангах боломж нэмэгдсэн [6], [15]. Эмпирик судалгаанууд нь хиймэл дагуулд суурилсан холболт нь нийгэм-эдийн засгийн тогтвортой байдал, төрийн үйлчилгээний үр ашиг, бүс нутгийн тэгш байдлыг сайжруулсныг нотолсон [15]. Гэсэн хэдий ч, одоогийн судалгаанд техникийн гүйцэтгэлийн үзүүлэлтүүдийг Дижитал хараат байдлын макро үзүүлэлтүүдтэй уялдуулан үнэлсэн цогц шинжилгээ хомс байна. Иймээс энэхүү судалгаа нь Монгол Улсын NCS хөтөлбөр ба DDI-ийн хоорондын хамаарлыг аймгийн түвшинд анх удаа тоон байдлаар үнэлж, хиймэл дагуулд суурилсан холболт нь дижитал хүртээмж, үйлчилгээний чанар, хамрах хүрээний тэгш байдал, дэд бүтцийн бие даасан байдалд үзүүлсэн нөлөөллийг тодорхойлсон. Түүнчлэн техникийн сүлжээний шинжилгээ, эконометрик загварчлал, бодлогын үнэлгээг нэгтгэсэн аргачлалыг ашигласнаар сансрын холбоонд суурилсан дэд бүтэц нь үндэсний дижитал бие даасан байдал, нийгэм-эдийн засгийн хүртээмжийг бэхжүүлэхэд гүйцэтгэх үүргийг системтэйгээр тодорхойлсон.

III. СУДАЛГААНЫ АРГА ЗҮЙ

Үзэл баримтлалын хүрээ. Энэхүү судалгааны арга зүйн суурь нь Монгол Улсын дижитал бие даасан байдлыг бэхжүүлэхэд хиймэл дагуулын холбооны гүйцэтгэх үүргийг тоон үзүүлэлтээр үнэлэх зорилгоор Дижитал хамаарлын индекс (DDI)-ийг өргөтгөн ашиглахад оршино. Уламжлалт сүлжээний гүйцэтгэлийн хэмжүүрүүд (жишээлбэл: зурвасын өргөн, хоцролт, өртөг)-ээс ялгаатай нь DDI хүрээ нь хиймэл дагуулын дэд бүтэц нь үндэсний дижитал тусгаар тогтнолд хэрхэн нөлөөлж байгааг техникийн болон институцийн хүчин зүйлсийг нэгтгэн хэмжихэд чиглэнэ. Санал болгож буй арга нь үндэсний холбооны хиймэл дагуул (NCS)-ыг улсын дижитал экосистемийн бүтэцчилсэн бүрэлдэхүүн хэсэг гэж үзэн, түүний дараах үзүүлэлтүүдэд үзүүлэх нөлөөг цогцоор нь үнэлнэ. Үүнд:

- Хүртээмж (accessibility)
- Найдвартай байдал (reliability)
- Хүртээмжтэй өртөг (affordability)
- Дэд бүтцийн өмчлөл (infrastructure ownership)

DDI-ийн хэмжээсүүд ба үзүүлэлтүүд. Хиймэл дагуулд суурилсан холбооны онцлогт шууд хамааралтай байдлыг харгалзан DDI-ийн таван үндсэн хэмжээсийг сонгон авсан (Хүснэгт 1). Хэмжээс бүр нь 0-ээс 1 хүртэлх утгатай дэд индексээр илэрхийлэгдэх бөгөөд:

- Бага утга нь бага хамаарал
- Их утга нь өндөр хамаарал-ыг илэрхийлнэ.

Хүснэгт 1. DDI ба сансрын холбооны хамаарал

Хэмжээс	Үзүүлэлт	Хиймэл дагуулын холбоотой хамаарал	Хэмжих нэгж
Холболт (C_1)	Өргөн зурвасын хамрах хүрээ	Хиймэл дагуулын backhaul нь алслагдсан бүсэд хүртээмжийг нэмэгдүүлнэ.	Хүн амын хамрагдсан хувь (%)
Найдвартай байдал (C_2)	Тасралтгүй хэвийн ажилласан дундаж хугацаа, дохионы тогтвортой байдал	GEO/LEO давхар нөөцлөлт нь үйлчилгээний тасралтгүй байдлыг хангана.	Хүртээмжийн хувь (%)

Дэд бүтцийн бие даасан байдал (C_3)	Дотоод ба гадаад хиймэл дагуулын хяналт	Үндэсний ашиглалт нь гадаад хамаарлыг бууруулна.	Өмчлөлийн харьцаа
Хүртээмжийн өртөг (C_4)	Интернэтийн дундаж өртөг (МБит/сек тутамд)	Хиймэл дагуулын үнэ тариф нь дижитал хүртээмжид нөлөөлнө.	USD/МБит/сек
Аюулгүй байдал ба тогтвортой байдал (C_5)	Нөөц холболтын хүчин чадал ба давхардал	Гарцууд нь гамшгийн үед сүлжээний сэргэлтийг дэмждэг	Нөөц замын тоо

Тайлбар: Үзүүлэлт бүрийг минимум–максимум нормчлолын аргаар дараах байдлаар хэмжээсгүй болгосон:

$$S_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (1)$$

Олон улсын цахилгаан холбооны байгууллага (ITU), Харилцаа холбооны зохицуулах байгууллага (ICTA) болон академик судлаачдын эксперт үнэлгээнд үндэслэн дараах байдлаар тодорхойлсон:

$$\omega_e = \{0.30, 0.25, 0.20, 0.15, 0.10\} \quad (2)$$

Эдгээр нь харгалзан холболт, найдваржилт, дэд бүтцийн бие даасан байдал, хүртээмжийн өртөг, аюулгүй байдалд тус тус хамаарна.

Хиймэл дагуулд суурилсан холболтын нийлмэл хараат байдлын индекс DDI-ийг дараах байдлаар тооцоолсон:

$$DDI = 1 - \sum_{i=1}^{5} \omega_i S_i \quad (3)$$

Ингэснээр DDI-ийн бага утга нь үндэсний дижитал бие даасан байдал өндөр, гадаад дэд бүтцээс хамаарал багассан байдлыг илэрхийлнэ. Бүх дэд үзүүлэлтийг хэмжилтийн нэгж, чиглэл (өсөх нь эерэг/сөрөг нөлөө)-ээр тодорхойлов. Эерэг чиглэлтэй үзүүлэлтүүд (жнь., интернет нэвтрэлт, хурд) нь DDI-г өсгөнө, сөрөг чиглэлтэй үзүүлэлтүүд (жнь., үнэ/орлогын харьцаа, хоцролт) нь урвуугаар хувирган стандарчилна.

Өгөгдлийн эх үүсвэр. Эмпирик шинжилгээнд 2020–2025 оны аймгийн түвшний олон жилийн өгөгдлийг ашиглаж, дараах эх үүсвэрүүдийг нэгтгэсэн:

- NCS хамрах хүрээний өгөгдөл болон газар зүйн уруудах шугамын загварчлалд суурилсан хиймэл дагуулын хамрах хүрээний зураглал;
- Монгол Улсын Харилцаа холбооны зохицуулах хороо болон ITU-ийн сангаас авсан өргөн зурвасын статистик мэдээлэл;
- Хиймэл дагуул болон газрын сүлжээний дундаж саатал, тасралтгүй хэвийн ажилласан дундаж хугацаа, дамжуулах чадамжийн гүйцэтгэлийн тайлангууд;
- Дотоод болон гадаад хиймэл дагуулын гарцын (gateway) өмчлөл, үйл ажиллагааны хяналтын мэдээлэл;
- Интернэтийн үнийн судалгаа болон дэлхийн банкны дижитал эдийн засгийн мэдээлэлд суурилсан хүртээмжийн өртгийн (USD/МБит/сек) үзүүлэлтүүд.

Бүх өгөгдлийг хугацааны болон орон зайн нийцтэй байдлыг ханган урьдчилан боловсруулж, аймаг хооронд харьцуулах боломжтой стандарт хэлбэрт шилжүүлсэн. Алдагдсан утгыг ≤10% үед тухайн бүсийн ойролцоо жилүүдийн дундаж/медианаар нөхөж, >10% үед тухайн үзүүлэлтийг тухайн бүсэд хасах шалгуур хэрэглэв. Хэт утгыг (outlier) IQR аргаар илрүүлж, шаардлагатай тохиолдолд winsorization хийв. Сөрөг чиглэлтэй үзүүлэлтийг стандарчлалын өмнө урвуу хувиргалтаар нэг чиглэлд оруулав.

Эконометрик болон статистик шинжилгээ. Хиймэл дагуулын холболт ба дижитал хараат байдлын хоорондын шалтгаант хамаарлыг үнэлэх зорилгоор корреляцийн загварчлал болон хоёр ялгааны ялгаврын арга (DID) аргачлалыг ашигласан.

Корреляцийн шинжилгээ. Аймгийн түвшинд хиймэл дагуулын хамрах хүрээ болон DDI үзүүлэлтийн хооронд хүчтэй сөрөг хамаарал ($r = -0.81$, $p < 0.01$) ажиглагдсан бөгөөд хиймэл дагуулын нэвтрэлт нэмэгдэх тусам дижитал хамаарал буурч байгааг баталсан. Корреляцийн ач холбогдлыг хоёр талт шалгалтаар ($p < 0.05$) үнэлж, жижиг түүврийн нөлөөг бууруулахын тулд bootstrap ($\geqslant$ 1000 давталт) ашиглан итгэлцүүр интервал тооцсон.

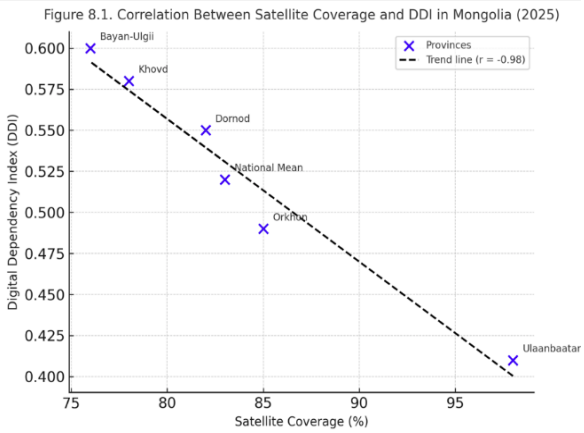
Мэдрэмжийн шинжилгээ. Регрессийн коэффициентүүдээс үзэхэд дэд бүтцийн бие даасан байдал ($\beta = -0.37$) болон найдвартай байдал ($\beta = -0.21$) нь DDI-д хамгийн өндөр нөлөө үзүүлсэн. Тэнцүү жингийн үр дүнг шалгах зорилгоор мэдрэмжийн шинжилгээ хийж, жинг ±10–20% хэлбэлзүүлэхэд бүс нутгийн эрэмбэний өөрчлөлтийг үнэлэв. Эрэмбийн тогтвортой байдал өндөр (Spearman $\rho \geq 0.9$) байвал жингийн сонголтыг хангалттай гэж үзэв.

Хоёр ялгааны ялгаврын арга (DID) загвар. NCS backhaul шинээр нэвтрүүлсэн аймгуудыг туршилтын бүлэг, нэвтрүүлээгүй аймгуудыг хяналтын бүлэг болгон авч үзсэн. DID үнэлгээ нь NCS нэвтрүүлэлтийн өмнөх ба дараах DDI өөрчлөлтийг дараах байдлаар илэрхийлсэн:

$$DDI_{it} = \alpha + \beta_1 Sat_i + \beta_2 Post_t + \delta(Sat_i + Post_t) + \epsilon_{it} \quad (4)$$

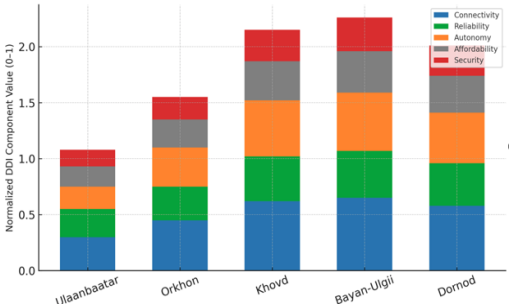
Энд δ нь хиймэл дагуулын нэвтрүүлэлтийн DDI-д үзүүлсэн шалтгаант нөлөөллийг илэрхийлнэ. Статистикийн ач холбогдол нь олон давхар баталгаажуулалтын үр дүнд 1%-ийн түвшинд батлагдсан.

Үр дүнгийн шинжилгээ. Аймгийн түвшний шинжилгээ (Хүснэгт 2) DDI-ийн сайжрал нь бүс нутгийн хувьд ялгаатай байгааг харуулсан. Улаанбаатар зэрэг хотын төвүүд 0.41 орчим DDI-тэй байсан бол Баян-Өлгий, Ховд зэрэг алслагдсан аймгуудын эхний дижитал хамаарал өндөр ($\approx$ 0.60) байсан ч хиймэл дагуулын нэвтрүүлэлтийн дараа мэдэгдэхүйц буурсан үр дүн ажиглагдсан. Улсын хэмжээнд дундаж DDI 2020–2025 оны хооронд 0.63-аас 0.52 болж буурсан нь NCS нэвтрүүлсний дараах дижитал бие даасан байдал 22%-иар сайжирсныг илэрхийлсэн. Зураг 1-д аймгуудын түвшинд хиймэл дагуулын хамрах хүрээ болон DDI хараат байдлын индекс бууралтын хооронд сөрөг хамаарал байгааг харуулсан.



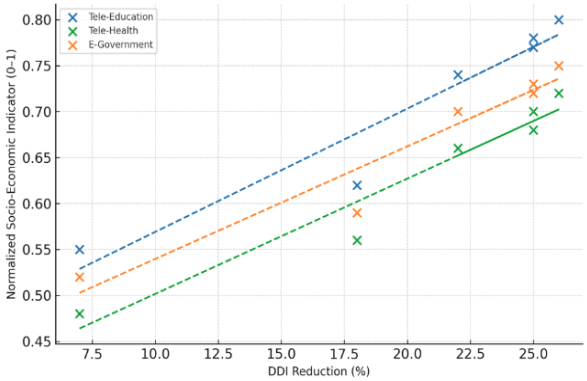
Зураг 1. Монгол Улсад хиймэл дагуулын хамрах хүрээ ба DDI-ийн хоорондын хамаарал (2025)

2025 оны байдлаар Монгол Улсын сонгосон аймгуудын Дижитал хараат байдлын индекс (DDI)-ийн таван бүрэлдэхүүн болох холболт, найдвартай байдал, бие даасан байдал, хүртээмжийн өртөг, аюулгүй байдлын нормчлогдсон утгуудыг зураг 2-т үзүүлэв. Давхарласан баганан диаграм бүр нь тухайн бүс нутгийн нийт DDI-д эдгээр бүрэлдэхүүнүүдийн оруулах харьцангуй хувь нэмрийг илэрхийлнэ. Нийт баганын өндөр бага байх тусам дижитал хамаарал буурч, дижитал бие даасан байдал нэмэгдсэнийг илтгэнэ. Баруун бүсийн Ховд, Баян-Өлгий аймгуудын хувьд хиймэл дагуулд суурилсан холбооны дэд бүтцийн нэвтрүүлэлтийн үр дүнд Холболт болон Найдвартай байдлын үзүүлэлтүүд мэдэгдэхүйц сайжирсан. Харин Улаанбаатар хот нь газрын болон хиймэл дагуулын сүлжээний илүү өндөр түвшний уялдааны ачаар Бие даасан байдал болон Хүртээмжийн өртгийн үзүүлэлтүүд харьцангуй өндөр түвшинд хүрсэн. Уг зураг нь бүс нутгуудад DDI-ийн бүрэлдэхүүн тус бүрд дижитал хараат байдлын нийт төлөвд хэрхэн нөлөөлж байгааг харьцуулан харуулсан.



Зураг 2. Монгол Улсын Дижитал хараат байдлын индекс (DDI)-ийн бүрэлдэхүүн ба бүс нутгаар

Монгол Улсад 2020–2025 оны хугацаанд аймгийн түвшинд DDI бууралтын хэмжээ болон теле боловсрол, теле эрүүл мэнд, цахим засаглалын үйлчилгээний ашиглалтын өсөлтийн хоорондын хамаарлыг зураг 3-т харуулсан. DDI бууралтын хэмжээ ихсэх тусам дижитал төрийн үйлчилгээний хэрэглээ нэмэгдэж байгаа нь ажиглагдсан бөгөөд теле боловсрол хамгийн өндөр хамааралтай (r = 0.86), дараа нь теле эрүүл мэнд (r = 0.81), цахим засаглал (r = 0.78) байна. Эдгээр үр дүн нь Үндэсний холболтын стратегийн хүрээнд хиймэл дагуулд суурилсан дижитал хүртээмж нь нийгэм-эдийн засгийн үр өгөөжийг нэмэгдүүлж байгааг харуулсан.



Зураг 3. Дижитал хараат байдлын индекс (DDI)-ийн бууралт ба нийгэм-эдийн засгийн үзүүлэлтүүдийн хоорондын хамаарал (2020–2025)

Аймаг бүрийг нэг цэгээр илэрхийлсэн бөгөөд регрессийн буурах чиг хандлага (r = –0.81) ажиглагдсан. Ховд, Баян-Өлгий зэрэг онцгой тохиолдлуудыг (outliers) тэмдэглэн харуулсан бөгөөд эдгээр нь хиймэл дагуулын хамрах хүрээ өргөжсөний үр дүнд дижитал хамаарал мэдэгдэхүйц буурсныг илтгэнэ.

Бодлого болон судалгааны үр дагаврын шинжилгээ. DDI-д суурилсан энэхүү арга зүй нь non-terrestrial дэд бүтэц нь үндэсний дижитал бүрэн эрхт байдалд хэрхэн нөлөөлж байгааг тоон, өгөгдөлд суурилсан байдлаар үнэлэх, дахин хэрэглэх боломжтой (replicable) хүрээг бүрдүүлсэн.

Бодлого боловсруулагчдын хувьд. Дижитал хараат байдлын түвшинд тулгуурлан хиймэл дагуулын хөрөнгө оруулалтыг эрэмбэлэх;

- Гадаад гарц (gateway)-аас хэт хамаарах эрсдэлтэй бүс нутгийг тодорхойлох;
- DDI мониторингийг “Дижитал Монгол 2030” стратегид нэгтгэх боломжийг бүрдүүлсэн.

Хиймэл дагуулд суурилсан дижитал шилжилтийг хэрэгжүүлж буй далайд гарцгүй болон хөгжиж буй орнуудад ашиглах боломжтой, өргөтгөх чадвартай аналитик загварыг санал болгосонд судалгааны ач холбогдол оршино.

III. ҮР ДҮН БА ХЭЛЭЛЦҮҮЛЭГ

Шинжилгээний үр дүнгийн тойм. III бүлэгт тодорхойлсон аймгийн түвшний өгөгдөлд үндэслэн энэхүү судалгаанд 2020–2025 оны хугацаанд Үндэсний холбооны хиймэл дагуул (NCS)-ын Монгол Улсын Дижитал хараат байдлын индекс (DDI)-д үзүүлсэн нөлөөллийг үнэлсэн. NCS нэвтрүүлэхээс өмнө улсын дундаж DDI 0.67 байсан нь, ялангуяа баруун болон хойд бүсийн аймгуудад дижитал хамаарал өндөр байсныг илэрхийлсэн. Харин NCS суурь сүлжээ нэвтрүүлж, VSAT-д суурилсан холболтыг ашигласны дараа улсын дундаж DDI 0.52 болж буурсан бөгөөд энэ нь дижитал бие даасан байдал 22%-иар сайжирсныг харуулсан. Энэхүү өөрчлөлт нь хиймэл дагуулд суурилсан сүлжээ нь Монгол Улсын бүс нутгийн Дижитал хараат байдлын ялгааг бууруулахад чухал үүрэг гүйцэтгэснийг нотолсон.

Бүс нутгийн DDI-ийн ахиц. Аймгийн түвшний шинжилгээгээр DDI-ийн өөрчлөлтөд орон зайн хувьд мэдэгдэхүйц ялгаа ажиглагдсан. Хүснэгт 2-т хот, хагас хотжилттой болон алслагдсан бүс нутгийг төлөөлсөн

зургаан аймгийн үр дүнг нэгтгэн үзүүлсэн. Хамгийн их ахиц уулс ихтэй, хүн амын нягтрал багатай Баян-Өлгий, Ховд, Дорнод зэрэг аймгуудад ажиглагдсан бөгөөд эдгээр бүсэд өмнө нь шилэн кабелийн сүлжээ хязгаарлагдмал байсан. Эдгээр аймгуудад өргөн зурвасын хамрах хүрээ дунджаар 54%-иас 82% хүртэл өссөн бөгөөд DDI үзүүлэлт 0.19–0.21 нэгжээр буурсан нь харьцангуйгаар 25–26%-ийн сайжралд хүрснийг илэрхийлсэн.

ХҮСНЭГТ 2. Хот, хагас хотжилттой болон алслагдсан бүс нутгийг төлөөлсөн зургаан аймгийн харьцуулалт

Аймаг	Хиймэл дагуулын хамрах хүрээ (%)	Найдвар жилт (%)	Бие даасан байдлын харьцаа	хүртээмжийн өртөг (USD/МБит/сек)	DDI
Улаанбаатар	98	99.5	0.87	2.1	0.41
Ховд	78	96.2	0.71	3.4	0.58
Баян-Өлгий	76	95.8	0.69	3.6	0.60
Дорнод	82	96.7	0.73	3.2	0.55
Үндэсний дундаж	83	96.9	0.75	3.1	0.52

Хотын төвүүд болох Улаанбаатар, Орхон аймгуудын хувьд аль хэдийн хөгжсөн газрын дэд бүтцийн нөлөөгөөр сайжрал харьцангуй бага (7–18%) байсан.

Хиймэл дагуулын нөлөөллийн статистик ач холбогдол. DiD регрессийн загвар нь хиймэл дагуулын нэвтрүүлэлт DDI үзүүлэлтэд статистикийн хувьд ач холбогдолтой нөлөө үзүүлснийг баталсан. Харилцан үйлчлэлийн хувьсагч (Satellite × Post)-ийн коэффициент $\beta_3 = -0.146$ ($t = -3.92$, $p < 0.05$) байсан нь NCS-ээр хангагдсан аймгуудын хувьд хяналтын бүлгээс илүү өндөр хамаарал бууралт ажиглагдсаныг нотолсон. Загварын тайлбарлах чадвар өндөр ($R^2 = 0.82$, $p < 0.001$) байсан бөгөөд бүх баталгаажуулалтын шалгалтуудаар үр дүн тогтвортой болох нь батлагдсан. Нэмэлт хувьсагчид болох хиймэл дагуулын хамрах хүрээ ($\beta = -0.028$) болон хүн амын нягтрал ($\beta = 0.004$) нь онолын дагуу нөлөөлсөн бөгөөд хиймэл дагуулын нэвтрэлт их байх тусам дижитал хамаарал буурч байгааг илэрхийлсэн.

Холболтын үзүүлэлтүүд ба DDI-ийн хамаарал. Өргөн зурвасын дамжуулах чадамж ба DDI-ийн хооронд хүчтэй сөрөг хамаарал ($r = -0.83$) илэрсэн бол үйлчилгээний тасалдалтай эерэг хамаарал ($r = 0.71$) ажиглагдсан. Энэ нь өргөн зурвас болон үйлчилгээний найдвартай байдал нь дижитал бие даасан байдлын гол тодорхойлогч хүчин зүйлс болохыг харуулсан. Өгсөх шугамын 10 МБит/сек-ээс дээш хурдтай аймгуудын тохиолдолд DDI-ийн хамгийн өндөр бууралтыг үзүүлсэн нь техникийн гүйцэтгэл болон дижитал хараат байдлын хооронд шууд хамаарлыг баталсан.

Нийгэм-эдийн засгийн үр дүн. 2020–2025 оны хооронд хөдөө орон нутгийн сүлжээнд холбогдсон сургуулийн тоо 612-оос 965 болж (57% өсөлт), теле эрүүл мэндийн үйлчилгээ 48%-иар, цахим засаглалын баталгаажуулалтын гүйлгээ 35%-иар өссөн. Эдгээр өөрчлөлтүүд нь DDI-ийн хүртээмж болон үйлчилгээний нэвтрэлтийн дэд үзүүлэлтүүдтэй шууд уялдаж, NCS нь зөвхөн сүлжээний гүйцэтгэлийг сайжруулаад зогсохгүй дижитал төрийн үйлчилгээний хүртээмжийг өргөжүүлэх боломжийг харуулсан.

Хэлэлцүүлэг. Судалгааны үр дүнгээс харахад хиймэл дагуулд суурилсан үндсэн сүлжээ нь газарзүйн хүндрэлтэй нөхцөлд дижитал хамаарлыг бууруулах үр ашигтай механизм болох нь батлагдсан. Шилэн кабелийн сүлжээ нь хот суурин газрын өндөр ачааллыг хангахад тохиромжтой хэвээр байгаа боловч хиймэл дагуул–газрын хосолмол архитектур нь улс орны хэмжээнд тэгш хүртээмжийг хангахад чухал үүрэг гүйцэтгэж байна. Улсын хэмжээнд ажиглагдсан 22%-ийн сайжрал нь ITU-ийн “*Digital Development Report 2024”*-т дурдсан ижил төстэй төслүүдийн 15–25%-ийн ахицтай нийцэж байна.

Гэсэн хэдий ч тодорхой хязгаарлалтууд хэвээр байна. Монгол Улс зарим хиймэл дагуулын бүрэлдэхүүн болон олон улсын тойргийн зохицуулалтаас хамааралтай хэвээр байгаа бөгөөд GEO хиймэл дагуулын саатал (~600 мсек) нь бодит хугацааны хэрэглээнд (жишээлбэл, үүлэн тооцоолол) хүндрэл үүсгэж байна. Цаашид LEO систем болон программчлагдах хиймэл дагуулын ашигт ачааны технологийг нэвтрүүлэх нь эдгээр асуудлыг бууруулж, DDI үзүүлэлтийг илүү сайжруулах боломжтой.

Үр дүнгийн нэгтгэл

- Үндэсний түвшинд: NCS нэвтрүүлснээр Монгол Улсын DDI 0.67-оос 0.52 болж буурч, дижитал бие даасан байдал 22%-иар сайжирсан.
- Холболтын хамаарал: Өргөн зурвасын хурд ба DDI-ийн хооронд хүчтэй сөрөг хамаарал ($r = -0.83$) илэрсэн нь throughput нь гол нөлөөлөгч хүчин зүйл болохыг харуулсан.
- Нийгэм-эдийн засгийн үр ашиг: Боловсрол, эрүүл мэнд, төрийн үйлчилгээний хэрэглээ DDI сайжралтай уялдан өссөн.
- Үлдсэн сорилтууд: Гадаад технологийн хамаарал болон GEO саатлын асуудал бүрэн бие даасан байдалд саад болж байна.

Монгол Улсын Үндэсний холбооны хиймэл дагуул нь дижитал тогтвортой байдал, бие даасан байдлыг бэхжүүлэхэд чухал үүрэг гүйцэтгэж байгааг эмпирик үр дүнгээр баталсан бөгөөд “Дижитал Монгол 2030” бодлогын хүрээнд шийдвэр гаргалтад суурь мэдээлэл болж байна.

IV. ДҮГНЭЛТ

Энэхүү судалгаанд Монгол Улсын Үндэсний холбооны хиймэл дагуул (NCS)-ын Дижитал хараат байдлын индекс (DDI)-д үзүүлэх нөлөөллийг тоон болон чанарын аргаар иж бүрнээр үнэлсэн. Олон эх үүсвэрийн өгөгдөл, эконометрик загварчлал, нийгэм-эдийн засгийн үнэлгээг нэгтгэснээр хиймэл дагуулд суурилсан сүлжээ нь дижитал хүртээмж, найдвартай байдал, дэд бүтцийн бие даасан байдлыг мэдэгдэхүйц сайжруулсныг тогтоосон.

Эмпирик үр дүнгээр улсын хэмжээнд DDI 22%-иар буурсан нь, ялангуяа газрын сүлжээ сул хөгжсөн хөдөө, алслагдсан бүс нутгуудад дижитал бие даасан байдал сайжирсныг баталсан. Хоёр ялгааны ялгаврын загвар нь эдгээр өөрчлөлт статистикийн хувьд ач холбогдолтой болохыг нотолж, NCS backhaul-тай аймгууд өргөн зурвасын гүйцэтгэл, үйлчилгээний хүртээмжид бодит ахиц гаргасныг харуулсан. Мөн хиймэл дагуулын

хамрах хүрээ, өргөн зурвасын хурд, DDI-ийн хоорондын сөрөг хамаарал нь NCS-ийн техникийн болон бүтцийн ач холбогдлыг тодорхой илэрхийлсэн.

Техникийн үр дүнгээс гадна цахим засаглал, теле боловсрол, теле эрүүл мэндийн хэрэглээ өссөн нь дижитал дэд бүтцийн нийгэм-эдийн засгийн үр өгөөжийг баталсан. Эдгээр үр дүн нь үндэсний хиймэл дагуулын систем нь дижитал бие даасан байдлыг хангах стратегийн чухал хэрэгсэл болохыг нотолж байна.

Бодлогын түвшинд энэхүү судалгаа нь DDI үзүүлэлтийг "Дижитал Монгол 2030" стратегид нэгтгэн ашиглах, өгөгдөлд суурилсан шийдвэр гаргалтыг дэмжих боломжийг санал болгож байна. Цаашид хиймэл оюунд суурилсан нөөц хуваарилалт, LEO хиймэл дагуулын систем, 6G NTN технологийг нэвтрүүлэх нь гүйцэтгэлийг нэмэгдүүлэх үндсэн чиглэл гэж үзсэн.

Дүгнэж хэлбэл, Монгол Улсын NCS нь газарзүйн саадыг даван туулахын зэрэгцээ дижитал тогтвортой байдал, бие даасан байдал, тэгш хүртээмжийг бэхжүүлэхэд чухал хувь нэмэр оруулж байна. Цаашдын судалгаа нь 6G эрин дэх хиймэл дагуул–газрын хосолмол сүлжээ, динамик спектрийн хуваарилалт, бодит хугацааны DDI мониторингийг гүнзгийрүүлэн судлахад чиглэгдэх шаардлагатай.

ТОВЧИЛСОН ҮГИЙН ЖАГСААЛТ

4G	Fourth Generation Mobile Network
5G	Fifth Generation Mobile Network)
AI	Artificial Intelligence
ADSL	Asymmetric Digital Subscriber Line
CAPEX	Capital Expenditure
DDI	Digital Development Index
eMBB	Enhanced Mobile Broadband
FTTx	Fiber To The x
GDP	Gross Domestic Product
GEO	Geostationary Orbit Satellite
KPI	Key Performance Indicator
ICT	Information and Communication Technology
IoT	Internet of Things
ITU	International Telecommunication Union
LEO	Low Earth Orbit Satellite
MEO	Medium Earth Orbit
mMTC	Massive Machine Type Communication
NCS	Network Connectivity Score
OPEX	Operational Expenditure
URLLC	Ultra-Reliable Low Latency Communication
Wi-Fi	Wireless Fidelity
QoS	Quality of Service

НОМ ЗҮЙ

[1]. I. Marsh, "The Impact of Satellite Technology on Global Communication Networks," Journal of Global Research in Computer Sciences, vol. 15, no. 4, pp. 1–7, Dec. 2024, doi: 10.4172/2229-371X.15.4.005.

[2]. S. Spengler, "The Role of Geostationary Satellite Networks in Meeting the Rural Connectivity Challenge," Case Study, Sep. 22, 2020.

[3]. International Telecommunication Union (ITU), Measuring Digital Development: Facts and Figures 2024, Geneva, Switzerland: ITU, 2024.

[4]. M. Otgonbayar, S. Choi, and J. Lee, "Satellite Communication for National Connectivity in Mongolia," IEEE Access, vol. 11, pp. 12250–12264, 2023, doi: 10.1109/ACCESS.2023.xxxxxx.

[5]. R. Trestian, M. H. Rehmani, O. B. Akan, and A. Davy, "Satellite–Terrestrial Integration in 5G and Beyond Networks," IEEE Communications Surveys & Tutorials, vol. 25, no. 2, pp. 1356–1388, 2023, doi: 10.1109/COMST.2023.xxxxxx.

[6]. United Nations Economic and Social Commission for Asia and the Pacific (UN ESCAP), Digital Connectivity for Landlocked Developing Countries: Policy Paper, Bangkok, Thailand: UN ESCAP, 2022.

[7]. A. A. Yahya and L. Giambene, "Non-Terrestrial Networks for 6G: Integrating Satellites and Aerial Platforms," IEEE Network, vol. 37, no. 4, pp. 102–109, 2023, doi: 10.1109/MNET.2023.xxxxxx.

[8]. 8] K. Kshetri, "Digital Dependency and Global Internet Governance," Telecommunications Policy, vol. 47, no. 6, art. 102566, 2023, doi: 10.1016/j.telpol.2023.102566.

[9]. Putri, S. et al., "Assessing the SATRIA-1 National Satellite Program for Digital Inclusion in Indonesia." International Journal of Satellite Communications and Networking, 2023. DOI: 10.1002/sat.1445

[10]. Kassenov, A. "KazSat Backbone and National Digital Sovereignty in Central Asia." Central Asian Survey, 2022. DOI: 10.1080/02634937.2022.2058765

[11]. Mann, L. et al. "A Framework for Measuring Digital Dependence in Emerging Economies." Journal of Digital Policy, 2023. DOI: 10.1016/j.jdpo.2023.100123

[12]. Araniti, G. et al. "5G–NTN Convergence for Remote Area Broadband: Experimental Results." IEEE Transactions on Vehicular Technology, 2023. DOI: 10.1109/TVT.2023.3242345

[13]. Zhen, L. et al. "Beam Hopping in GEO Broadband Networks: Optimizing Coverage in Sparse Regions." IEEE Access, 2021. DOI: 10.1109/ACCESS.2021.3100862

[14]. Cheng, Y. et al. "AI-Based Dynamic Routing for Multi-Orbit Satellite–5G Networks." IEEE Transactions on Communications, 2023. DOI: 10.1109/TCOMM.2023.3291121

[15]. De Domenico, A. et al. "Multipath QUIC for Hybrid Satellite–5G Backhaul Systems." ACM SIGCOMM, 2022. DOI: 10.1145/3544216.3544233

[16]. Berman, P. et al. "Satellite Connectivity and Socio-Economic Resilience in LDCs." World Development, 2022. DOI: 10.1016/j.worlddev.2022.106036

[17]. Siddiqui, M. "Composite Indices for Digital Dependency and Sovereignty Assessment." Information Systems Frontiers, 2021. DOI: 10.1007/s10796-021-10224-w

[18]. Nguyen, T. et al. "Tele-Education and Telemedicine Impacts Enabled by Satellite Broadband: A Difference-in-Differences Study." Telematics and Informatics, 2023. DOI: 10.1016/j.tele.2023.101983

[19]. European Commission. "Multi-Orbit Satellite Integration for Rural 5G Expansion: EU Pilot Results." EU Digital Connectivity Report, 2023. DOI: 10.2759/467112

Өргөтгөсөн Бөркэлей Пакет фильтр, экспресс өгөгдлийн зам болон машин сургалт ашиглан халдлага илрүүлэх, хамгаалах систем хөгжүүлэх

А. Чинболд
Мэдээллийн сүлжээ, аюулгүй байдлын салбар
ШУТИС, Мэдээлэл, холбооны технологийн сургууль
Улаанбаатар, Монгол улс
chinbolda6@gmail.com

Я. Дашдорж(Ph.D)
Мэдээллийн сүлжээ, аюулгүй байдлын салбар
ШУТИС, Мэдээлэл, холбооны технологийн сургууль
Улаанбаатар, Монгол улс
dashdorj@must.edu.mn

Ц.Энхтөр(Ph.D)
Мэдээллийн сүлжээ, аюулгүй байдлын салбар
ШУТИС, Мэдээлэл, холбооны технологийн сургууль
Улаанбаатар, Монгол улс
enkhtur@must.edu.mn

Х.Уянгаа
Мэдээллийн сүлжээ, аюулгүй байдлын салбар
ШУТИС, Мэдээлэл, холбооны технологийн сургууль
Улаанбаатар, Монгол улс
uyangaakh@must.edu.mn

Хураангуй—2024 оны эхний хагаст DDoS халдлагын тоо өмнөх оны мөн үетэй харьцуулахад 46%-иар нэмэгдсэн бөгөөд Gcore Radar-ийн тайланд тэмдэглэснээр тоглоом болон технологийн салбар хамгийн их халдагдсан байна. Уламжлалт халдлага илрүүлэх систем (IDS)-үүд болох Snort, Suricata нь урьдчилан тодорхойлсон гарын үсгийн дүрмэнд тулгуурладаг тул шинэ, хувирсан халдлагуудыг илрүүлэхэд хязгаарлагдмал байдаг. Энэ дутагдлыг нөхөхийн тулд сүлжээний урсгалаас суралцах чадвартай машин сургалт (ML) болон гүн сургалт (DL)-ын аргачлалууд IDS-д нэвтэрч эхэлсэн нь судалгааны хамгийн идэвхтэй чиглэл болсон байна. Гүн сургалтад суурилсан IDS-ийн судалгааны хувийн жин 2016 онд бараг 0%-аас 2024 он гэхэд 65.7%-д хүрсэн нь энэ чиглэлийн хурдацтай өсөлтийг харуулж байна. Энэхүү судалгааны ажилд машин сургалтад суурилсан халдлага илрүүлэх системийн аргачлалуудыг судалж, eBPF (Extended Berkeley Packet Filter) болон XDP (eXpress Data Path) технологийг машин сургалтын Random Forest алгоритм болон CIC-DDoS2019 өгөгдлийн санг ашиглан хослуулсан бодит цагийн IDS/IPS системийг хэрэгжүүлж, уламжлалт халдлага илрүүлэх системээс хурдан ба санах ойн зарцуулалт болон цөмийн ашиглалт илт багассан үр дүнг гаргасан.

Түлхүүр үгс—ML, IDS, IPS, eBPF, XDP

I. Удиртгал

Орчин үеийн мэдээллийн технологийн дэд бүтэц нь интернетийн хурдацтай хөгжилтэй зэрэгцэн кибер аюулгүй байдлын чиглэлд ноцтой сорилтуудтай тулгарч байна. Европын Кибер Аюулгүй Байдлын Агентлаг (ENISA)-ийн 2024 оны тайланд дурдснаар 2023 оны 7 дугаар сараас 2024 оны 6 дугаар сарын хооронд дэлхий даяарх халдлагад нийтийн захиргаа (19%), тээвэр (11%), санхүүгийн салбар (9%) хамгийн их өртсөн бөгөөд DDoS халдлага нь бүх салбарыг хамарсан гол аюул хэвээр байсан болно [14]. 2023 онд 100Гбит/с-аас 250Гбит/с-ын өндөр хурдтай DDoS халдлага өмнөх оноос 177%-иар нэмэгдсэн бөгөөд 500Гбит/с-аас их хэмжээний томоохон халдлага 150%-иар өссөн байна. Дэлхийн бүх DDoS халдлагын бараг тал хувь нь Америкийг чиглэсэн байжээ [2]. Эдгээр өгөгдлүүд нь сүлжээний аюулгүй байдлын систем тасралтгүй хөгжүүлэх шаардлагатайг тодорхой харуулж байна.

Халдлага илрүүлэх систем (IDS) нь сүлжээний урсгалыг байнга хянаж, хортой үйл ажиллагааг илрүүлдэг кибер аюулгүй байдлын суурь бүрэлдэхүүн хэсэг юм. IDS нь сүлжээний урсгалыг хянан шинжлэх замаар хортой үйл ажиллагааны шинж тэмдгийг илрүүлж, аюулгүй байдлын заналхийлэлд хариу арга хэмжээ авахад чиглэсэн кибер аюулгүй байдлын дэд бүтцийн тэргүүлэх бүрэлдэхүүн хэсэг бөгөөд хиймэл оюун ухаан, машин сургалтын дэвшилтэт аргачлалууд IDS-ийн гүйцэтгэлийг мэдэгдэхүйц сайжруулсан байна [4].

Уламжлалт гарын үсэгт суурилсан IDS-үүд нь тодорхой дүрэм, хэв маягийн мэдээллийн санг ашигладаг тул урьд мэдэгдэж байсан халдлагыг илрүүлэхэд үр дүнтэй байдаг. Гэсэн хэдий ч шинэ болон хувирсан (zero-day) халдлагын эсрэг хамгаалалт дутмаг байх нь эдгээр системийн гол дутагдал юм. Энэ хязгаарлалтыг даван туулахын тулд машин сургалтын аргачлалуудыг ашигласан аномалийн илрүүлэлт (anomaly-based detection) улам бүр түгэн дэлгэрч байгаа бөгөөд энэ нь урьд тодорхойлогдоогүй халдлагуудыг ч хэв маягаас нь тодорхойлох боломжийг бүрдүүлдэг. eBPF нь Linux цөмд байгаа зааврын багц болон гүйцэтгэлийн орчин юм. eBPF нь XDP (eXpress Data Path)-г програмчлах боломжийг олгодог бөгөөд XDP нь сүлжээний интерфейсийн карт (NIC)-д ойр байрлах цөмийн сүлжээний давхаргад пакетийг хурдан боловсруулдаг. Энэ технологийг Facebook, Cloudflare, Netronome зэрэг томоохон компаниуд сүлжээний мониторинг, ачааллын тэнцвэржүүлэлт, аюулгүй байдлын гүйцэтгэлд өргөн

ашигладаг [7].

eBPF/XDP-г ашигласан хурдан пакет боловсруулалтын гүйцэтгэлийн давуу талыг судлаачид олон судалгаандаа тэмдэглэсэн бөгөөд уламжлалт аргуудтай харьцуулахад нөөцийн хэмнэлт болон гүйцэтгэлийн давуу тал нь мэдэгдэхүйц байдаг [9].

II. Сэдэл болон зорилго

Сүлжээний аюулгүй байдлын орчин улам бүр хурдацтай хувирч байгаа нь уламжлалт гарын үсэгт суурилсан IDS системүүдийн хязгаарлагдмал байдлыг илүү тодоор харуулж байна. Шифрлэгдсэн өгөгдлийн хэрэглээ, шинэ протоколуудын олон янз байдал, дэлхий даяар хортой үйл ажиллагааны тоо нэмэгдэж байгаа нь IDS системүүдэд шинэ сорилт тулгарч байна. Одоо байгаа гарын үсэгт суурилсан IDS нь үл мэдэгдэх хортой үйл ажиллагааг хэв маягт нь тулгуурлан илрүүлж чадахгүй байгаагаас гадна машин сургалтад суурилсан IDS нь гарын үсэгт суурилсан IDS-ийг шинэ хортой үйл ажиллагааг тодорхойлоход мэдэгдэхүйц давж гарч байгааг судалгаанууд харуулж байна.

Халдлага илрүүлэх системийн хөгжлийн гурван үе шат байдаг: эхний үед гарын үсгийн хэв маяг ашигладаг байсан бол, хоёр дахь үед аномалийн илрүүлэлт нэвтэрсэн боловч буруу дохиоллын хувь өндөр байсан, харин гуравдах үед машин сургалт болон гүн сургалт нэвтэрч эхэлсний дараа системүүд тохируулан суралцах чадвартай болсон. Тус судалгааны сэдэл нь кибер аюулгүй байдлын заналхийллүүдийн нарийн төвөгтэй байдал нэмэгдэж байгаа, уламжлалт IDS аргачлалуудын дутагдал, мөн урьдчилан сэргийлэхэд чиглэсэн аюулгүй байдлын хандлагыг нэвтрүүлэх хэрэгцээтэй холбоотой байна.

Техникийн хувьд eBPF/XDP технологийн нэвтрэлт нь ML-IDS-ийн гүйцэтгэлийн тэнцвэрийн асуудлыг шийдвэрлэх шинэ боломж нээсэн. Уламжлалт пакет барих аргачлалуудад (tcpdump гэх мэт) шинж чанарыг гаргаж авах нь бүх пакет ирж дуустал хүлээх шаардлагатай байдаг. Харин XDP нь пакет бүрийг хүлээн авсан даруй боловсруулах программчлалын боломж олгодог тул шинж чанарыг гаргах хугацааг нэлээд богиносгодог.

A. Энэхүү судалгааны гол сэдэл

Нэгдүгээрт, гүйцэтгэлийн хэт өртөг. Snort, Suricata зэрэг уламжлалт IDS-үүд нь хэрэглэгчийн орчинд ажилладаг тул контекст шилжүүлэлт (context switch), буфер хуваарилалт зэрэг нэмэлт ачааллыг авч явдаг. Өндөр хурдны сүлжээнд энэ нь пакет алдалтад хүргэдэг бөгөөд аюулгүй байдлын тасалдлыг бий болгодог. Хоёрдугаарт, zero-day халдлагад эмзэг байдал. гарын үсгийн мэдээллийн санд байхгүй шинэ халдлагуудыг уламжлалт IDS илрүүлж чадахгүй. ML болон гүн сургалт зэрэг дэвшилтэт технологиудыг IDS-д нэгтгэснээр дасан зохицох чадвар, иж бүрэн мониторинг болон zero-day халдлагыг илрүүлж, бууруулах чадварыг хангадаг болохыг Springer Discover Artificial Intelligence дээр нийтлэгдсэн иж бүрэн судалгаанд онцлон тэмдэглэсэн.

Энэхүү ажлын ерөнхий зорилго нь eBPF/XDP технологийг Random Forest машин сургалтын загвартай хослуулсан бодит цагийн DDoS халдлага илрүүлэх болон блоклодог систем боловсруулж, Ubuntu 20.04 орчинд хэрэгжүүлэх явдал юм.
Тодорхой зорилтууд:

- ML-д суурилсан IDS/IPS системийн аргачлалуудын иж бүрэн судалгаа хийж, уламжлалт ML алгоритм болон гүн сургалтын загварууд, eBPF/XDP-д суурилсан системүүдийн харьцуулалтын үнэлгээ гаргах.
- CIC-DDoS2019 benchmark өгөгдлийн санг ашиглан Random Forest алгоритмыг сургах, классын тэнцвэргүй байдлыг SMOTE аргаар шийдвэрлэх, загварын нарийвчлал болон F1-оноог үнэлэх.
- Linux цөмийн XDP hook-д eBPF программ боловсруулж, пакетийн шинж чанарыг (src/dst IP, протокол, пакетийн хэмжээ, inter-arrival хугацаа) BPF ring buffer ашиглан хэрэглэгчийн орчинд дамжуулах.
- Хэрэглэгчийн орчинд ажиллах Python daemon хэрэгжүүлж, ring buffer-ийн мэдээллийг Random Forest загвараар ангилан, DDoS гэж таньсан урсгалыг BPF blocklist map-д бичиж, kernel-д буцааж илгээх.

III. Өмнө судлагдсан ажлууд

Сүлжээний халдлага илрүүлэхэд машин сургалтыг ашиглах нь 2000-аад оноос эхэлсэн судалгааны чиглэл боловч сүүлийн жилүүдэд гүн сургалтын аргачлалуудын хурдацтай хөгжил, шинэ benchmark өгөгдлийн сангуудын гарч ирснээр энэ чиглэл дахин сэргэн эрчимжсэн байна.

A. Уламжлалт ML алгоритм дээр суурилсан судалгаанууд

Pinto (2023) MDPI Sensors дээр нийтэлсэн судалгаандаа чухал дэд бүтцийн хамгаалалтад зориулсан ML-д суурилсан IDS системүүдийн иж бүрэн тойм дүгнэлт хийсэн бөгөөд үйлдвэрийн удирдлагын систем (ICS), SCADA болон тархмал удирдлагын системд чиглэсэн халдлагуудыг илрүүлэхэд машин сургалт хэрхэн ашиглагдаж байгааг задлан шинжилсэн байна [?].

Akers болон Salman нарын 2024 оны судалгаанд таван ансамбль алгоритм буюу Random Forest, AdaBoost, XGBoost, LightGBM, CatBoost-ыг IoT-д зориулсан IDS-д ашигласны дараа XGBoost нь CIC-IDS2017 өгөгдлийн сан дээр 99.91% нарийвчлал, 99.88% AUC-ROC-г харуулсан нь бусад бүх алгоритмуудыг давсан байна. Судалгаанд мөн загварын тайлбарлах чадварыг сайжруулахын тулд дүрмийн дедукцийн аргыг

нэмж ашигласан нь практикт хэрэгжүүлэх боломжийг нэмэгдүүлсэн байна [?]. Springer-ийн Journal of Network and Systems Management дээр нийтлэгдсэн судалгаанд (2024) CICIDS-2017 өгөгдлийн сан дээр Random Forest классификаторыг permutation feature importance аргачлалтай хослуулан ашигласны үр дүнд 26 гол шинж чанарыг тодорхойлж, жинлэгдсэн F1-оноог 99.8%, макро F1-оноог 93.31%-д хүргэж, уламжлалт арга болон ML-ийн аль алиныг давсан байна [15]. ScienceDirect дээр нийтлэгдсэн IoT орчинд зориулсан судалгаанд (2025) CICIDS2017 дээр шинж чанар сонголттой XGBoost нь 99.5% нарийвчлал, 99.99% precision, 99.91% recall-г харуулсан бөгөөд гүйцэтгэлийн хувьд LightGBM болон Random Forest-ийг давсан байна. Тус судалгаа нь 40 шинж чанарыг ашигласнаар загварын бодит цагт ажиллах боломжийг нэмэгдүүлсэн гэж онцолжээ [16].

B. eBPF/XDP болон ML-ийн хослолд суурилсан судалгаанууд

Bachl, Fabini, Zseby нарын 2021–2022 онд ArXiv-д нийтэлсэн "A flow-based IDS using Machine Learning in eBPF"бүтээл нь энэ чиглэлийн суурь ажил болсон юм. Шийдэх мод (Decision Tree) алгоритмыг бүрэн eBPF дотор хэрэгжүүлж, урсгалд суурилсан IDS-ийг хэрэгжүүлэх боломжтой болохыг нотолсон бөгөөд хэрэглэгчийн орчинд ажиллах ижил хэрэгжүүлтийг 20%-иас дээш гүйцэтгэлээр давсан байна. eBPF-д ML загвар байрлуулахад тулгардаг stack хэмжээ, зааврын тоо, хөвөгч цэгийн арифметикийн хязгаарлалтуудыг тодорхойлж, шийдлийг санал болгосон нь цаашдын судалгааны чиглэлийг тавьсан гэж үзэж болно [8].

Hadi-гийн 2024 оны "iKern"системд eBPF болон Virtual Network Functions (VNF)-ийг хослуулан цөмийн түвшинд пакетийг шалгадаг IDPS-ийг боловсруулсан. 1Гбит/с сүлжээнд туршихад iKern нь TCP SYN пакетийн 98.9%, UDP пакетийн 97.4%, DNS өсгөлтийн пакетийн 99.4%-ийг амжилттай шүүж унагаасан байна. Тус систем нь PF_RING socket ашиглан олон цөм дэх пакет дамжуулалтыг тэнцвэржүүлж, ачааллын пик үед buffer overflow-оос ангид байсан нь Snort, Suricata зэрэг уламжлалт IDPS-ийн гол сул тал болох дохиолол алдалтыг шийдсэн гэж үзжээ [10].

DSN 2024 бага хурлаар нийтлэгдсэн NN-eBPF судалгаанд XDP-г ашиглан бодит цагийн халдлага илрүүлэх болон урьдчилан сэргийлэх системийг хэрэгжүүлсэн. eBPF-ийн бүхэл тооны арифметикийн хязгаарлалтыг шийдвэрлэхийн тулд хэрэглэгчийн орчинд сургасан загварыг квантизаци хийж цөмд байрлуулах шинэ арга нэвтрүүлсэн. Офлайн болон онлайн өгөгдлийн санд тус тус 0.933 болон 0.992 F1-оноог харуулж, санах ойн зарцуулалтыг 5KB хүртэл бууруулж, нэг урсгалд 3000–5000нс дүгнэлтийн хугацааг хангасан байна [11].

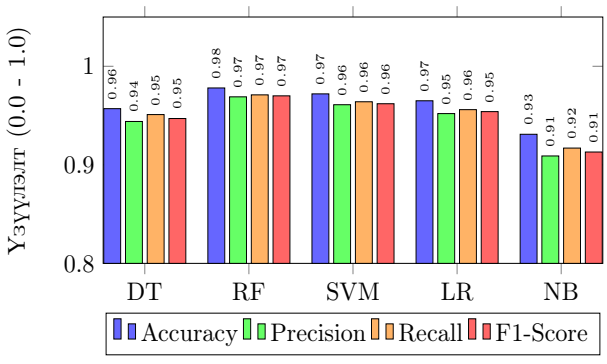
IV. Машин сургалтын аргууд ба алгоритм сонголт

A. Машин сургалтын аргачлалуудын ерөнхий тойм

Машин сургалтын алгоритмуудыг IDS-д ашиглах нь суралцах аргачлалын дагуу гурван үндсэн ангилалд хуваагддаг: хяналттай сургалт, хяналтгүй сургалт болон хагас хяналттай сургалт. Хяналттай сургалт нь шошголсон өгөгдлийн санд тулгуурлан хэвийн болон хортой урсгалыг ялгаж сурдаг бол хяналтгүй сургалт нь аномалийг урьдчилсан шошго шаардалгүйгээр хэв маягт суурилан илрүүлдэг. Гүн сургалт нь IDS-ийн орчинд ялангуяа CNN болон LSTM загварууд нь нарийн хэв маяг болон аномалиудыг автоматаар сурч, уламжлалт систем илрүүлж чадахгүй байгаа халдлагуудыг ч олж тогтоодог.

AI-д суурилсан IDS системүүдийн дүн шинжилгээнд дурдснаар, гарын үсэгт суурилсан илрүүлэлт нь зөвхөн тодорхой хэв маягийг тодорхойлж ажиглагдсан үйл явдалтай харьцуулдаг тул мэдэгдэж буй халдлагыг илрүүлэхэд үр дүнтэй боловч шинэ халдлагын эсрэг хангалтгүй байдаг. Харин аномалид суурилсан аргачлал нь хэвийн зан авирын үндсэн шугамаас хазайлтыг илрүүлдэг бөгөөд машин сургалт нь энэ аргачлалыг автоматжуулж, динамик сурах чадвараар хангадаг [11].

Зураг 1-т CIC-DDoS2019 өгөгдлийн сан дээр сургасан машин сургалтын алгоритмуудын (DT, RF, SVM, LR, NB) гүйцэтгэлийн үзүүлэлтүүдийг харьцуулан харуулсан. Энэхүү туршилтын нь сүлжээний урсгалд суурилсан DDoS халдлагыг илрүүлэхэд аль алгоритм нь илүү тохиромжтойг тодорхойлох зорилготой. Туршилтад CIC-DDoS2019 өгөгдлийн сангийн урсгалын шинж чанаруудыг ашигласан бөгөөд эдгээр нь пакетийн хэмжээ, хугацааны интервал, эх үүсвэр/хүлээн авагчийн хаяг, порт зэрэг олон хэмжээст шинжүүдээс бүрдэнэ. Сургалтад CIC-DDoS2019 өгөгдлийн сангаас 6 сая урсгал ашиглаж, 80% сургалт, 20% тест хуваалт хийсэн.



Зураг 1. Алгоритмуудын үзүүлэлтийн нэгдсэн харьцуулалт (CIC-DDoS2019)

Дээрх харьцуулалтуудаас харахад Random Forest нь нарийвчлал, дүгнэлтийн хурд, тайлбарлах чадвар бо-

лон байрлуулах тохиромжтой байдлын хувьд хамгийн оновчтой тэнцвэрийг хангадаг бөгөөд eBPF/XDP-тай хослуулсан бодит цагийн DDoS илрүүлэлтийн системд тохиромжтой сонголт болно.

B. Алгоритм сонголтын үндэслэл

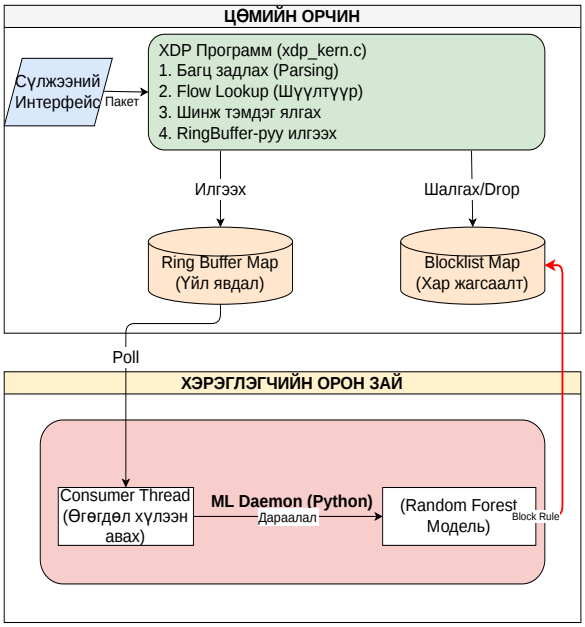
Энэхүү судалгаанд Random Forest-ийг үндсэн алгоритм болгон сонгосон нь дараах үндэслэлтэй:

- Хүснэгтэн өгөгдөлд тохирох оновчтой байдал. Сүлжээний урсгалын IDS-ийн хувьд тооцооллын болон ангиллын шинж чанарууд давамгайлдаг бөгөөд тийм бүтэцтэй хүснэгтэн өгөгдөлд RF болон XGBoost зэрэг уламжлалт ML загварууд нь LSTM, GRU, CNN зэрэг гүн сургалтын загваруудаас давж гарч байна гэдгийг ScienceDirect дээрх харьцуулалтын судалгаа баталсан байна. Сонгодог загваруудын — ялангуяа шийдэх мод болон ансамбль аргачлалуудын дотоод архитектур нь шинж чанарын чухлыг болон харилцан хамаарлыг ашигласнаар хэв маягийг үр дүнтэй таньдаг.
- Бодит цагийн байрлуулах тохиромжтой байдал. Random Forest-ийг бодит цагийн DDoS илрүүлэлтэд ашигласан туршилт дээр тархмал болон ирмэгийн тооцоолох хослол нь хурдан болон нарийн илрүүлэлтийн хэрэгцээг хангасан байна. MDPI joblib ашиглан сериализацилагдсан RF загвар нь Python sklearn-д inference дуусгах хугацаа хэдэн миллисекунд байдаг тул XDP ring buffer-ийн дамжуулалтаас хоцрохгүй ажиллах боломжтой.
- eBPF-д хялбар байрлуулах боломж. eBPF/XDP-ийн хязгаарлагдмал стек хэмжээ болон зааврын тооноос шалтгаалан CNN, LSTM зэрэг гүн сургалтын загварыг шууд kernel-д байрлуулах нь квантизаци болон нарийн техникийн ажиллагаа шаарддаг. Харин Random Forest-ийг хэрэглэгчийн орчинд байрлуулан eBPF/XDP-г шинж чанар гаргах болон урьдчилан шүүлтийн давхарга болгон ашиглах нь практикт хэрэгжүүлэхэд илүү тохиромжтой.

V. eBPF/XDP программ хөгжүүлэх нь

- Цөмийн орчин (Kernel Space): eBPF/XDP Программ
  Энэ хэсэг нь сүлжээний карт (NIC)-аас пакет орж ирэх хамгийн эхний цэг дээр ажилладаг.
  - NIC (eth0): Сүлжээний интерфэйсээр дамжин ирж буй "түүхий"пакетууд.
  - XDP Kern Программ:
    * Задлан шинжилгээ (Parsing): Ethernet, IP, болон TCP/UDP толгой мэдээллүүдийг задалж уншина.
    * Шүүлтүүр (Lookup): blocklist_map-аас тухайн урсгал хаагдсан эсэхийг шалгана. Хэрэв олдвол XDP_DROP үйлдлээр пакетыг цөмөөс цааш нэвтрүүлэхгүй.
    * Шинж тэмдэг цуглуулах (Feature Extraction): Загвард хэрэгцээтэй 5-tuple (src/dst IP, port, proto) болон пакетын хэмжээ, ирэх хугацааны зөрүүг тооцоолно.
    * Өгөгдөл дамжуулах: Цуглуулсан шинж тэмдгүүдийг ring_buffer ашиглан User Space руу илгээнэ.
- Хэрэглэгчийн орчин (User Space): ML Daemon
  Энэ хэсэгт илүү нарийн төвөгтэй аналитик болон шийдвэр гаргалт явагдана.
  - Consumer Thread: ring_buffer__poll ашиглан цөмөөс ирэх үйл явдлуудыг тасралтгүй хүлээн авч дараалалд оруулна.
  - Inference Thread: Random Forest алгоритм ашиглан тухайн урсгалыг "Хэвийн", "DDoS", эсвэл "Мэдээлэл алдагдалт"гэж ангилна.
  - Хариу үйлдэл (Feedback Loop): Хэрэв халдлага илэрвэл тухайн урсгалын мэдээллийг цөмийн blocklist_map-д буцааж бичнэ. Ингэснээр дараагийн пакет ирэхэд XDP түвшинд шууд хаагдах болно [15].

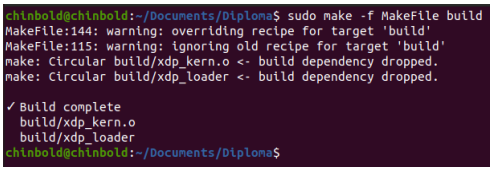
Доорх зурагт харуулсан архитектурын дагуу программ ажиллана.



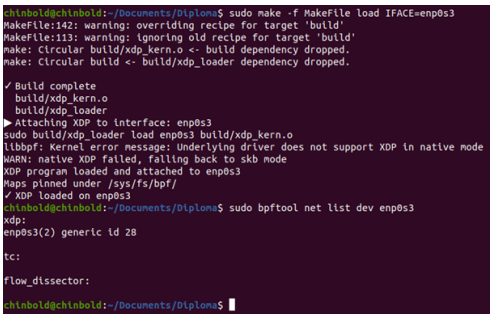
Зураг 2. Программын архитектур

Доорх хоёр зураг дээр make команд ашиглан программыг угсарч, дараа нь сүлжээний карт руу (enp0s3) XDP программыг ачаалж байна. Төхөөрөмж нь native mode дэмжихгүй байгаа тул автоматаар "skb"буюу

ерөнхий горим руу шилжин ажиллаж байгааг харж болно.



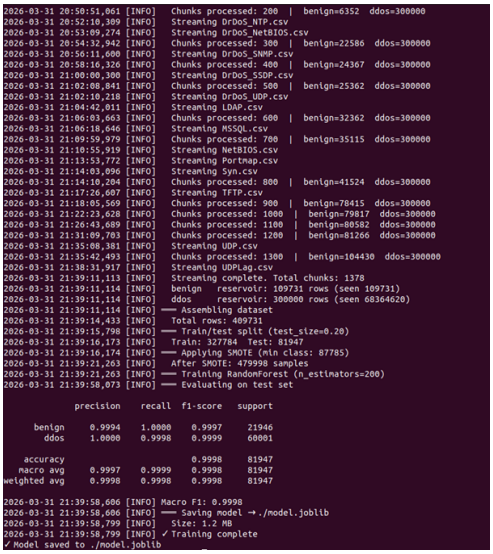
Зураг 3. XDP Программыг хөрвүүлэх



Зураг 4. Программыг сүлжээний интерфэйс дээр байрлуулах

RandomForest алгоритм ашиглан халдлагыг таних загварыг сургаж байна.

- Төрөл бүрийн DDoS халдлагуудын өгөгдөл (NTP, DNS, SNMP гэх мэт) дээр сургалт хийгдсэн.
- Үр дүн нь маш өндөр буюу 99.9% нарийвчлалтай (accuracy) гарч, загвар нь model.joblib файл болон хадгалагдсан.



Зураг 5. ML загварыг сургах

Доорх зургуудад bpftool ашиглан программын дотоод өгөгдлийг шалгаж байна.

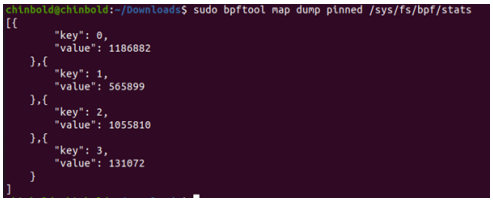


Зураг 6. Халдлага илрүүлж буй байдал

stats: Сүлжээгээр дамжиж буй пакетуудын тоо болон хаагдсан пакетуудын статистик мэдээлэл. Үүнд:

- key 0: Нийт урссан пакетийн тоо
- key 1: Халдлага эсвэл сэжигтэй урсгал гэж тодорхойлогдсон пакет.
- key 2: Мэдээлэл цуглуулсан пакет
- key 3: Хаагдсан IP хаягаас ирсэн пакетууд

Одоогийн системийн лог нь нийт пакетын тоо болон блоклогдсон пакетын статистик мэдээллийг агуулж байгаа тул халдлага илрэх хүртэлх пакетын тоог ойролцоогоор үнэлэх боломжтой. Гэвч урсгал тус бүрийн эхлэх хугацаа болон илрүүлэлтийн яг цаг хугацааг бүртгэдэггүй тул илрүүлэх хоцролт болон блоклох хугацааг нарийвчлалтай хэмжих боломж хязгаарлагдмал байна.



Зураг 7. Илрүүлсэн пакетуудын тоо

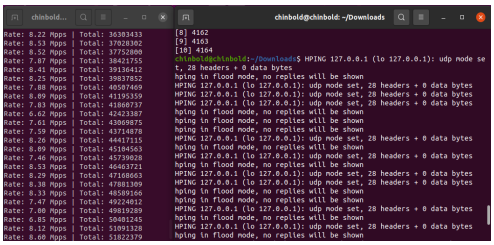
Сүлжээний урсгалаас гаргасан шинж чанаруудыг ашиглан тухайн урсгалыг ангилж, DDoS халдлага гэж тодорхойлсон тохиолдолд эх үүсвэрийн IP хаягийг eBPF-ийн blocklist map-д бичдэг. Дараагийн пакетууд XDP hook дээр ирэх үед kernel space-д blocklist map-аас уншиж, тухайн IP хаяг бүртгэлтэй байвал пакетуудыг XDP_DROP үйлдлээр шууд блоклоно. Доорх зурагт blocklist харагдаж байна. src_ip хэсэгт халдагчийн IP хаяг аравтын тооллын системээр хадгалагдсан байна.



Зураг 8. Хаалт хийсэн IP хаяг

Туршилтыг Ubuntu 20.04 үйлдлийн систем дээр, 4 цөм, 6 GB RAM бүхий орчинд гүйцэтгэсэн. Туршилтад сүлжээний ачааллыг үүсгэх зорилгоор hping хэрэгслийг

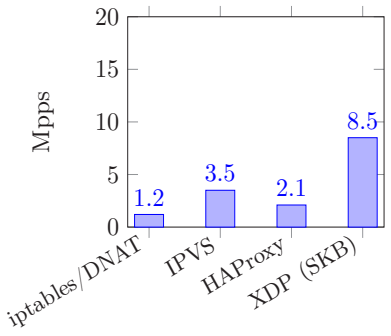
ашиглан тогтмол хурдтай пакет урсгал үүсгэж, секундэд боловсруулсан пакетийн тоо (Mpps)-г хэмжсэн.



Зураг 9. XDP хурдны хэмжилт

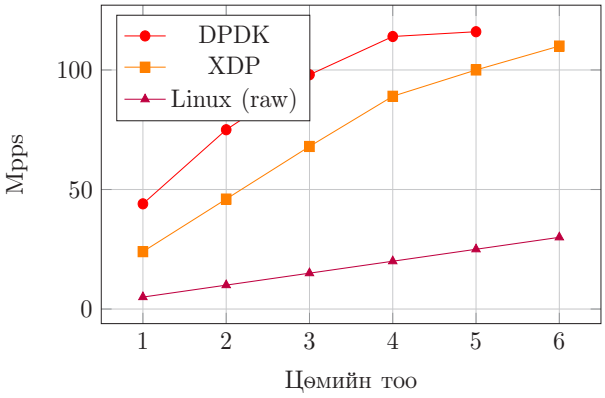
VI. Туршилтын үр дүн

Доорх диаграмм нь Линукс үйлдлийн систем дээр ачаалал тэнцвэржүүлэх янз бүрийн аргуудын гүйцэтгэлийг секундэд боловсруулсан пакетийн тоо (Mpps)-оор харьцуулсан. Туршилтыг ижил техник хангамжийн орчинд гүйцэтгэж, pktgen хэрэгслээр тогтмол хурдтай сүлжээний урсгал үүсгэн хэмжилт хийсэн. Бүх аргуудыг ижил пакетийн хэмжээ болон ижил ачааллын нөхцөлд туршсан. Хэмжилтийн үр дүнгээс харахад уламжлалт Линукс цөмийн сүлжээний стек дээр суурилсан iptables/DNAT (1.2 Mpps) болон хэрэглэгчийн орчны HAProxy (2.1 Mpps) нь хамгийн бага гүйцэтгэлтэй байв. Харин XDP (skb) горим нь 8.5 Mpps-д хүрч, бусад аргуудаас 4-7 дахин өндөр гүйцэтгэл үзүүлсэн.



Зураг 10. Ачаалал тэнцвэржүүлэгчдийн хүчин чадлын харьцуулалт [19]

Системийн CPU-ийн цөмийн тоо нэмэгдэхэд өгөгдөл дамжуулах хүчин чадал хэрхэн нэмэгдэж байгааг гурван өөр технологиор доорх графикт харьцуулсан. DPDK технологи нь нэг цөм дээр хамгийн өндөр буюу 44 Mpps-ээс эхэлж, 5 цөм дээр 116 Mpps хүртэл хамгийн хурдан өсөлт үзүүлсэн. XDP технологийн хувьд гүйцэтгэл нь цөмийн тоотой шууд хамааралтайгаар шугаман байдлаар өсөж, 6 цөм дээр 110 Mpps-д хүрч байгаа нь стандарт Линукс системээс бараг 3-4 дахин илүү хүчин чадалтай болохыг харуулж байна [19].



Зураг 11. Цөмийн тооноос хамаарсан нэвтрүүлэх чадварын харьцуулалт

VII. Дүгнэлт

Кибер аюулгүй байдлын заналхийлэл нэмэгдэж буй өнөө үед машин сургалтад суурилсан халдлага илрүүлэх систем нь уламжлалт шийдлүүдийн хязгаарлалтыг даван туулах чухал хэрэгсэл болж байна. eBPF болон XDP нь дараагийн үеийн сүлжээний дэд бүтцийн загварчлалд гол үүрэг гүйцэтгэх технологиуд бөгөөд хоцролтыг бууруулах, дэвшилтэт программчлах чадвар олгох хувьд ирээдүйтэй шийдэл болох нь iptables-тай харьцуулалтад пакет боловсруулах хугацааг 40%-иар бууруулсан туршилтаар нотлогдсон. Энэхүү судалгааны ажил нь eBPF/XDP технологийг Random Forest ML загвартай хослуулсан иж бүрэн, ажиллах боломжтой прототип системийг хэрэгжүүлэн, CIC-DDoS2019 орчин үеийн өгөгдлийн сан дээр сургаж, бодит цагт DDoS халдлагыг илрүүлж блоклодог систем бүрдүүлсэн. Гэсэн хэдий ч энэхүү систем нь урсгал тус бүрийн илрүүлэлтийн хоцролтыг нарийвчлалтай хэмжих боломжгүй байгаа нь цаашдын судалгааны чиглэл болно. Тус судалгаа нь eBPF+ML IDS-ийн чиглэлд хийгдсэн peer-reviewed нийтлэлүүдийн цоорхойг нөхөхөд хувь нэмрэн оруулж, цаашдын судалгааны суурь болж чадна гэдэгт итгэлтэй байна.

Ашигласан ном

[1] ENISA, “ENISA Threat Landscape 2024,” Европын Кибер Аюулгүй Байдлын Агентлаг, 2024. [Онлайн]. Хандсан: https://securitydelta.nl/media/com_hsd/report/690/document/ENISA-Threat-Landscape-2024.pdf [Хандсан огноо: 2026 оны 4-р сар 26 өдөр].

[2] Radware / Cisco, “2024 Global Threat Analysis Report,” Cisco, 2024. [Онлайн]. Хандсан: https://www.cisco.com/c/dam/m/en_in/events/security-conclave-2024/radware-threat-report-summary-2024.pdf [Хандсан огноо: 2026 оны 4-р сар 26 өдөр].

[3] Gcore Radar, “DDoS Attack Statistics Q1–Q2 2024,” Gcore, 2024. [Онлайн]. Хандсан: https://nordlayer.com/blog/cybersecurity-statistics-of-2024/ [Хандсан огноо: 2026 оны 4-р сар 26 өдөр].

[4] "A comprehensive survey on intrusion detection systems with advances in machine learning, deep learning and emerging cybersecurity challenges," Discover Artificial Intelligence, Springer, 2025. doi: 10.1007/s44163-025-00578-1.
[5] A. Pinto, L.-C. Herrera, Y. Donoso, and J. A. Gutierrez, "Survey on Intrusion Detection Systems Based on Machine Learning Techniques for the Protection of Critical Infrastructure," Sensors, vol. 23, no. 5, p. 2415, 2023. doi: 10.3390/s23052415.
[6] "Deep Learning-based Intrusion Detection Systems: A Survey," arXiv:2504.07839, 2025. [Онлайн]. Хандсан: https://arxiv.org/html/2504.07839v3 [Хандсан огноо: 2026 оны 4-р сар 26 өдөр].
[7] M. A. Vieira et al., "Fast Packet Processing with eBPF and XDP: Concepts, Code, Challenges, and Applications," ACM Computing Surveys, vol. 53, no. 1, pp. 1–36, 2020. doi: 10.1145/3371038.
[8] M. Bachl, J. Fabini, and T. Zseby, "A flow-based IDS using Machine Learning in eBPF," arXiv:2102.09980, 2021. [Онлайн]. Хандсан: https://arxiv.org/abs/2102.09980 [Хандсан огноо: 2026 оны 4-р сар 26 өдөр].
[9] T. Farasat, J. Kim, and J. Posegga, "SmartX Intelligent Sec: A Security Framework Based on Machine Learning and eBPF/XDP," arXiv:2410.20244, 2024. [Онлайн]. Хандсан: https://arxiv.org/abs/2410.20244 [Хандсан огноо: 2026 оны 4-р сар 26 өдөр].
[10] H. J. Hadi et al., "iKern: Advanced Intrusion Detection and Prevention at the Kernel Level Using eBPF," Technologies, vol. 12, no. 8, p. 122, 2024. doi: 10.3390/technologies12080122.
[11] "Real-Time Intrusion Detection and Prevention with Neural Network in Kernel using eBPF," DSN 2024 Proceedings, 2024. [Онлайн]. Хандсан: https://dsn2024uq.github.io/Proceedings/pdfs/DSN2024-6rvE3SSpzFYmysif75Dkid/410500a416/410500a416.pdf [Хандсан огноо: 2026 оны 4-р сар 26 өдөр].
[12] "Advanced IDS: a comparative study of datasets and machine learning algorithms for network flow-based intrusion detection systems," Applied Intelligence, Springer, 2025. [Онлайн]. Хандсан: https://link.springer.com/article/10.1007/s10489-025-06422-4 [Хандсан огноо: 2026 оны 4-р сар 26 өдөр].
[13] H. A. Ahmed, A. Hameed, and N. Z. Bawany, "Network intrusion detection using oversampling technique and machine learning algorithms," PeerJ Computer Science, vol. 8, p. e820, 2022. doi: 10.7717/peerj-cs.820.
[14] Astra Security, "100+ Cybersecurity Statistics 2024: Trends, Data and Insights," 2024. [Онлайн]. Хандсан: https://www.getastra.com/blog/security-audit/cyber-security-statistics/ [Хандсан огноо: 2026 оны 4-р сар 26 өдөр].
[15] S. S. S. Abidi et al., "Deep Learning-Based Intrusion Detection Systems for 5G and Beyond Networks: A Comprehensive Survey," Journal of Network and Systems Management, vol. 32, no. 98, 2024. doi: 10.1007/s10922-024-09874-0.
[16] "A hybrid deep learning approach for intrusion detection in Internet of Things (IoT) networks," Informatics in Medicine Unlocked, vol. 54, 2025. [Онлайн]. Хандсан: https://www.sciencedirect.com/science/article/pii/S2307187725010259 [Хандсан огноо: 2026 оны 4-р сар 26 өдөр].
[17] M. Z. Al-Faiz et al., "Enhanced network intrusion detection system using deep learning and ensemble techniques," Scientific Reports, vol. 15, no. 6241, 2025. doi: 10.1038/s41598-025-07706-y.
[18] M. Satyajeet Sharma and J. Singh, "A Feed Forward Deep Neural Network Model Using Feature Selection for Cloud Intrusion Detection System," Applied Sciences, vol. 14, no. 16, p. 6967, 2024. doi: 10.3390/app14166967.
[19] N. Dhandala, "Building a High-Performance Load Balancer with eBPF and XDP," 2026. [Онлайн]. Хандсан: https://oneuptime.com/blog/post/2026-01-07-ebpf-xdp-load-balancer/view [Хандсан огноо: 2026 оны 4-р сар 26 өдөр].

САЛБАР ХУРАЛДААН - II

Холбоо, сүлжээ

№	**ҮБХИС-ийн хичээлийн I байр, 306 тоот, Их-Эе танхим** *Хуралдаан даргалагч: Г.Отгонбаатар /МУИС/, Ю.Отгонбаатар /Хүрээ МХТДС/*
1	“Когнитив радио сүлжээний спектр мэдрэлтэнд анхдагч хэрэглэгчийн дараагийн слотын төлөвийг таамаглах нь” *Erdenebayar Lamjav and Otgonbayar Bataa*
2	“Харилцаа холбооны санал болгох системийг сайжруулах загвар, түүний туршилт” *Solongo Tumurbaatar, Bilguun Dagvadorj and Uyanga Sambuu*
3	“5G, WiFi6 сүлжээнд зориулсан давтамжийн дахин тохируулгатай антенны загварчлал” *Batjargal Enkhbold, Saruulbuyan Azkhuu, Urangua Bayanbat, Баярсайхан Пүрэвцэрэн and Пүрэвдорж Уганбаяр*
4	“FSO холбооны системийн бит алдааны магадлалын үнэлгээ” *Ulzii-Orshikh Manjaa, Khurelbaatar Tseveenjav and Munkhbayar Adiya*
5	“IPv6 шилжилтийн механизмын гүйцэтгэлийн харьцуулалт: Их сургуулийн кампус сүлжээнд” *Enkhbayar Batjargal, Bayasgalan P, Davgasuren Ts, Zolzaya Kh and Battulga Davaasambuu*
6	“Чөлөөт орон зайд ОӨМ бүхий дохиог дамжуулахад турбуленсийн нөлөөлөх нөлөөллийн судалгаа” *Namuun G and Munkhbayar Adiya*

Когнитив радио сүлжээний спектр тандалтад үндсэн хэрэглэгчийн дараагийн слотын төлөвийг таамаглах нь

Л.Эрдэнэбаяр
Холбооны инженерчлэлийн тэнхим
ШУТИС, МХТС
Монгол улс, Улаанбаатар хот
Erdenebayar.L@must.edu.mn

Б.Отгонбаяр
Холбооны инженерчлэлийн тэнхим
ШУТИС, МХТС
Монгол улс, Улаанбаатар хот
Otgonbayar.B@must.edu.mn

***Хураангуй*— Когнитив радио сүлжээнд спектрийн тандалтын уламжлалт аргууд нь ихэвчлэн тухайн агшны зурвасын төлвийг тогтоодог бөгөөд дараагийн хугацааны слот дахь төлвийг шууд таамаглахгүй. Иймээс энэ өгүүлэлд энергийн илрүүлэлт, зохицсон шүүрийн илрүүлэлт, циклостатистик шинжид суурилсан илрүүлэлтийн гаралтыг шийдвэрийн түвшний шинж болгон нэгтгэж, лицензтэй үндсэн зангилааны дараагийн слотод идэвхтэй байх магадлалыг үнэлэх асуудлыг авч үзэв. Таамагласан магадлалд тулгуурлан хоёрдогч зангилаа дамжуулах, хүлээх, эсвэл дахин тандах гэсэн гурван үйлдлийн аль нэгийг сонгох шийдвэрийн загвар байгуулж, параметрийн хайлтыг генетик алгоритм (GA), бөөмийн сүргийн оновчлол (PSO), симуляцлагдсан хөргөлт (SA), шоргоолжны колонийн оновчлолын (ACO) тасралтгүй хувилбараар тус тус хэрэгжүүлэн ижил нөхцөлд харьцууллаа. Туршилтын үр дүнгээс генетик алгоритм болон бөөмийн сүргийн оновчлол нь шинжилгээний нарийвчлалын хувьд ижил (0.7611), ROC AUC-ийн хувьд мөн ижил (0.7605) үзүүлэлттэй байсан бол генетик алгоритм сорилтын кросс энтропын утгаар бага алдагдал (0.5670) үзүүлэв. Харин бөөмийн сүргийн оновчлол дундаж ашигт үнэлгээгээр хамгийн өндөр (0.4267) гарсан нь шийдвэрийн практик ашигт чанарын хувьд давуу байж болохыг харууллаа. Үүний эсрэгээр ACO арга бүх үндсэн хэмжүүрээр харьцангуй доогуур үр дүнтэй байв. Иймээс нэг арга бүх хэмжүүрээр туйлын давуу гэсэн дүгнэлт хийхээс илүүтэй, магадлалын үнэлгээний чанар ба үйлдлийн ашигт байдлын хоорондын солбицлыг харгалзан алгоритм сонгох нь зүйтэй гэж үзэв.**

Түлхүүр үг— когнитив радио, метахэвристик арга, шинэ сүлжээ, уламжлалт арга, оновчлол

I. Удиртгал

Радио давтамжийн зурвасын статик хуваарилалт нь бодит ашиглалтын хугацааны хувьсамтгай шинжтэй үргэлж нийцдэггүй. Иймээс лицензтэй зурваст бодит сул чөлөөг илрүүлж, саад учруулахгүй нөхцөлд түр ашиглах боломжийг эрэлхийлдэг когнитив радио сүлжээ нь спектрийн үр ашигтай менежментийн чухал шийдэл болж хөгжсөн [1]–[3]. Гэвч энэ орчинд шийдвэр гаргалтын чанар нь зөвхөн тухайн мөчийн илрүүлэлтээр хязгаарлагдахгүй, харин дараагийн хугацааны алхамд зурвас хэрхэн өөрчлөгдөхийг тооцсон урьдчилсан үнэлгээнээс ихээхэн хамаарна [4], [5]. Энэ өгүүлэлд primary user (PU)-ийг хүн бус, харин тухайн давтамжийн зурвасыг эрхийн дагуу ашиглаж буй лицензтэй үндсэн радио зангилаа гэж; Хоёрдогч зангилаа (SU)-ийг үндсэн зангилаанд саад учруулахгүй нөхцөлд боломжит сул чөлөөг ашиглахыг зорьж буй хоёрдогч радио зангилаа гэж тодорхойлов.

Энергийн илрүүлэлт, зохицсон шүүрийн илрүүлэлт, циклостатистик шинжид суурилсан илрүүлэлт зэрэг уламжлалт аргууд нь хэрэгжилтийн хувьд өргөн хэрэглэгддэг боловч тэдгээрийн үндсэн зорилго нь ихэвчлэн одоогийн слот дахь зурвасын завгүй/сул төлөвийг тогтоох явдал байдаг [2], [3]. Сүүлийн үеийн судалгаанууд спектрийн хэрэглээний хугацааны хамаарал, динамик сувгийн нөхцөл, бага дохио-шуугианы харьцаа, олон хэрэглэгчтэй орчны тодорхойгүй байдлаас шалтгаалан дараагийн слотын төлвийг таамаглах асуудал улам чухал болж байгааг харуулж байна [4]–[7], [13].

Үүнтэй зэрэгцэн хиймэл оюун болон өгөгдөлд суурилсан аргууд спектрийн тандалтын судалгаанд эрчимтэй нэвтэрч, LSTM, GAN, граф нейрон сүлжээ, гүн сургалт зэрэг аргуудыг багтаасан шинэ чиглэл үүсээд байна [4], [6], [7], [14]. Гэсэн хэдий ч ижил шийдвэрийн түвшний томьёолол дээр, өөрөөр хэлбэл илрүүлэгчийн гаралт→төлөвийн вектор→дараагийн слотын идэвхийн магадлал→үйлдлийн шийдвэр гэсэн нэгэн ижил урсгалд генетик алгоритм (GA), бөөмийн сүргийн оновчлол (PSO), симуляцлагдсан хөргөлт (SA), ACO аргуудыг зэрэгцүүлэн үнэлсэн ажил хомс хэвээр байна [8]–[12].

Энэ өгүүлэл нь дээрх хоосон орон зайг нөхөх зорилгоор дараах гурван хувь нэмрийг дэвшүүлэв. Нэгдүгээрт, илрүүлэгчийн гаралтыг шийдвэрийн түвшний нэгдсэн шинж болгон байгуулж, дараагийн слотын төлөвийг таамаглах асуудлыг тодорхой томьёолов. Хоёрдугаарт, GA, PSO, SA, ACO аргуудыг ижил өгөгдөл, ижил төлөвийн орон зай, ижил үйлдлийн дүрэм, ижил үнэлгээний хэмжүүрээр харьцуулсан. Гуравдугаарт, магадлалын үнэлгээний чанар болон практик үйлдлийн ашигт байдал бүхэлдээ нэг алгоритмаар шийдэгдэхгүй байж болохыг туршилтын үр дүнгээр нотлон тайлбарлав.

II. Ижил судалгааны ажлууд

Уламжлалт судалгаанд энерги илрүүлэлт, зохицсон шүүрийн илрүүлэлт, циклостатистик шинжид суурилсан илрүүлэлт гэсэн гурван гол чиглэл зонхилж ирсэн бөгөөд эдгээр аргууд нь хэрэгжилтийн нийлмэл байдал, шаардлагатай урьдчилсан мэдээллийн хэмжээ, шуугианд тэсвэртэй байдал зэргээрээ ялгаатай [2], [3]. Энерги илрүүлэлт нь хэрэгжүүлэхэд хялбар, тооцооллын хувьд бага зардалтай тул бодит хэрэглээнд өргөн ашиглагддаг боловч шуугианы түвшний тодорхойгүй байдал болон бага SNR нөхцөлд гүйцэтгэл нь эрс мууддаг [2], [3]. Зохицсон шүүрийн илрүүлэлт нь PU дохионы бүтэц урьдчилан мэдэгдэж байгаа үед онолын хувьд өндөр

мэдрэмжтэй байж болох ч иймурьдчилсан мэдээлэл практикт үргэлж хангагддаггүй [2]. Харин циклостатистик шинжид суурилсан илрүүлэлт нь дохион дахь циклостатистик шинжийг ашигласнаар стационар шуугианаас PU дохиог ялгах чадвартай боловч тооцооллын зардал илүү өндөр байдаг [3]. Иймээс эдгээр аргуудын аль нь ч бүх нөхцөлд дангаараа хангалттай шийдэл болж чаддаггүй нь олон судалгаагаар нотлогдсон [2], [3].

Дээрх уламжлалт аргуудын нийтлэг онцлог нь тухайн агшны PU төлөвийг илрүүлэхэд чиглэдэгт оршино. Өөрөөр хэлбэл, тэдгээрийн үндсэн зорилго нь PU_t-ийн idle/active төлөвийг тогтоох бөгөөд дараагийн слотын төлөвийн талаар шууд дүгнэлт гаргахгүй [4], [5]. Гэвч бодит радио орчинд PU-ийн идэвхжил нь bursty, цаг хугацааны хамааралтай байдаг тул зөвхөн одоогийн слот илрүүлэлт хийх нь спектр хүртэмжийн хувьд реактив шийдвэр гаргалтад хүргэдэг [4]. Ийм реактив хандлага нь хоёр үндсэн сул талтай. Нэгдүгээрт, PU-ийн идэвхжил гэнэт өөрчлөгдөх үед SU нь аль хэдийн дамжуулах шийдвэр гаргасан байж болзошгүй тул интерференцийн эрсдэл өснө [5]. Хоёрдугаарт, илүү болгоомжтой хандах үүднээс тандалтын хугацааг хэт уртасгах нь нэвтрүүлэлтийн алдаа болон тандалтын алдааг нэмэгдүүлнэ [1], [4]. Иймээс сүүлийн үеийн судалгаанд спектр тандалтыг дан илрүүлэгчийн асуудал бус, харин таамаглалд суурилсан шийдвэрийн асуудал гэж авч үзэх шаардлагатай гэж үзэх болсон [4], [5].

Нөгөө талаас, когнитив радиогийн орчинд хоёрдогч зангилаанууд нэг ижил сувгийн нөөцийг ашиглахын тулд хоорондоо шууд бус өрсөлдөөн үүсгэдэг. Иймээс спектр хүртэмжийг зөвхөн дохионы боловсруулалтын асуудал гэж үзэхээс гадна шийдвэр гаргалтын асуудал гэж авч үзэх шаардлагатай. Энэ хүрээнд хамтын биш тоглоомын ойлголт ашиглагдаж ирсэн. Ийм төрлийн ажлуудад хэрэглэгч бүр өөрийн ашиг сонирхолд нийцүүлэн шийдвэр гаргах боловч тухайн шийдвэр нь бусад хэрэглэгчийн болон үндсэн хэрэглэгчийн төлөвтэй холбогддог гэж үздэг.

Өмнөх судалгаануудын нэг хэсэг нь сувгийн сонголт, дамжуулах чадал, хүлээлтийн бодлого зэрэг асуудлыг оновчлохын тулд хэвристик болон метахэвристик аргуудыг ашигласан байдаг. Ялангуяа генетик алгоритм нь параметрийн өргөн орон зайд хайлт хийх чадвартай тул спектр хуваарилалт, параметр тохируулга зэрэг асуудалд хэрэглэгдэж ирсэн. PSO нь тасралтгүй параметрийн орон зайд хурдан шийдэл олох давуу талтай учраас когнитив радио орчинд мөн өргөн хэрэглэгдсэн. SA нь локал шийдэлд гацахаас зайлсхийх боломжтой тул төвөгтэй зорилтот функцтэй асуудалд ашиглагддаг. ACO болон түүний тасралтгүй орон зайн хувилбарууд ч мөн адил нөөц хуваарилалт, замналын сонголт, параметрийн оновчлолд хэрэглэгдэж байсан.

Гэсэн хэдий ч эдгээр аргуудыг нэг ижил таамаглалын спектр тандалтын томъёолол дээр зэрэгцүүлэн авч үзсэн судалгаа харьцангуй цөөн байдаг. Өөрөөр хэлбэл, илрүүлэлтийн гаралтыг шийдвэрийн түвшний төлөв болгон нэгтгээд, үүнээс дараагийн слотын PU идэвхжлийн магадлалыг үнэлж, дараа нь Дамжуулалт, Хүлээх, Дахин тандалт гэсэн үйлдлийн аль нэгийг сонгох асуудал дээр GA, PSO, SA, ACOR аргуудыг ижил нөхцөлөөр харьцуулсан ажил дутмаг байна. Ихэнх тохиолдолд хэвристик аргууд нь зөвхөн туслах параметр сонгох эсвэл нөөц хуваарилалт хийхэд хэрэглэгдсэнээс биш, дараагийн слотын таамаглалын үндсэн шийдэгч болгон авч үзээгүй байдаг.

III. Системийн загвар

Энэ судалгаанд когнитив радиогийн орчинд үндсэн хэрэглэгч (PU) болон хоёрдогч зангилаа (SU)-ийн харилцааг слотчилсон хугацааны загвараар авч үзсэн. Хугацааг t=1,2,…,T гэж тэмдэглэсэн дараалсан слотуудад хуваана. Слот бүрт үндсэн хэрэглэгч нь тухайн сувгийг ашиглаж байгаа эсэхээс хамааран хоёр төлвийн аль нэгэнд байна. Үүнийг

$$PU_t \in \{0,1\}$$

гэж тэмдэглэнэ. Энд PU_t=1 бол зурвас завгүй, PU_t=0 бол зурвас сул байгааг илэрхийлнэ.

Хоёрдогч зангилаа нь үндсэн хэрэглэгчийн төлвийг шууд мэдэхгүй. Харин спектрийг ажигласны үндсэн дээр илрүүлэлтүүдийн гаралтыг ашиглан шийдвэр гаргана. Энэ ажилд дараах үндсэн гаралтуудыг авч үзсэн:

- Энерги илрүүлэлтийн гаралт,
- Зохицсон шүүрийн гаралт,
- Циклостатистик шинжийн гаралт.

Эдгээр хэмжигдэхүүнүүд нь тухайн слотын спектрийн нөхцлийн талаар мэдээлэл өгнө. Гэхдээ эдгээрийг дангаар нь ашиглахын оронд шийдвэрийн түвшний төлөв үүсгэх зорилгоор нэгтгэн боловсруулна. Иймд тухайн слот дахь төлвийн векторыг дараах байдлаар тодорхойлж болно.

$$S_t \in f\left(ED_t, MF_t, Cyclo_t\right)$$

Хоёрдогч зангилаа нь PU_t-ийг шууд ажиглахгүй, харин энергийн статистик, зохицсон шүүрийн гаралт, циклостатистик шинжийн хэмжигдэхүүн зэрэг илрүүлэгчийн гаралтаас үүсгэсэн S_t төлвийн векторт тулгуурлан дараагийн слотын төлвийг үнэлнэ.

Судалгааны гол зорилго нь одоогийн төлвийг ангилах бус, харин $\hat{p}_{t+1}$ магадлалыг үнэлэх явдал юм. Энд θ нь таамаглагчийн параметрийн вектор. Энэхүү магадлал нь зөвхөн онолын үнэлгээ биш; харин дараагийн алхамд дамжуулах эсэх, хүлээх эсэх, эсвэл дахин тандалт хийх эсэх шийдвэрийн үндэс болно.

$$\hat{p}_{t+1} = P(PU_{t+1} = 1 | S_t; \theta)$$

Иймд үйлдлийн орон зайг

$$A_t \in \{\Pr oceed, Stop, Creep\}$$

гэж тодорхойлов. Хэрэв $\hat{p}_{t+1}$ бага байвал дамжуулах, өндөр байвал хүлээх, харин тодорхой бус завсрын мужид байвал дахин тандалт хийх болгоомжтой үйлдлийг сонгоно. Энэхүү томьёолол нь спектрийн боломжит сул чөлөөг ашиглах сонирхол ба үндсэн зангилаанд саад учруулах эрсдэлийн хоорондын тэнцвэрийг шийдвэрийн түвшинд авч үзэж байгаагаараа онцлог юм.

Сонгосон үйлдлийн үр дүнг үнэлэхийн тулд ашигт функц тодорхойлно. Ашигт функц нь зөвхөн зөв таамагласан эсэхийг бус, мөн практик утгаар нь ашигтай байсан эсэхийг хамтад нь илэрхийлнэ. Жишээлбэл:

- Дамжуулалт хийж, үндсэн хэрэглэгч үнэхээр сул зогсонги байвал эрэг ашиг,
- Дамжуулалт хийсэн ч дараагийн слотод PU active болсон бол сөрөг ашиг,
- Хүлээх хийж, PU active байвал эерэг ашиг,
- Дахин тандалт хийвэл бага хэмжээний зардал

өгч болно.

Иймээс энэ судалгааны үндсэн асуудал нь илрүүлэлтүүдийн гаралтуудаас бүрдсэн шийдвэрийн түвшний төлөвийн орон зай дээр үндсэн хэрэглэгчийн дараагийн слотын төлөвийг урьдчилан үнэлж, түүнд тохирсон үйлдлийг сонгох параметрийн тохиргоог олох асуудал юм. Үүнийг дараах байдлаар ерөнхийлөн бичиж болно.

$$\min_{\theta} \mathrm{K}(\theta)$$

Эсвэл

$$\max_{\theta} E[R_t]$$

Энд $\mathrm{K}(\theta)$ нь таамаглалын алдаа болон шийдвэрийн чанарыг илэрхийлэх зорилтот функц юм.

Энэхүү асуудал нь шугаман бус, олон боломжит шийдэлтэй, параметрийн орон зай төвөгтэй бүтэцтэй тул шууд аналитик аргаар шийдэхэд хүндрэлтэй. Иймээс энэ ажилд параметрийн векторыг олохын тулд генетик алгоритм, бөөмийн сүргийн оновчлол, симуляцлагдсан хөргөлт, шоргоолжны колонийн оновчлол гэсэн дөрвөн аргыг ашиглаж харьцуулсан.

IV. Оновчлолын алгоритмуудыг ашиглах нь

Энэ ажлын хүрээнд доорх аргуудын үндсэн зорилго нь таамаглагчийн параметрүүдийг оновчтой сонгох явдал юм. Өөрөөр хэлбэл, илрүүлэлтүүдийн гаралтууд болон шийдвэрийн түвшний шинжүүдээс үндсэн хэрэглэгчийн дараагийн слотын төлвийг илүү зөв таамаглахад тохирох параметрийн векторыг хайж олно. Бүх алгоритм ижил өгөгдөл, ижил төлвийн орон зай дээр ажиллах тул тэдгээрийн хоорондын ялгаа нь зөвхөн хайлтын аргаар ялгарна. Эдгээр аргууд нь:

- Genetic Algorithm (GA)
- Particle Swarm Optimization (PSO)
- Simulated Annealing (SA)
- Ant Colony Optimization (AC)

юм.

Бүх алгоритм дараах нэг ижил урсгал дээр ажиллана:

- Илрүүлэлтүүдийн гаралтууд (ED, MF, Cyclo) авах
- Шийдвэрийн түвшний төлөв S_t байгуулах
- Параметрийн вектор θ-г ашиглан p_{t+1} тооцох
- Дамжуулалт/Хүлээх/Дахин тандалт үйлдэл сонгох
- θ-г шинэчлэх

A. Генетик алгоритм

Генетик алгоритм нь олон боломжит шийдлийг нэгэн зэрэг авч үздэг популяцид суурилсан хайлтын арга юм. Генетик алгоритм энэ ажлын хүрээнд параметрийн олон хувилбарыг зэрэг шалгах үүрэгтэй оролцоно. Эхний шатанд олон өөр θ вектор үүсгэж, тус бүрийг спектр тандалт дүрмээр хэрэгжүүлнэ. Өөрөөр хэлбэл нэг θ-г нэг бүрэн тандалт, шийдвэр, үнэлгээ гэсэн байдлаар бүх популяцийг үнэлнэ. Дараа нь сайн циклийн гарсан θ-уудыг сонгож, crossover хийж параметрүүдийг хольж шинэ θ үүсгэх замаар mutation хийж зарим коэффициентийг бага зэрэг өөрчилнө. Ингэснээр GA нь шийдвэрийн түвшний mapping $S_t \rightarrow p_{t+1}$-ийг илүү зөв болгох параметрийн хослолыг аажмаар сайжруулж олдог. Өөрөөр хэлбэл GA нь global search хийж, дүрэмд хамгийн тохирох параметрийн бүтэц хайж байна.

B. PSO алгоритм

Бөөмийн сүргийн оновчлол нь параметрийн орон зайд хөдөлж буй олон “бөөм”-ийн санаанд тулгуурласан арга юм. Бөөм бүр нэг боломжит шийдлийг илэрхийлэх бөгөөд өөрийн өмнө олсон хамгийн сайн байрлал болон нийт сүргийн хамгийн сайн байрлалыг ашиглан дараагийн алхмаа шинэчилдэг. PSO нь параметрийн орон зайд чиглэлтэй хөдөлгөөн хийх үүрэгтэй оролцоно. Бөөм бүр нэг θ вектортэй ба өөрийн өмнөх хамгийн сайн θ-г санаж, нийт бөөмүүдийн хамгийн сайн θ-г давхар үнэлнэ. Үүний дагуу θ нь θ+(өөрийн туршлага)+(бусдын туршлага) гэсэн байдлаар шинэчлэгдэнэ. Ингэснээр PSO нь шийдвэр boundary (Дамжуулалт/Хүлээх/Дахин тандалт ялгах хил)-ийг илүү зөв байрлуулах параметрүүдийг хурдан олно.

C. Симуляцлагдсан хөргөлт

Симуляцлагдсан хөргөлт нь нэг шийдлийг үе шаттайгаар сайжруулдаг хайлтын арга юм. Эхлээд θ-д бага өөрчлөлт өгч шинэ θ-г дахин үнэлнэ. Хэрэв үнэлгээ сайжирвал шууд сонгож, муудвал заримдаа зөвшөөрнө. Өөрөөр хэлбэл температурын жишээгээр санавал түүний утга буурах тусам муу шийдэл зөвшөөрөх нь багасч θ тогтвортой болно. Ингэснээр SA нь суурин шийдэл орчмоос бага багаар сайжруулж, шийдвэрийн дүрмийн тогтвортой хувилбарыг олдог

D. Шоргоолжны колонийн оновчлол

Шоргоолжны колонийн аргыг тасралтгүй параметрийн орон зайд хэрэглэх хувилбар юм. Энэ аргын үндсэн санаа нь өмнөх сайн шийдлүүдийг архивт хадгалж, дараагийн шинэ шийдлүүдийг тэдгээрийн ойролцоо үүсгэхэд оршино. Шоргоолжны колонийн арга нь сайн параметрийн бүсийг төвлөрүүлж судлах үүрэгтэй оролцоно. Өмнө туршсан θ-уудыг түр хадгалж, хамгийн сайн үнэлгээ авсан θ-уудыг эрэмбэлж, шинэ θ-уудыг эдгээрийн ойролцоо үүсгэнэ. Өөрөөр хэлбэл θ_{new} нь сайн шийдлийн орчим гэж ажиллах ба ингэснээр шийдвэрийн түвшний оноолтын “сайн ажиллаж буй бүс”-ийг нарийсган хайдаг. Гэхдээ эхний сайн бүс буруу байвал, эсвэл хайлт хангалттай өргөн биш байвал гүйцэтгэл буурах магадлалтай.

V. Үр дүн

Энэ бүлэгт GA, PSO, SA, AC гэсэн дөрвөн аргын гүйцэтгэлийг харьцуулан авч үзэв. Үнэлгээнд дөрвөн

хэмжүүр ашиглав. Accuracy нь хоёр ангиллын зөв таамаглалын түвшинг, кросс энтроп нь магадлалын үнэлгээний чанарыг, ROC AUC нь ангиллын ялгах чадварыг, дундаж ашиг нь сонгосон үйлдлийн практик ашигт байдлыг тус тус илэрхийлнэ. Иймээс олон хэмжүүртэй үнэлгээ нь зөвхөн "аль нь хамгийн өндөр нарийвчлалтай вэ" гэсэн нэг талт дүгнэлтээс зайлсхийж, алгоритм бүрийн давуу ба сул талыг илүү тэнцвэртэй тайлбарлах боломж олгоно. Бүх алгоритмыг ижил өгөгдөл, ижил төлөвийн орон зай, ижил үйлдлийн дүрэм, ижил үнэлгээний хэмжүүрийн дор туршсан. Хүснэгт 1-т гарсан үр дүн нь алгоритмуудын ялгааг зөвхөн хайлтын стратегийн түвшинд харьцуулах боломж олгож байна.

TABLE I. АРГУУДЫН ХАРЬЦУУЛСАН УТГУУД

Алгоритм	Train Acc	Train CE	Train AUC	Test Acc	Test CE	Test AUC	Average Reward
GA	0.81667	0.47293	0.82528	0.76111	0.56701	0.7605	0.4
PSO	0.81667	0.47177	0.82313	0.76111	0.56834	0.7605	0.42667
SA	0.81667	0.47274	0.82268	0.76111	0.57232	0.75649	0.38667
AC	0.80952	0.53236	0.82151	0.72778	0.70203	0.74558	0.37056

Хүснэгтэд үзүүлсэн үр дүнгээс харахад GA, PSO, SA гурвын гүйцэтгэл хоорондоо нэлээд ойролцоо байгаа бол AC бусдаасаа бага зэрэг доогуур үзүүлэлттэй байна. Энэ нь эхний гурван арга энэ асуудал дээр илүү тогтвортой, харин AC нь тухайн тохиргоонд төдийлөн сайн нийцээгүй байж болохыг харуулж байна.

A. Test accuracy-ийн харьцуулалт

Хүснэгт 1-ээс харахад GA, PSO, SA гурван арга сорилтын нарийвчлалын хувьд маш ойролцоо ажилласан бөгөөд гурвуулаа 0.76111 гэсэн ижил accuracy үзүүлэв. Энэ нь хоёр ангиллын хатуу шийдвэрийн түвшинд эдгээр гурван арга ойролцоо ялгах чадвартай байгааг харуулна. Харин ACO-ын сорилтын accuracy 0.72778 болж буурсан нь тухайн асуудал дээрх параметрийн хайлтын тогтвортой байдал харьцангуй сул байсныг харуулна.

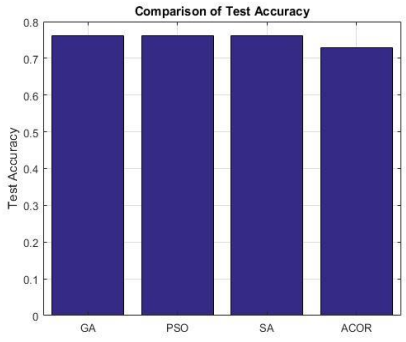


Fig. 1. Шинжилгээний нарийвчлал

Accuracy-ийн утга ижил төстэй байгаа нь GA, PSO, SA нь хоёртын шийдвэрийн түвшинд адилхан сайн ажиллаж байгааг харуулж байна. Гэхдээ зөвхөн accuracy-аар аль арга хамгийн сайн бэ гэж шууд дүгнэхэд хангалтгүй. Учир нь нэг арга зөв таамаглалыг их хийсэн ч магадлалын чанар, ашгийн хувьд өөр байж болно. Иймээс бусад хэмжүүрүүдийг хамт авч үзэх хэрэгтэй.

B. Кросс энтропын алдааны харьцуулалт

Магадлалын үнэлгээний чанарыг cross-entropy хэмжүүрээр авч үзвэл GA хамгийн бага сорилтын алдагдал (0.56701) үзүүлсэн бөгөөд PSO (0.56834)-оос үл ялиг давуу байна. SA-ийн cross-entropy 0.57232, ACOR-ынх 0.70203 байгаа нь сүүлийн арга магадлалын үнэлгээний хувьд илүү их алдаатай байгааг илтгэнэ. Өөрөөр хэлбэл GA нь зөвхөн ангиллын эцсийн шийдвэрийн түвшинд бус, таамагласан магадлалын чанарын түвшинд мөн тогтвортой үр дүн үзүүлсэн гэж тайлбарлаж болно.

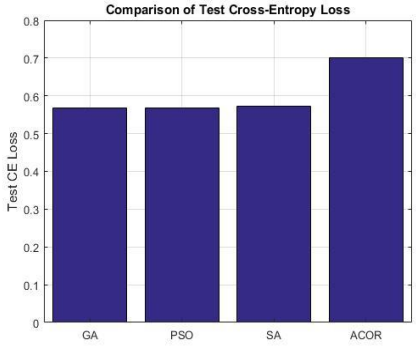


Fig. 2. Кросс энтроп шинжилгээ

PSO болон SA-ийн алдаа мөн ойролцоо тул эдгээр хоёр арга магадлалын чанарын хувьд боломжийн түвшинд ажиллаж байна гэж үзэж болно. Харин AC-ийн алдаа өндөр байгаа нь таамаглалын магадлалууд харьцангуй сул, зарим тохиолдолд хэт итгэлтэй мөртлөө буруу таамаглал өгсөн байж болохыг харуулна.

C. ROC AUC-ийн харьцуулалт

ROC AUC үзүүлэлтээр GA ба PSO сорилтын олонлог дээр ижил 0.76050 утгатай гарсан бол SA 0.75649, ACOR 0.74558 утгатай байв. Энэ нь GA ба PSO нь шийдвэрийн босго өөрчлөгдсөн нөхцөлд ч идэвхтэй ба сул төлөвийг ялгах суурь чадвараараа ижил түвшинд байгааг харуулна. Иймээс нөхцөлд "PSO нь AUC-аар илэрхий давуу" гэх мэт дүгнэлт хийх нь үндэслэлгүй бөгөөд хүснэгтийн бодит тоон үр дүнтэй нийцэхгүй.

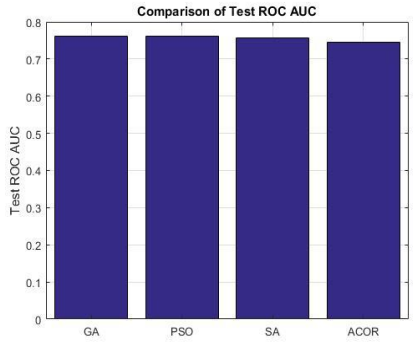


Fig. 3. ROC AUC шинжилгээ

Энэ нь PSO алгоритм шийдвэр босго өөрчлөгдсөн ч идэвхтэй болон сул зогсонги төлвийг харьцангуй сайн

ялгаж чаддаг гэсэн үг. Харин GA accuracy болон loss-ийн хувьд сайн байсан ч ROC AUC-аар PSO-оос үл ялиг доогуур гарсан байна. SA дундаж түвшний ялгах чадвартай, ACOR энэ тал дээр мөн сул байгааг үр дүн харуулж байна.

D. Дундаж үнэлгээний харьцуулалт

Дундаж ашигт үнэлгээгээр тайлбарлавал PSO хамгийн өндөр 0.42667 утгатай, GA 0.40000, SA 0.38667, ACOR 0.37056 байна. Иймээс бодлогын шийдвэрийн ашигт байдлыг илүү чухалчилсан тохиолдолд PSO давуу сонголт байж болно. Харин магадлалын үнэлгээний чанарыг илүү чухалчилбал GA бага кросс энтропигээр давуу талтай. Энэ нь нэг алгоритм бүх хэмжүүрээр туйлын давамгайлна гэхээс илүүтэй, сонголт нь зорилтот хэрэглээний шалгуураас хамаарна гэсэн илүү хязгаарлагдмал боловч нотолгоонд суурилсан дүгнэлт хийх шаардлагатайг харуулж байна.

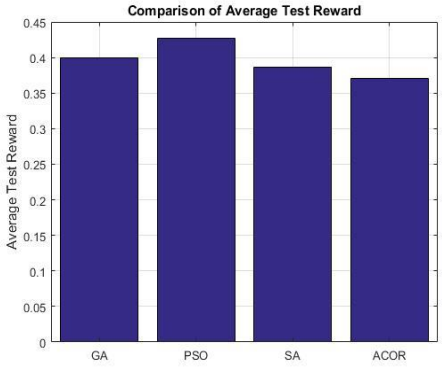


Fig. 4. Дундаж ашгийн шинжилгээ

Энэ судалгааны ажил нь хамтын спектр тандалтын бүхэл бүтэн нэгдмэл архитектур, мэдээлэл дамжуулалтын бүтэц, эсвэл олон төвтэй хамтын шийдвэрийн протоколыг иж бүрнээр шинжилсэн судалгаа биш; харин илрүүлэгчийн гаралтуудаас бүрдсэн шийдвэрийн түвшний төлөв дээр дараагийн слотын төлөвийг таамаглах харьцуулсан загварчлалд төвлөрсөн судалгаа юм.

VI. Дүгнэлт

Энэ өгүүлэлд когнитив радио сүлжээнд спектрийн төлөвийг одоогийн мөчид илрүүлэхээс илүүтэйгээр дараагийн слотын төлөвийг урьдчилан таамаглах асуудлыг авч үзэж, илрүүлэгчийн гаралтад тулгуурласан шийдвэрийн түвшний төлвийн загварыг санал болгов. Уг загвар дээр GA, PSO, SA, ACO гэсэн дөрвөн мэтаэвристик аргыг ижил нөхцөлд харьцуулан үнэлэв.

Туршилтын үр дүнгээс GA ба PSO хоёр сорилтын accuracy болон ROC AUC хэмжүүрээр ижил түвшний үр дүн үзүүлсэн бол GA магадлалын үнэлгээний чанарын хувьд бага зэрэг давуу, PSO ашигт үнэлгээгээр илүү сайн байв. SA тогтвортой боловч дээрх хоёр аргын түвшинд хүрээгүй, ACO тухайн тохиргоонд хамгийн сул үзүүлэлттэй гарлаа. Иймээс алгоритмын сонголтыг бүх хэмжүүрийг үл харгалзан ерөнхийлөхөөс илүүтэй, бодит хэрэглээ аль хэмжүүрийг тэргүүлэх шалгуур болгохоос хамааран хийх нь илүү үндэслэлтэй гэж дүгнэв.

Ингэснээр энэхүү судалгаа нь шийдвэрийн түвшний шинжид суурилсан дараагийн слотын спектр илрүүлэлтийн асуудалд мэтаэвристик аргуудын харьцуулсан дүр зургийг тодруулж, магадлалын чанар ба үйлдлийн ашигт байдлын хоорондын солбицлыг илүү нягт авч үзэх шаардлагатайг харуулсан гэж үзэж байна.

Ном зүй

[1] S. Haykin, “Cognitive radio: Brain-empowered wireless communications,” IEEE Journal on Selected Areas in Communications, vol. 23, no. 2, pp. 201–220, 2005.

[2] T. Yücek and H. Arslan, “A survey of spectrum sensing algorithms for cognitive radio applications,” IEEE Communications Surveys & Tutorials, vol. 11, no. 1, pp. 116–130, 2009.

[3] A. Nasser, H. Al Haj Hassan, J. Abou Chaaya, A. Mansour, and K. C. Yao, “Spectrum sensing for cognitive radio: Recent advances and future challenges,” Sensors, vol. 21, no. 7, Art. no. 2408, 2021.

[4] M. G. Sumithra and M. Suriya, “Improved spectrum prediction model for cognitive radio networks using hybrid deep learning technique,” International Journal of Intelligent Networks, 2024, doi: 10.1016/j.ijin.2024.05.003.

[5] K. Lahrouni, H. Semlali, G. Andrieux, J.-F. Diouris, and A. Ghammaz, “A systematic literature review on spectrum detection for Cognitive Radio-Internet of Things networks,” Ad Hoc Networks, vol. 166, Art. no. 103857, 2025, doi: 10.1016/j.adhoc.2025.103857.

[6] M. Falco, A. Pagano, and D. Croce, “AI-driven spectrum sensing: An in-depth meta-analysis of trends, challenges and opportunities,” Computer Networks, vol. 275, Art. no. 111789, 2026, doi: 10.1016/j.comnet.2025.111789.

[7] P. Yi and Y.-C. Liang, “Collaborative Spectrum Sensing in Cognitive and Intelligent Wireless Networks: An Artificial Intelligence Perspective,” arXiv:2602.09615, 2026, doi: 10.48550/arXiv.2602.09615.

[8] X. Xing, T. Jing, W. Cheng, Y. Huo, and X. Cheng, “Spectrum prediction in cognitive radio networks,” IEEE Wireless Communications, vol. 20, no. 2, pp. 90–96, 2013.

[9] J. Elhachmi and Z. Guennoun, “Cognitive radio spectrum allocation using genetic algorithm,” EURASIP Journal on Wireless Communications and Networking, Art. no. 133, 2016.

[10] A. E. Paschos, V. M. Kapinas, G. D. Ntouni, L. J. Hadjileontiadis, and G. K. Karagiannidis, “Dynamic spectrum sensing through accelerated particle swarm optimization,” arXiv:1510.03840, 2015.

[11] K. Kaur and A. K. Sharma, “Optimization of cognitive radio system using simulated annealing,” Wireless Personal Communications, vol. 72, no. 2, pp. 1025–1045, 2013.

[12] L. Liu, S. Zhang, and H. Zhang, “A novel spectrum scheduling scheme with ant colony optimization algorithm in cognitive radio networks,” Algorithms, vol. 11, no. 2, Art. no. 16, 2018.

[13] M. Falco, A. Scarvaglieri, F. Busacca, F. Nosrati, and D. Croce, “The Evolution of Dynamic Spectrum Sensing: A Two-Decade Survey from Foundations to Frontiers,” Computer Networks, 2026, doi: 10.1016/j.comnet.2026.112095.

[14] G. Narmadha, M. Jeyalakshmi, M. Ponnrajakumari, N. Duraichi, and B. Sakthivel, “Enhancing spectrum prediction in cognitive radio networks using an optimized generative adversarial network,” Results in Engineering, vol. 26, Art. no. 105270, 2025, doi: 10.1016/j.rineng.2025.105270.

Харилцаа холбооны санал болгох системийн шийп сайжруулах загвар, түүний туршилт

Төмөрбаатарын Солонго
МУИС-ийн МТЭС-ийн Мэдээлэл, компьютерын ухааны тэнхим,
Улаанбаатар хот, Монгол,
23M1NUM0751@stud.num.edu.mn

Дагвадоржийн Билгүүн*
МУИС-ийн МТЭС-ийн Мэдээлэл, компьютерын ухааны тэнхим,
Улаанбаатар хот, Монгол,
25M1NUM0709@stud.num.edu.mn

Самбуугийн Уянга
МУИС-ийн МТЭС-ийн Мэдээлэл, компьютерын ухааны тэнхим ,
Улаанбаатар хот, Монгол,
uyanga@seas.num.edu.mn

*Холбоо барих зохиогч

Хураангуй **- Сүүлийн жилүүдэд харилцаа холбооны салбарт хэрэглэгчийн өгөгдлийн хэмжээ эрс нэмэгдэж, хэрэглэгчийг тогтоон барих болон хувь хүнд тохирсон үйлчилгээ үзүүлэх шаардлага өссөөр байна. Гэвч уламжлалт хамтын шүүлтүүрт суурилсан аргууд нь хэрэглэгч хоорондын нийгмийн харилцаа, нөлөөллийг бүрэн тусгаж чаддаггүй. Иймээс энэхүү судалгаанд Монгол Улсын харилцаа холбооны операторын дуудлагын өгөгдөлд тулгуурлан хэрэглэгч хоорондын харилцааг жигнэсэн чиглэлтэй граф G=(V,E) хэлбэрээр загварчилж, дуудлагын өгөгдөл дээр суурилсан нийгмийн сүлжээний шинжилгээ(SNA) болон машин сургалтын аргуудыг нэгтгэсэн шийдэл мөн системийн загварыг боловсруулсан. Бүлгийг "K-Clique" алгоритмаар хэрэглэгчдийг бүлэглэж, SNA хэмжүүрүүдийг загварын шинж болгон ашиглав. Үр дүнд нь Community "Random forest" загвар churn таамаглалд AUC үзүүлэлтийг статистикийн хувьд сайжруулсныг DeLong тестээр баталсан. Харин бүтээгдэхүүн санал болгох даалгаварт Gradient Boosting + SNA хослол хамгийн өндөр гүйцэтгэлийг үзүүлэв. GraphSAGE загвар нь суурь аргуудыг давсан боловч Community RF-ийг давж чадаагүй нь графын сийрэг бүтэцтэй холбоотой гэж үзэв. Эдгээр үр дүнг нэгтгэн дараагын хамгийн сайн шийдвэр гаргах системийн загварыг боловсруулж, алдагдах хэрэглэгчийг эрт илрүүлэх, бизнесийн алдагдлыг бууруулах боломжийг харууллаа.**

Түлхүүр үг: Санал болгох систем, Нийгмийн сүлжээний шинжилгээ, Хэрэглэгч алдагдалтын таамаглал, GraphSAGE, K-Clique,Машин сургалт, Харилцаа холбооны өгөгдөл

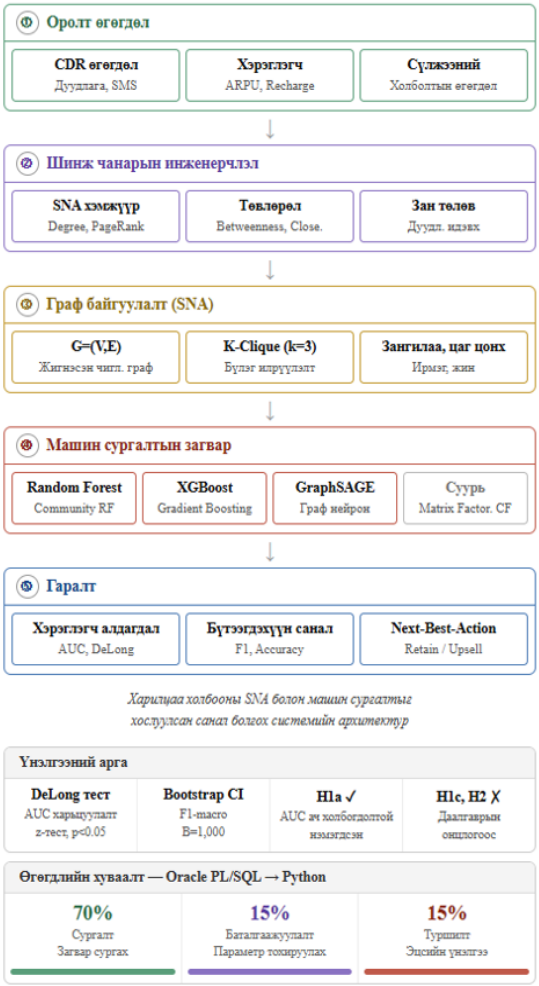
Удиртгал

Орчин үеийн харилцаа холбооны салбарт операторуудын хоорондын өрсөлдөөн эрчимжихийн хэрээр хэрэглэгчийг тогтоон барих болон тэдэнд тохирсон үйлчилгээг санал болгох нь стратегийн гол зорилт болж байна. Хэрэглэгч үйлчилгээ үзүүлэгчээ өөрчлөх буюу хэрэглэгч алдагдах үзэгдэл нь операторын орлогод шууд нөлөөлдөг бөгөөд шинэ хэрэглэгч татахаас одоо байгаа хэрэглэгчийг хадгалах нь хэд дахин бага зардалтай байдаг нь салбарын практикаас тодорхой байна[1]. Иймд хэрэглэгчийн гарах эрсдэлийг урьдчилан таамаглах болон хэрэглэгч бүрт тохирсон бүтээгдэхүүн санал болгох нь нэгдмэл байдлаар шийдвэрлэхийг шаарддаг тулгамдсан асуудал болоод байна. Уламжлалт санал болгох системүүд нь хамтын шүүлтүүр болон агуулгад суурилсан шүүлтүүрийн аргуудад тулгуурлан хэрэглэгчийн өмнөх хэрэглээний түүхэнд үндэслэн санал боловсруулдаг[2] Эдгээр аргууд нь хэрэглэгч хоорондын бодит нийгмийн харилцаа, нөлөөллийг бүрэн тусгаж чаддаггүй хязгаарлалттай хэвээр байна[3]. Нийгмийн нөлөөллийн онолоос харахад хэрэглэгчийн шийдвэр нь зөвхөн өөрийн хэрэглээний хэв маягаас бус, харилцаж буй хөрш хэрэглэгчдийн зан төлөв болон харьяалагдах нийгмийн бүлгийн нөлөөллөөс шууд хамаардаг[4]. Тухайлбал, нэг хэрэглэгч алдагдах үед тухайн хэрэглэгчийн хөршийн мөн адил шийдвэр гаргах магадлал нэмэгддэг бөгөөд ижил бүлэгт харьяалагдах хэрэглэгчид ижил төстэй бүтээгдэхүүн сонгох хандлагатай байдаг[5].

Харилцаа холбооны операторт бэлэн байдаг дуудлагын дэлгэрэнгүй бичлэгийн өгөгдөл (CDR) нь хэрэглэгчдийн хоорондын бодит харилцааны бүтцийг граф хэлбэрт буулгах боломжийг олгодог чухал эх үүсвэр юм[6]. Гэвч CDR өгөгдлийг нийгмийн сүлжээний шинжилгээ болон машин сургалтын аргатай нэгдсэн байдлаар хослуулж, хэрэглэгч алдагдах таамаглал болон бүтээгдэхүүн санал болгох хоёр даалгаврыг нэгдсэн шийдвэрийн системд нэгтгэсэн судалгаа, ялангуяа Монгол улсын хэмжээнд, хангалттай хийгдэлгүй байна[7].

Судалгааны шинэлэг тал нь гурван үндсэн чиглэлд оршино. Нэгдүгээрт, CDR өгөгдлөөс бүтээсэн граф дахь SNA хэмжүүрүүдийг ангиллын загварын оролтоор системтэйгээр ашигласан нь хамтын шүүлтүүрт аргын хязгаарыг сүлжээний бүтцийн мэдээллээр нөхөх боломжийг нотолсон. Хоёрдугаарт, хэрэглэгч алдагдах таамаглал болон бүтээгдэхүүн санал болгох хоёр даалгаврыг нэг нэгдсэн өгөгдлийн урсгал болон хамгийн тохиромжтой дараагийн үйлдлийн шийдвэрийн системд нэгтгэн хэрэгжүүлсэн нь ихэнх судалгаа эдгээрийг тусад нь авч үздэгтэй харьцуулахад практик ач холбогдол бүхий нэгдсэн framework болж байна. Гуравдугаарт, GraphSAGE загварыг уламжлалт машин сургалтын загваруудтай харьцуулан DeLong тест болон Bootstrap итгэлцлийн интервалаар статистикийн хувьд баталгаажуулсан нь нийгмийн сүлжээний мэдээллийн нөлөөлөл даалгаврын төрлөөс хамаардаг болохыг нотолж, цаашдын судалгааны чиглэлийг тодорхойлсон. Дээрх арга зүйн хандлага нь харилцаа холбооны операторуудад хэрэглэгчийн зан төлөвийг сүлжээний бүтцийн мэдээлэлтэй хослуулан шинжлэх боломжийг бүрдүүлж, шийдвэр гаргалтын үр ашгийг дээшлүүлэхэд чухал хувь нэмэр оруулна.

Иймд энэхүү судалгаанд CDR өгөгдөлд тулгуурлан хэрэглэгч хоорондын харилцааг жигнэсэн чиглэлтэй граф G=(V,E) хэлбэрт загварчилж, K-Clique алгоритмаар харилцааны бүлгүүдийг илрүүлж, Degree Centrality, Betweenness, Closeness болон PageRank зэрэг SNA хэмжүүрүүдийг машин сургалтын загварын шинж чанар болгон ашигласан. Random Forest, XGBoost болон GraphSAGE загваруудыг хэрэгжүүлж, загваруудын гүйцэтгэлийн ялгааг DeLong тест болон Bootstrap итгэлцлийн интервалаар статистикийн хувьд баталгаажуулсан [8]. Эцэст нь хэрэглэгч алдагдах таамаглал болон бүтээгдэхүүн санал болгох даалгаврыг нэгдсэн Next-Best-Action шийдвэрийн системд нэгтгэсэн. Санал болгож буй системийн архитектурыг Зураг 1-т үзүүлэв.



Харилцаа холбооны SNA болон машин сургалтыг хослуулсан санал болгох системийн архитектур

Зураг 1. SNA болон машин сургалтад суурилсан санал болгох системийн архитектур.

I. СУДЛАГДСАН БАЙДАЛ

Санал болгох систем нь хэрэглэгчийн зан төлөв, сонирхолд тулгуурлан тухайн хэрэглэгчид тохирох бүтээгдэхүүн, үйлчилгээ, контентыг санал болгох зорилготой арга, шийдлүүдийн цогц юм. Ийм системүүд нь хэрэглэгчийн шийдвэр гаргалтыг дэмжихээс гадна дижитал орчин дахь мэдээллийн илүүдлийг бууруулахад чухал үүрэг гүйцэтгэдэг. Сүүлийн жилүүдэд өгөгдлийн хэмжээ огцом өссөнтэй холбоотойгоор хэрэглэгчид өөрт хамааралтай мэдээллийг ялган сонгох шаардлага нэмэгдсэн нь системийн ач холбогдлыг улам бүр нэмэгдүүлж байна . Уг системүүдийг ашиглаж буй өгөгдөл болон алгоритмын онцлогоос нь хамааран агуулгад суурилсан болон хамтарсан шүүлтүүрт(CF) зэрэг үндсэн төрлүүдэд ангилдаг [9]. Эдгээрээс хамтарсан шүүлтүүрт арга нь хэрэглэгч-бүтээгдэхүүний харилцан үйлчлэлд тулгуурлан хэрэглэгчдийн ижил төстэй байдлыг тодорхойлж, тухайн хэрэглэгчид тохирох санал гаргадаг өргөн хэрэглэгддэг арга юм [10]. Цаашлаад матриц задлалд суурилсан хандлагууд нь хэрэглэгчийн далд сонирхлыг илүү нарийвчлалтай илрүүлэх боломжийг бүрдүүлж, CF аргын гүйцэтгэлийг сайжруулахад чухал үүрэг гүйцэтгэж байна. Гэсэн хэдий ч хамтарсан шүүлтүүрт арга нь хэрэглэгч хоорондын бодит харилцаа, нийгмийн нөлөөллийг шууд тусгадаггүй нь бодит хэрэглээний орчинд, ялангуяа харилцаа холбооны салбарт, чухал хязгаарлалт болдог [11]. Иймээс хэрэглэгч хоорондын харилцаа, нөлөөллийг харгалзан үзсэн recommendation аргуудыг судлах шаардлага үүссэн. Тухайлбал, нийгмийн итгэлцэл олон хэрэглэгч хоорондын харилцааг загварт нэгтгэсэн аргууд нь хэрэглэгчийн сонирхол, төстэй байдал болон нөлөөллийг хамтад нь авч үзэх боломжийг бүрдүүлж, санал болгох гүйцэтгэлийн чанарыг сайжруулах боломжтойг харуулсан. Эдгээр хандлагууд нь ихэвчлэн харилцааг хязгаарлагдмал байдлаар тусгаж, сүлжээний бүтцийг бүрэн ашиглаж чаддаггүй хэвээр байна. Сүүлийн үеийн судалгаанд графт суурилсан болон гүн сургалтын аргуудыг санал болгох системд өргөн ашиглах болсон. Тухайлбал, Graph Neural Network (GNN)-д суурилсан загварууд нь хэрэглэгч болон бүтээгдэхүүний хоорондын харилцааг граф бүтэц дээр илэрхийлж, нарийн төвөгтэй хамаарлыг илүү үр дүнтэй загварчлах боломжийг бүрдүүлдэг [12]. Үүний нэг жишээ болох Neural Graph Collaborative Filtering (NGCF) арга нь уламжлалт CF аргуудтай харьцуулахад илүү өндөр гүйцэтгэл үзүүлснийг судалгаанууд харуулж байна. Нөгөө талаас, SNA-д суурилсан хандлага нь хэрэглэгчдийн харилцааг граф бүтэц хэлбэрээр илэрхийлж, сүлжээн дэх байр суурь, нөлөөллийг тодорхойлох боломжийг олгодог. Ялангуяа centrality болон k-shell decomposition зэрэг зангилааны үнэлгээний аргууд нь сүлжээн дэх нөлөө бүхий хэрэглэгчдийг тодорхойлоход үр дүнтэй бөгөөд recommendation системд ашиглах боломжтой. CDR-д суурилсан хэрэглэгчдийн хамаарлын шинжилгээнд тулгуурласан судалгаанууд нь хэрэглэгч-хэрэглэгч харилцаанд суурилсан загварууд тогтвортой харилцааг ашиглан илүү үр дүнтэй санал хийх боломжтойг харуулсан. Сүүлийн жилүүдэд системд машин сургалтын аргуудыг өргөн ашиглах болсон бөгөөд эдгээр нь хэрэглэгчийн зан төлөв болон олон эх үүсвэрт өгөгдлөөс үүсэх нарийн төвөгтэй, шугаман бус хамаарлыг илрүүлэхэд илүү үр дүнтэй болохыг судалгаанууд харуулж байна [13]. Машин сургалтын аргууд нь нэг загвараас илүү тогтвортой, ерөнхийлсөн гүйцэтгэл үзүүлэх чадвартайгаараа онцлогтой. Тухайлбал, Random Forest болон Gradient Boosting зэрэг алгоритмууд нь decision tree-д суурилсан боловч сургалтын стратегиараа ялгаатай бөгөөд өөр өөр давуу талуудыг бий болгодог. Random Forest нь олон модыг параллель байдлаар сургах замаар variance-ийг бууруулж тогтвортой гүйцэтгэл үзүүлдэг бол Gradient Boosting нь алдааг дараалсан байдлаар засах замаар илүү өндөр нарийвчлалд хүрэх боломжтой. Эдгээр аргуудыг гүн сургалтын болон граф нейрон сүлжээний аргуудтай харьцуулахад илүү бага өгөгдөл дээр тогтвортой ажиллах, параметрийн тохируулга харьцангуй энгийн, мөн тайлбарлах боломж өндөр зэрэг давуу талтай [14]. Гэсэн хэдий ч дээр дурдсан судалгаануудын ихэнх нь collaborative filtering, social trust, graph-based болон machine learning аргуудыг тус тусад нь авч үзсэн бөгөөд хэрэглэгчдийн бодит харилцаа, нөлөөллийг CDR өгөгдөлд тулгуурлан нэгдсэн байдлаар нэгтгэсэн санал болгох загвар харьцангуй хязгаарлагдмал хэвээр байна. Иймээс энэхүү судалгаанд хэрэглэгчдийн харилцааг

граф бүтэц хэлбэрээр загварчилж, SNA-н хэмжигдэхүүнүүдийг хэрэглэгчийн зан төлөвийн өгөгдөлтэй нэгтгэн машин сургалтын загварт ашиглах замаар санал системийн гүйцэтгэлийг сайжруулахыг зорьж байна.

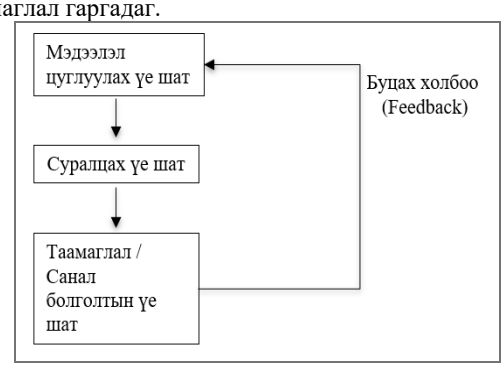
II. ОНОЛЫН ХЭСЭГ

A. Санал болгох системийн ерөнхий ойлголт

Санал болгох систем нь хэрэглэгч болон бүтээгдэхүүн, үйлчилгээний хоорондын хамаарлыг загварчлах замаар тухайн хэрэглэгчид тохирох бүтээгдэхүүнийг тодорхойлох алгоритмын систем юм. Энэхүү систем нь хэрэглэгч–бүтээгдэхүүний харилцан үйлчлэлд тулгуурлан хэрэглэгчийн сонирхлыг илэрхийлэх функц байгуулахад чиглэдэг. Санал болгох системийн үндсэн бүрэлдэхүүн хэсгүүд нь дараах байдалтай байна. Үүнд:

- Хэрэглэгч
- Бүтээгдэхүүн
- Харилцан үйлчлэл

Эдгээр бүрэлдэхүүнүүдийн хоорондын хамаарлыг ашиглан хэрэглэгчийн сонирхлыг тодорхойлж, таамаглал гаргадаг.

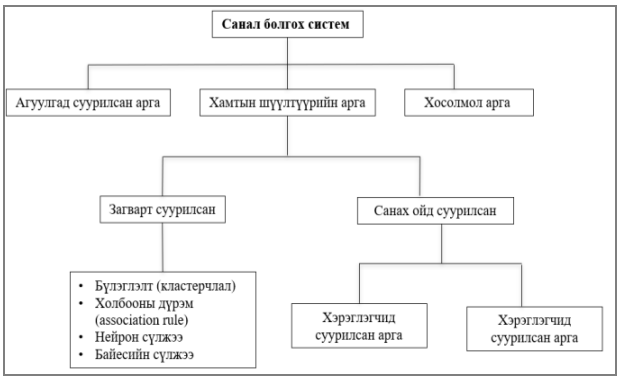


Зураг 2. Санал болгох системийн боловсруулалтын үе шат.

Зураг 2-т санал болгох системийн ерөнхий боловсруулалтын гурван үе шатыг үзүүлэв. Эхний үе шатанд хэрэглэгчийн зан төлөв, хэрэглээний түүх зэрэг шаардлагатай өгөгдлийг цуглуулж, хэрэглэгчийн профайл үүсгэнэ. Дараагийн үе шатанд цуглуулсан өгөгдөлд тулгуурлан машин сургалтын алгоритмыг ашиглан загварыг сургаж, хэрэглэгчийн сонирхол болон хэв маягийг тодорхойлно. Эцсийн үе шатанд сурсан загвараар дамжуулан хэрэглэгчид тохирох санал гаргах бөгөөд хэрэглэгчийн хариу үйлдэл буцах холбоосоор дамжин системийг тасралтгүй сайжруулж байдаг давталтат процесс юм.[15]

B. Санал болгох системийн арга зүй

Санал болгох систем нь хэрэглэгчийн өмнөх зан төлөв болон өгөгдөлд тулгуурлан тухайн хэрэглэгчид тохирох бүтээгдэхүүн, үйлчилгээ, контентыг тодорхойлох алгоритмын арга зүй юм. Ашиглагдаж буй өгөгдөл болон загварчлалын аргаас хамааран эдгээр системийг агуулгад суурилса, хамтарсан шүүлтүүрт болон хосолмол гэсэн үндсэн төрлүүдэд ангилдаг.



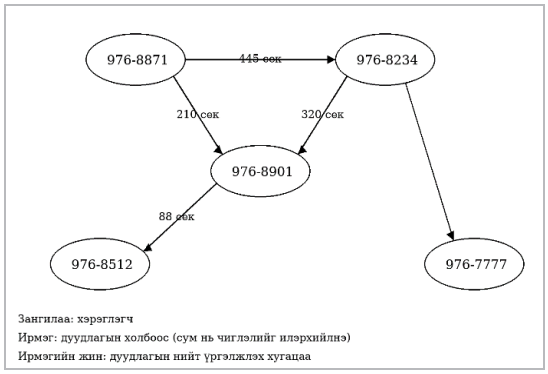
Зураг 3. Санал болгох системийн ангилал

- Агуулгад суурилсан хамтарсан шүүлтүүрийг арга нь хэрэглэгчийн өмнө сонирхож байсан бүтээгдэхүүний шинж чанарт тулгуурлан төстэй бүтээгдэхүүнийг тодорхойлж санал болгодог. Энэ нь хэрэглэгч бүрийн профайлд суурилсан тул хувь хэрэглэгчийн онцлогийг сайн тусгах давуу талтай боловч санал нэг төрлийн байх хандлагатай байдаг [16].

- Хамтарсан шүүлтүүрт санал болгох арга нь хэрэглэгч хоорондын зан төлөвийн төстэй байдлыг ашиглан санал гаргадаг бөгөөд “төстэй хэрэглэгчид төстэй сонголт хийх хандлагатай” гэсэн үндсэн зарчимд тулгуурладаг. Энэ арга нь санах ойд болон загварт суурилсан арга гэсэн хоёр хэлбэрт хуваагдана. Санах ойд суурилсан арга нь төстэй байдлын хэмжүүрүүдийг ашиглан шууд санал гаргадаг бол загварт суурилсан арга нь машин сургалтын загвар ашиглан хэрэглэгч-бүтээгдэхүүний хамаарлыг суралцдаг.

- Хосолмол арга нь агуулгад суурилсан болон хамтарсан шүүлтүүрт санал болгох аргуудыг хослуулан ашиглаж, тус бүрийн сул талыг бууруулахад чиглэдэг. Ийм хандлага нь илүү тогтвортой, оновчтой үр дүнд хүрэх боломжтой боловч загварын бүтэц болон тооцооллын нарийн төвөгтэй байдлыг нэмэгдүүлдэг.

Санал болгох системд ашиглагдах мэдээллийг ерөнхийд нь хэрэглэгчийн үйлдлийг илэрхийлэх бүтээгдэхүүний шинж чанарыг илэрхийлэх контент, мөн хэрэглэгч хоорондын харилцааг илэрхийлэх нийгмийн мэдлэгийн холбоос гэж ангилж болно [17]. Ялангуяа нийгмийн мэдлэгийн мэдлэг нягт холбоотой бөгөөд хэрэглэгч хоорондын харилцаа, нөлөөллийг граф бүтцээр илэрхийлэх боломжийг олгодог. Ийм төрлийн мэдээллийг ашиглах нь хэрэглэгчийн зан төлөвийг илүү өргөн хүрээнд ойлгож, санал болгох системийн үр дүнг сайжруулах чухал үндэс болдог.



Зураг 4: Дуудлагын бүртгэлд суурилсан хэрэглэгч хоорондын жигнэсэн чиглэлтэй графын жишээ.

C. Нийгмийн сүлжээний шинжилгээний үндсэн хэмжигдэхүүнүүд

Нийгмийн сүлжээний шинжилгээ нь хувь хүн, байгууллага болон бусад нийгмийн субъект хоорондын харилцааг граф онол, магадлал статистик болон матрицын алгебрийн аргыг ашиглан тоон хэлбэрт буулгаж, бүтэц болон динамикийг нь судлах арга зүйн нэгдсэн хүрээ юм. Энэхүү шинжилгээний онолын үндэс нь 1930-аад оны нийгэм судлалын судалгаанаас эхтэй бөгөөд J.L. Moreno судлаачийн боловсруулсан социограм аргачлалаас үүсэлтэй. Сүлжээг дараах байдлаар тодорхойлно:

$$G=(V,E) \qquad (1)$$

Энд:

- V={v1,v2,…,vn} зангилааны олонлог бөгөөд телекомын хэрэглэгч тус бүрийг нэг зангилаагаар төлөөлнө;
- E⊆V×V- ирмэгийн олонлог бөгөөд хоёр хэрэглэгчийн хоорондын харилцааг илэрхийлнэ;

Энэхүү судалгаанд сүлжээний бүтцийн шинж чанарыг илэрхийлэх үндсэн 6 метрик ашигласан. Эдгээр метрикүүдээс дурдвал:

- Зэрэглэлийн төвлөрөл нь сүлжээний шинжилгээнд хамгийн өргөн хэрэглэгддэг бөгөөд хамгийн энгийн төвлөрлийн хэмжүүрт тооцогддог. Энэ хэмжүүр нь тухайн зангилаа бусад зангилаатай хэр олон шууд холбоотой байгааг илэрхийлдэг бөгөөд хэрэглэгчийн сүлжээн дэх шууд идэвх, хүртээмжийг үнэлэхэд ашиглана.

$$C_D(v) = \frac{\deg(v)}{|V|-1} \qquad (2)$$

Энд, deg(v) нь |V| зангилаатай шууд холбогдсон хөршийн тоо, |V| нь графын нийт зангилааны тоо, |V|-1 нь нэг зангилааны хувьд боломжит дээд холболтын тоо тул хэмжүүрийг [0, 1] завсарт нормчилно.

- Зуучлалын төвлөрөл нь тухайн зангилаа бусад зангилаануудын хоорондох мэдээлэл дамжих замд хэр чухал байр суурь эзэлж байгааг хэмждэг.

$$C_B(v) = \sum_{s\neq v\neq t} \frac{\sigma\, st(v)}{\sigma\, st} \qquad (3)$$

Энд s-ээс t хүртэлх хамгийн богино замын нийт тоо, нь эдгээр замуудаас v зангилааг дайран өнгөрөх замын тоо юм. Иймд энэ харьцаа нь v зангилаа сүлжээнд хэр олон хос зангилааны хооронд зуучлагч үүрэг гүйцэтгэж байгааг илэрхийлнэ.

- Ойрын төвлөрөл нь тухайн зангилаа сүлжээний бусад зангилаанд дунджаар хэр ойр байрлаж байгааг илэрхийлдэг хэмжүүр юм. Энэ нь зангилаа мэдээлэл, мэдээ, нөлөөллийг сүлжээ даяар хэр хурдан түгээх боломжтойг үнэлэхэд чухал ач холбогдолтой.

$$C_C(v) = (|V|-1) \Big/ \sum_{u\in V, u\neq v} d(v,v) \qquad (3)$$

D. Машин сургалтын аргууд

Судалгаанд хэрэглэгчийн зан төлөв болон нийгмийн сүлжээний шинжилгээнээс гарган авсан шинж чанаруудыг нэгтгэн ангиллын асуудлыг шийдвэрлэх зорилгоор ансамбл машин сургалтын аргуудыг ашигласан. Ансамбл арга нь олон сул суралцагчийг нэгтгэх замаар загварын ерөнхийлөх чадварыг сайжруулж, өгөгдлийн хэлбэлзлээс үүсэх алдааг бууруулах давуу талтай [18].

1. Random Forest алгоритм нь шийдвэрийн мод суурилсан ансамбл арга бөгөөд сургалтын өгөгдлөөс санамсаргүй түүвэр үүсгэн олон шийдвэрийн модыг бие даан сургадаг . Эдгээр мод бүр нь өөр өөр шинж чанарын дэд багц дээр сургагддаг тул загварын хэлбэлзэл буурч, илүү тогтвортой таамаглал гаргах боломж бүрддэг. Эцсийн үр дүн нь бүх шийдвэрийн модны таамаглалыг нэгтгэх замаар тодорхойлогдоно:

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} f_t(x) \qquad (4)$$

Энд:$f_t(x)$ нь t дахь шийдвэрийн модны таамаглал, T нь нийт модны тоо юм.

2. XGBoost нь Gradient Boosting алгоритмын сайжруулсан хувилбар бөгөөд тооцооллын үр ашиг, тогтвортой байдал болон хэт тохируулалтаас сэргийлэх чадварыг нэмэгдүүлсэн орчин үеийн ансамбль арга юм. Энэхүү алгоритм нь шийдвэрийн моднуудыг дараалсан байдлаар сургах замаар өмнөх загварын алдааг үе шаттайгаар бууруулдаг боловч уламжлалт Gradient Boosting-оос ялгаатай нь регуляризаци болон хоёрдугаар эрэмбийн ойролцооллыг ашигладгаараа онцлогтой.

$$L = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \qquad (5)$$

Энд l(yi,y^i)алдааны функц, Ω(fk) нь k-р модны төвөгшлийг хянах регуляризацийн гишүүнээр илэрхийлэгдэж байна.

- GraphSAGE нь граф нейрон сүлжээний архитектур бөгөөд зангилаануудын хоорондын харилцан хамаарал, хөршийн нөлөөг шууд загварчилдаг.

$$\mathbf{m}_u^{(l)} = \mathrm{MSG}^{(l)}\left(\mathbf{h}_u^{(l-1)}\right)$$
$$\mathbf{h}_v^{(l)} = \mathrm{AGG}^{(l)}\left(\left\{\mathbf{m}_u^{(l)}, u \in N(v)\right\}\right) \quad (6)$$

GraphSAGE загвар нь зангилааны төлөөллийг хөрш зангилаануудын мэдээллийг нэгтгэх замаар давхарга бүрт шинэчилдэг. Өөрөөр хэлбэл, тухайн зангилааны хөршүүдийн эмбэддинг агрегат функцээр нэгтгэн завсрын төлөөлөл үүсгээд, үүнийг өөрийн өмнөх төлөөлөлтэй нэгтгэн жингийн матриц болон идэвхжүүлэх функцээр дамжуулж шинэ embedding-ийг тооцоолдог

Эдгээр алгоритмуудыг сонгосон нь нийгмийн сүлжээний шинжилгээнээс гарсан шинж чанарууд нь шугаман бус хамааралтай, мөн хоорондын харилцан үйлчлэл ихтэй байгаатай холбоотой. Random Forest нь өгөгдлийн хэлбэлзлийг бууруулж тогтвортой гүйцэтгэл үзүүлэх бол Gradient Boosting нь илүү нарийн хамаарлыг илрүүлэх замаар таамаглалын нарийвчлалыг нэмэгдүүлэх давуу талтай. Иймээс эдгээр аргуудыг хослуулан ашиглах нь хэрэглэгчийн зан төлөв болон нийгмийн нөлөөллийг хамтад нь загварчлахад тохиромжтой гэж үзсэн.

Е. Шинж чанарын инженерчлэл

Судалгаанд хэрэглэгчийн зан төлөв болон нийгмийн сүлжээний бүтцийн шинж чанаруудыг нэгтгэн машин сургалтын загварт ашиглах зорилгоор шинж чанарын инженерчлэлийг гүйцэтгэсэн. Шинж чанаруудыг зан төлөвт суурилсан болон сүлжээний бүтцэд суурилсан гэж хоёр үндсэн бүлэгт ангилсан. Зан төлөвийн шинж чанарууд нь хэрэглэгчийн үйлчилгээний хэрэглээ, идэвх болон харилцааны давтамжийг илэрхийлдэг бол сүлжээний шинж чанарууд нь хэрэглэгчийн нийгмийн сүлжээн дэх байр суурь, холбогдлын түвшинг тодорхойлдог. Ялангуяа зэрэглэлийн төвлөрөл болон хөршийн зэрэглэл зэрэг хэмжүүрүүд нь хэрэглэгчийн шууд холболт болон түүний орчны бүтцийн онцлогийг илэрхийлэх боломжийг олгодог бөгөөд эдгээрийг нэгтгэн ашигласнаар хэрэглэгчийн зан төлөвийг илүү өргөн хүрээнд тодорхойлох нөхцөл бүрддэг. Эцэст нь эдгээр шинж чанаруудыг нэгтгэн нэгэн төрлийн шинж чанарын вектор болгон хувиргаж машин сургалтын загварт оролт болгон ашигласан бөгөөд олон эх үүсвэрт өгөгдлийг нэгтгэн ашиглах нь загварын таамаглалын чадварыг сайжруулахад чухал үүрэг гүйцэтгэдэг [19].

V. Туршилт

Энэхүү туршилтад Монгол улсын харилцаа холбооны операторын нэргүйжүүлсэн бодит өгөгдлийг ашиглан санал болгох системийн үндсэн процессын дагуу]туршилтын судалгааг хэрэгжүүлсэн. Өгөгдлийн сан нгурван үндсэн эх үүсвэрээс бүрдэж байгаа бөгөөд эдгээр нь хэрэглэгчийн зан төлөв болон харилцааны орчны мэдээллийг олон өнцгөөс илэрхийлэх боломжийг бүрдүүлдэг. Тухайлбал, дуудлагын дэлгэрэнгүй өгөгдөл нь хэрэглэгчдийн хоорондын харилцааг, хэрэглэгчийн мэдээллийн сан нь хувь хүний зан төлөвийн шинж чанарыг, харин эдгээрээс үүсгэсэн сүлжээний граф нь харилцааны бүтцийн онцлогийг тус тус илэрхийлнэ. Судалгаанд нийт 75,701 хэрэглэгчийн 4,914,906 дуудлагын харилцаанд суурилсан 34 хувьсагч бүхий нэгдсэн өгөгдлийн санг ашигласан бөгөөд ажиглалтын хугацаа нь 2025 оны 10–11-р сарыг хамарч, дараагийн сарын хэрэглэгчийн зан төлөвийг таамаглах нэг сарын таамаглалын цонхыг тодорхойлсон [20].

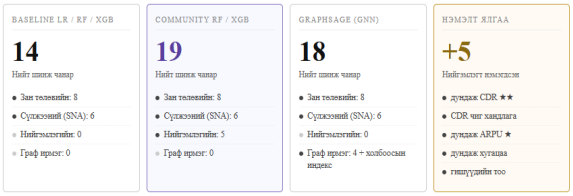
Туршилтын хүрээнд дуудлагын давтамжид тулгуурлан хэрэглэгч хоорондын харилцааг жигнэсэн чиглэлтэй граф G=(V,E) хэлбэрээр илэрхийлж, K-Clique (k=3) алгоритмаар 67,918 харилцааны бүлгийг илрүүлсэн. Зэрэглэлийн төвлөрөл, зуучлалын төвлөрөл, ойрын төвлөрөл болон хуудасны зэрэглэл зэрэг SNA хэмжүүрүүдийг хэрэглэгчийн зан төлөвийн өгөгдөлтэй нэгтгэн шинж чанарын инженерчлэл хийсэн. Дараагийн шатанд боловсруулсан шинж чанаруудыг машин сургалтын загварт оруулж ойн олон мод, XGBoost болон граф нейрон сүлжээний алгоритмуудаар хэрэглэгч алдагдах болон бүтээгдэхүүний сонирхлыг таамагласан. Хэрэглэгч алдагдалтыг зөвхөн үйлчилгээ бүрэн зогсоох байдлаар бус, хэрэглээ өмнөх сартай харьцуулахад 30%-иас дээш буурсан тохиолдлыг хамруулсан өргөтгөсөн тодорхойлолтоор тодорхойлсон нь хэрэглэгчийн зан төлөвийн өөрчлөлтийг илүү эрт илрүүлэх боломжийг бүрдүүлсэн. Судалгаанд хэрэглэгч алдагдлын хувь 5.8% байгаа нь ангиллын тэнцвэргүй байдал үүссэнийг харуулж байгаа тул загвар сургалтын үед жинлэлтийн тохиргоог ашиглан энэхүү нөлөөллийг бууруулсан. Өгөгдлийг сургалт (70%), баталгаажуулалт (15%), туршилт (15%) гэсэн харьцаагаар хувааж, цаг хугацааны дарааллыг хадгалсан байдлаар бэлтгэсэн нь өгөгдлийн алдагдлаас сэргийлэх, загварын үнэлгээг бодит нөхцөлд ойртуулах ач холбогдолтой. Туршилтын ерөнхий архитектур болон өгөгдлийн боловсруулалтын үе шатыг Зураг 1, Зураг 4-т тус тус үзүүлсэн.



Зураг 5: Өгөгдлийн боловсруулалтын үе шат

	CALLING_NUMBER	CALLER_NUMBER	CALL_DURATION	CALL_COUNT
1	8886****	9946****	392	12
2	8679****	9360****	1279	12
3	8956****	9128****	43	2
4	8902****	8802****	3268	87
5	9907****	8042****	1808	9

Хүснэгт 1: Өгөгдөлд хэрэглэгч хооронд харилцааг илэрхийлэх дуудлага хийсэн болон хүлээн авсан дугаар, дуудлагын үргэлжлэх хугацаа, давтамж зэрэг үндсэн шинж чанарууд агуулагдана.



Зураг 6: Загвар бүрийн шинж чанарын ангилал

Зураг 6-д загвар бүрийн шинж чанарын нийт тоог харьцуулан үзүүлэв. Суурь загварууд (логист регресс, ойн олон мод, XGBoost) нийт 14 шинж чанар ашигласан бөгөөд эдгээр нь зан төлөвийн 8 болон нийгмийн сүлжээний 6 шинж чанараас бүрдэнэ. Харилцааны бүлэгт суурилсан загварууд (Community RF, Community XGBoost) нийт 19 шинж чанар ашигласан бөгөөд нийгэмлэгийн 5 нэмэлт шинж чанараар баяжуулсан байна. Граф нейрон сүлжээний загвар нийт 18 шинж чанар ашигласан бөгөөд граф ирмэгийн 4 шинж чанар болон холбоосын индексийг нэмэлтээр ашигласан нь онцлог юм. Нэмэлт ялгааны баганаас харахад харилцааны бүлэгт суурилсан загварууддад нэмэгдсэн 5 шинж чанар нь дундаж CDR, CDR-ийн чиг хандлага, дундаж ARPU, дундаж хугацаа болон гишүүдийн тоо байгаа нь хэрэглэгчийн нийгмийн орчны мэдээллийг загварт тусгасан гол хувь нэмэр болж байна.

A. Сургалтын орчин

Нийгмийн сүлжээний шинжилгээ болон бүлэглэлийн тооцооллыг 8 процессор, 384 GB санах ой, RTX график процессор бүхий өндөр хүчин чадлын серверт гүйцэтгэсэн. Машин сургалтын загварын сургалт, туршилтыг Intel Core Ultra 5 125H процессор, 32 GB RAM санах ой бүхий Lenovo Gen 2 зөөврийн компьютер дээр Windows 11 Pro үйлдлийн системд хэрэгжүүлсэн. Туршилтад Python хэлийг ашиглаж, графын шинжилгээнд NetworkX, машин сургалтанд Scikit-learn, граф нейрон сүлжээний загварт PyTorch Geometric сангуудыг тус тус ашигласан.

B. Үнэлгээний үзүүлэлтүүд

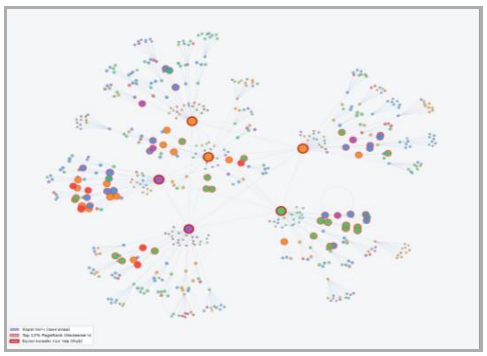
Санал болгох системийг ангиллын асуудал хэлбэрээр загварчилсан бөгөөд загварын гүйцэтгэлийг үнэлэхдээ AUC-ROC, нийт зөвийн үзүүлэлт, F1-оноо, нарийвчлал болон бүрхэц зэрэг үзүүлэлтүүдийг цогц байдлаар ашигласан. AUC-ROC нь загварын ялгах чадварыг нэг тоогоор илэрхийлэх бөгөөд ангиллын тэнцвэргүй байдлын үед найдвартай үнэлгээний үзүүлэлт болдог. Нарийвчлал нь тухайн ангилалд хамаарах гэж таамагласан утгуудын дундаас үнэхээр зөв байсан хувийг, бүрхэц нь бодит утгуудын дундаас зөв ангилсан хувийг тус тус илэрхийлнэ. F1-оноо нь эдгээр хоёр үзүүлэлтийн гармоник дундаж бөгөөд загварын гүйцэтгэлийг тэнцвэртэй үнэлэхэд ашиглагдана. Загваруудын гүйцэтгэлийн ялгааны статистикийн найдвартай байдлыг баталгаажуулахын тулд 95% итгэлцлийн интервалыг Bootstrap аргаар тооцоолсон бөгөөд AUC-ROC үзүүлэлтүүдийн харьцуулалтад DeLong тестийг ашигласан. Энэхүү судалгаанд олон ангиллын нөхцөлд загварын гүйцэтгэлийг илүү бодитой үнэлэх зорилгоор жинлэсэн F1-оноог голлон ашигласан.

C. Туршилтын үр дүн

Хэрэглэгч хоорондын дуудлагын харилцаанд тулгуурлан үүсгэсэн нийгмийн сүлжээний дэд графыг Python хэлний NetworkX санг ашиглан дүрсэлж, Fruchterman–Reingold байршуулалтын аргыг хэрэглэсэн бөгөөд энэ нь сүлжээн дэх бүлэглэлүүдийн бүтцийг тодорхой илэрхийлэх боломжийг олгосон. Бүлэглэлийг илрүүлэхдээ K-Clique (k=3) алгоритмыг ашигласан бөгөөд энэ нь гурав ба түүнээс дээш зангилааны бүрэн холбогдсон дэд графуудыг үндэслэн харилцааны нягт бүлгүүдийг тодорхойлох боломжийг олгодог. Уг аргаар нийт 67,918 харилцааны бүлгийг илрүүлсэн нь хэрэглэгчдийн нийгмийн орчны бүтцийг тодорхойлоход чухал үндэс болсон. Графын элементүүдийг зангилаа нь хэрэглэгчийг, холбоос нь хэрэглэгч хоорондын дуудлагын харилцааг, зангилааны өнгө нь хэрэглэгчийн ашиглаж буй бүтээгдэхүүний төрлийг, зангилааны хэмжээ нь сүлжээн дэх нөлөөллийн түвшинг тус тус илэрхийлнэ. Улаан хүрээтэй зангилаанууд нь өндөр нөлөө бүхий хэрэглэгчдийг (хуудасны зэрэглэлийн дагуу шилдэг 12%), төв зангилаа нь бүлэглэлийн гол хэрэглэгч буюу олон холболттой зангилааг илэрхийлнэ. Зураг 5-д дүрслэлийг харуулсан бөгөөд сүлжээ нь тодорхой бүлэглэлүүдэд хуваагдаж, өндөр нөлөө бүхий хэрэглэгчид сүлжээн дэх стратегийн байрлал эзэлж байгаа нь ажиглагдаж байна. Ийнхүү K-Clique алгоритмд суурилсан бүлэглэл илрүүлэлт болон графт суурилсан боловсруулалт нь хэрэглэгч хоорондын харилцааны бүтэц, нөлөөллийг илэрхийлэх боломжийг бүрдүүлж, санал болгох загварт ашиглах шинж чанаруудыг тодорхойлох үндэс болсон.

Графын элементүүдийг дараах байдлаар дүрсэлсэн:

- Зангилаа: Хэрэглэгч
- Холбоос: Хэрэглэгч хоорондын дуудлагын харилцаа
- Зангилааны өнгө: Хэрэглэгчийн ашиглаж буй бүтээгдэхүүний төрөл
- Зангилааны хэмжээ: Сүлжээн дэх нөлөөллийн түвшин (PageRank)
- Улаан хүрээтэй зангилаа: Өндөр нөлөө бүхий хэрэглэгчид (Top 12% PageRank)
- Төв зангилаа: Бүлэглэлийн гол хэрэглэгч буюу олон холболттой зангилаа



Зураг 7: Харилцаа холбооны өгөгдөлд суурилсан хэрэглэгчдийн нийгмийн сүлжээний дэд граф

Дээрх дүрслэд сүлжээ нь тодорхой бүлэглэлүүдэд хуваагдсан бөгөөд төв зангилаанууд нь тухайн бүлэглэл

дэх харилцааг төвлөрүүлж байна. Мөн өндөр нөлөө бүхий хэрэглэгчид нь сүлжээн дэх стратегийн байрлал эзэлж, нийгмийн нөлөөллийг тодорхойлох чухал үүрэгтэй байна. Иймхүү графт суурилсан боловсруулалт нь хэрэглэгч хоорондын харилцааны бүтэц, нөлөөллийг илэрхийлэх боломжийг бүрдүүлж, санал болгох загварт ашиглах шинж чанаруудыг тодорхойлох үндэс болсон.



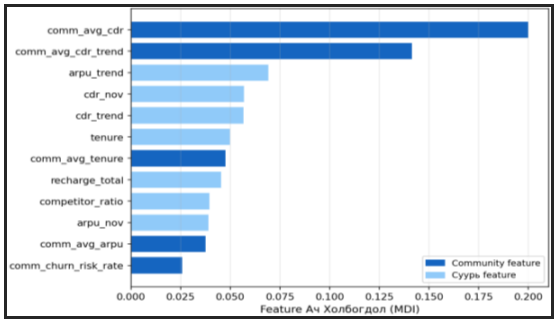
Зураг 8: Бүтээгдэхүүн шинэчлэлтийн загварын шинж чанарын ач холбогдлын үнэлгээ

Зураг 8-д бүтээгдэхүүн шинэчлэлтийн загварын шинж чанарын ач холбогдлын үнэлгээг харуулав. Үр дүнгээс харахад дахин цэнэглэлтийн тоо (MDI ≈ 0.28) болон нийт дахин цэнэглэлтийн дүн (MDI ≈ 0.22) хамгийн өндөр ач холбогдолтой шинж чанар болж гарсан нь хэрэглэгчий н санхүүгийн идэвх бүтээгдэхүүний сонголтод шийдвэрлэх нөлөө үзүүлж байгааг илтгэнэ. Үүний дараагаар ARPU болон ARPU-ийн чиг хандлага (MDI ≈ 0.08–0.10), өрсөлдөгчийн харьцаа болон хэрэглэгчийн хугацаа (MDI ≈ 0.05–0.07) харьцангуй өндөр ач холбогдолтой гарсан б а й н а . Улаанаар тэмдэглэсэн нийгмийн бүлгийн шинж чанарууд болох хөрш зангилааны дундаж зэрэглэл (MDI ≈ 0.04), хуудасны зэрэглэл, ойрын болон зуучлалын төвлөрөл зэрэг SNA хэмжүүрүүд дунд зэргийн ач холбогдолтой (MDI≈0.02–0.05) гарсан байна. Хүснэгт 2-ын үр д үн т э й харьцуулахад Community XGBoost загвар F1-macro үзүүлэлтээр (0.5870) суурь XGBoost-оос (0.5951) бага зэрэг доогуур гарсан нь бүтээгдэхүүн шинэчлэлтийн даалгаварт нийгмийн бүлгийн шинж чанарууд хязгаарлагдмал нэмэлт мэдээлэл өгч байгааг харуулж байна.

Загвар	Accuracy	F1-macro	95% CI	F1-weighted
Baseline LR	0.5736	0.4978	[0.48, 0.50]	0.5810
Baseline RF	0.7199	0.5805	[0.57, 0.58]	0.6844
Baseline XGBoost	0.7211	0.5951	[0.58, 0.60]	0.6944
Community RF	0.7166	0.5758	[0.56, 0.58]	0.6805
Community XGBoost	0.7152	0.5870	[0.57, 0.59]	0.6880
GraphSAGE	0.6859	0.5147	[0.50, 0.52]	0.6376

Хүснэгт 2: Бүтээгдэхүүн шинэчлэлтийн загваруудын гүйцэтгэлийн харьцуулалт

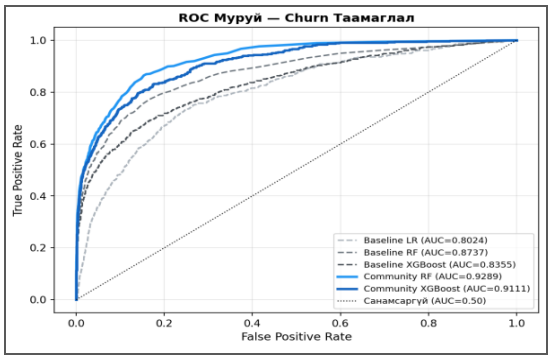
Хүснэгт 2-д бүтээгдэхүүн шинэчлэлтийн загваруудын нийт зөвийн үзүүлэлт, F1-macro, 95% итгэлцлийн интервал болон жинлэсэн F1-оноог харьцуулан үзүүлэв. Суурь загваруудаас логист регресс хамгийн бага гүйцэтгэлтэй (нийт зөв=0.5736, F1-macro=0.4978) байсан бол суурь XGBoost нийт зөвийн үзүүлэлтээр (0.7211) болон F1-macro-гаар (0.5951) суурь загваруудаас хамгийн өндөр гүйцэтгэл үзүүлсэн байна. Харилцааны бүлэгт суурилсан загваруудаас Community XGBoost нь F1-macro үзүүлэлтээр (0.5870) суурь ойн олон мод загвараас (0.5805) өндөр гарсан боловч суурь XGBoost-оос (0.5951) бага зэрэг доогуур гарсан байна. Community RF загвар нийт зөвийн үзүүлэлтээр (0.7166) суурь ойн олон мод загвартай (0.7199) ойролцоо гүйцэтгэл үзүүлсэн нь нийгмийн бүлгийн шинж чанарууд бүтээгдэхүүн сонгоход хязгаарлагдмал нэмэлт мэдээлэл өгч байгааг илтгэнэ. GraphSAGE загвар нийт зөвийн үзүүлэлтээр (0.6859) болон F1-macro-гаар (0.5147) бүх загваруудаас хамгийн бага гүйцэтгэлтэй гарсан бөгөөд 95% итгэлцлийн интервал [0.50, 0.52] нь уг үр дүн статистикийн хувьд тогтвортой болохыг баталж байна.



Зураг 9: Хэрэглэгч алдагдах таамаглалын загварын шинж чанарын ач холбогдлын үнэлгээ

Зураг 9-д хэрэглэгч алдагдах таамаглалын загварын шинж чанарын ач холбогдлын үнэлгээг харуулав. Үр дүнгээс харахад харилцааны бүлгийн дундаж CDR (MDI ≈0.200) болон CDR-ийн чиг хандлага (MDI ≈0.150) хамгийн өндөр ач холбогдолтой шинж чанар болж гарсан нь хэрэглэгчийн нийгмийн орчны харилцааны идэвх хэрэглэгч алдагдлыг таамаглахад голлох үүрэгтэй болохыг илтгэнэ. Энэ нь бүтээгдэхүүн шинэчлэлтийн загварын үр дүнтэй харьцуулахад нийгмийн бүлгийн шинж чанарууд хэрэглэгч алдагдлын таамаглалд хамаагүй өндөр нөлөөтэй болохыг тодорхой харуулж байна. Үүний дараагаар ARPU-ийн чиг хандлага (MDI≈ 0.080), 11-р сарын дуудлагын тоо (MDI≈0.060), дундаж хугацаа (MDI≈0.055) зэрэг шинж чанарууд өндөр ач холбогдолтой гарсан байна. Харилцааны бүлгийн дундаж хугацаа, нийт дахин цэнэглэлт, өрсөлдөгчийн харьцаа, дундаж ARPU зэрэг шинж чанарууд (MDI≈ 0.030–0.050) дунд зэргийн ач холбогдолтой гарсан бол хэрэглэгч алдагдлын хувь харьцангуй бага ач холбогдолтой (MDI≈0.020) гарсан байна.

Нийгмийн бүлгийн шинж чанарууд суурь шинж чанаруудтай харьцуулахад илүү өндөр нөлөөтэй гарсан нь нийгмийн орчны мэдээллийг загварт нэгтгэсний үр дүнд хэрэглэгч алдагдлын таамаглалын чанар мэдэгдэхүйц сайжирсан болохыг нотолж байна.



Зураг 10. Хэрэглэгч алдагдах таамаглалын загваруудын ROC муруйн харьцуулалт

Зураг 10-д хэрэглэгч алдагдах таамаглалын загваруудын ROC муруйн харьцуулалтыг үзүүлэв. Харилцааны бүлэгт суурилсан ойн олон мод загвар хамгийн өндөр гүйцэтгэлтэй бөгөөд суурь загваруудтай харьцуулахад мэдэгдэхүйц сайжирсан байна. Community XGBoost болон граф нейрон сүлжээний загвар ч өндөр гүйцэтгэл үзүүлсэн бол логист регресс суурь загвар хамгийн бага гүйцэтгэлтэй гарсан нь шугаман загвар хэрэглэгч алдагдлын нарийн төвөгтэй хамаарлыг бүрэн илрүүлж чадахгүй байгааг харуулж байна. ROC муруйн дүрслэлээс харахад нийгмийн бүлгийн шинж чанаруудыг нэмсэн загваруудын муруй суурь загваруудынхаас тодорхой дээш байрлаж байгаа нь нийгмийн мэдээллийн нөлөөллийг тодорхой харуулж байна.

Загвар	AUC-ROC	95% CI	PR-AUC	Precision	Recall
Baseline LR	0.8024	[0.78 0.8202]	0.2290	0.1692	0.6530
Baseline RF	0.8737	[0.85,0.8901]	0.5206	0.8053	0.2871
Baseline XGBoost	0.8355	[0.81, 0.8533]	0.4713	0.4702	0.4606
Community RF	0.9289	[0.91, 0.9382]	0.6124	0.8488	0.3454
Community XGBoost	0.9111	[0.89, 0.9227]	0.5819	0.6034	0.5110
GraphSAGE	0.9012	[0.88, 0.9143]	0.5492	0.3059	0.7319

Хүснэгт3: Хэрэглэгч алдагдах таамаглалын загваруудын гүйцэтгэлийн харьцуулалт

Хүснэгт 3-д хэрэглэгч алдагдах таамаглалын загваруудын AUC-ROC, 95% итгэлцлийн интервал, нарийвчлал болон бүрхэцийн харьцуулалтыг үзүүлэв. Community RF загвар AUC-ROC үзүүлэлтээр хамгийн өндөр гүйцэтгэлтэй бөгөөд нарийвчлал өндөр боловч бүрхэц харьцангуй бага байгаа нь хэрэглэгч алдагдлын тэнцвэргүй ангиллын нөлөөллийг илтгэнэ. Граф нейрон сүлжээний загвар AUC-ROC үзүүлэлтээр өндөр гарсан боловч нарийвчлал харьцангуй бага, бүрхэц өндөр байгаа нь загвар илүү олон хэрэглэгчийг алдагдах эрсдэлтэй гэж таамаглаж байгааг харуулж байна. 95% итгэлцлийн интервалуудаас харахад бүх загваруудын үр дүн статистикийн хувьд тогтвортой болохыг баталж байна.

D. Туршилтын дүгнэлт:

Нэгдүгээрт, нийгмийн сүлжээний шинжилгээний шинж чанарууд хэрэглэгч алдагдлын таамаглалын гүйцэтгэлийг мэдэгдэхүйц сайжруулсан бөгөөд харилцааны бүлэгт суурилсан ойн олон мод загвар суурь загвартай харьцуулахад AUC-ROC үзүүлэлтийг статистикийн хувьд ач холбогдолтойгоор нэмэгдүүлсэн. Харилцааны бүлгийн дундаж дуудлагын тоо болон түүний чиг хандлага хамгийн өндөр ач холбогдолтой шинж чанар болж гарсан нь нэг бүлэгт харьяалагдах хэрэглэгчдийн алдагдах зан төлөв харилцан нөлөөлдөг болохыг бодит өгөгдлөөр нотолж байна. Энэ нь нийгмийн сүлжээний бүтцийн мэдээллийг загварт нэгтгэх нь хэрэглэгч алдагдлын таамаглалын чанарыг сайжруулахад үр дүнтэй болохыг баталж байна.

Хоёрдугаарт, бүтээгдэхүүн шинэчлэлтийн даалгаварт нийгмийн сүлжээний шинж чанарууд таамаглалын гүйцэтгэлийг мэдэгдэхүйц сайжруулж чадаагүй бөгөөд харилцааны бүлэгт суурилсан загварууд суурь загваруудтай харьцуулахад F1-macro үзүүлэлтээр ижил түвшинд буюу бага зэрэг доогуур гарсан байна. Энэ нь бүтээгдэхүүн сонголт нийгмийн орчны нөлөөллөөс илүүтэй хувь хүний санхүүгийн идэвх болон хэрэглээний хэв маягтай хамааралтай болохыг харуулж байна. Иймд нийгмийн сүлжээний мэдээллийн нөлөөлөл нь даалгаврын төрлөөс хамааран ялгаатай байдаг болохыг энэхүү үр дүн тодорхой илтгэж байна.

Гуравдугаарт, граф нейрон сүлжээний загвар суурь загваруудыг AUC-ROC үзүүлэлтээр давж чадсан боловч харилцааны бүлэгт суурилсан ойн олон мод загварыг давж чадаагүй байна. Графын сийрэг бүтэц нь граф нейрон сүлжээний загварын зангилааны орчны мэдээллийг нэгтгэх чадварыг хязгаарлаж байгаа нь цаашдын судалгаанд анхаарах чухал дүгнэлт болж байна. Ийнхүүгийн сийрэг графын нөхцөлд уламжлалт машин сургалтын аргууд граф нейрон сүлжээний загвараас илүү тогтвортой гүйцэтгэл үзүүлж байгааг энэхүү туршилт харуулж байна.

Дүгнэлт

Энэхүү судалгаагаар харилцаа холбооны салбарт уламжлалт санал болгох аргуудын хязгаарлалтыг шийдвэрлэх зорилгоор CDR өгөгдөлд тулгуурлан хэрэглэгч хоорондын харилцааг жигнэсэн чиглэлтэй граф G=(V,E) хэлбэрт загварчилж, K-Clique (k=3) аргаар харилцааны бүлгүүдийг илрүүлж, SNA хэмжүүрүүдийг машин сургалтын загварт нэгтгэсэн цогц framework-ийг боловсруулж туршсан. Хэрэглэгч алдагдах таамаглал болон бүтээгдэхүүн санал болгох хоёр даалгаврыг нэгдсэн хамгийн тохиромжтой дараагийн үйлдлийн шийдвэрийн системд нэгтгэсэн нь практик хэрэглээтэй нэгдсэн framework болж байна.Судалгаанд нийт 75,701 хэрэглэгчийн 4,914,906 дуудлагын харилцаанд суурилсан 34 хувьсагч бүхий нэгдсэн өгөгдлийн санг ашигласан бөгөөд K-Clique (k=3) аргаар 67,918 харилцааны бүлгийг илрүүлсэн. Хэрэглэгч алдагдлын хувь 5.8% байгаа нь ангиллын тэнцвэргүй байдал үүссэнийг харуулж байгаа тул загвар сургалтын үед

жинлэлтийн тохиргоог ашиглан энэхүү нөлөөллийг бууруулсан. Өгөгдлийг сургалт (70%), баталгаажуулалт (15%), туршилт (15%) гэсэн харьцаагаар хуваасан бөгөөд ажиглалтын хугацаа нь 2025 оны 10–11-р сар, таамаглалын цонх нь 2025 оны 12-р сар байв.Туршилтын үр дүнгээс харахад харилцааны бүлэгт суурилсан ойн олон мод загвар хэрэглэгч алдагдлын таамаглалд AUC-ROC үзүүлэлтийг суурь загвартай харьцуулахад статистикийн хувьд ач холбогдолтойгоор сайжруулсан нь DeLong тестээр баталгаажсан бөгөөд харилцааны бүлгийн дундаж дуудлагын тоо хамгийн нөлөөтэй шинж чанар болж гарсан нь нийгмийн нэгэн төрөл байдлын үзэгдлийг бодит өгөгдлөөр нотолж байна. Харин бүтээгдэхүүн шинэчлэлтийн даалгаварт нийгмийн сүлжээний шинж чанарууд таамаглалын гүйцэтгэлийг мэдэгдэхүйц сайжруулж чадаагүй нь нийгмийн мэдээллийн нөлөөлөл даалгаврын төрлөөс хамааралтай болохыг илтгэж байна. Граф нейрон сүлжээний загвар суурь загваруудыг давсан боловч харилцааны бүлэгт суурилсан загварыг давж чадаагүй нь графын сийрэг бүтцийн нөхцөлд уламжлалт машин сургалтын аргууд илүү тогтвортой гүйцэтгэл үзүүлдгийг харуулж байна. Цаашид бодит цагийн өгөгдөл дээр туршилт хийх, граф нейрон сүлжээний загварт зангилааны шинж чанарыг баяжуулах болон хэрэглэгчийн бусад шинж чанаруудыг хамруулах замаар загварын гүйцэтгэлийг улам сайжруулах боломжтой гэж үзэж байна.

АШИГЛАСАН НОМ ЗҮЙ

[1] S. P. Borgatti, M. G. Everett, and J. C. Johnson, Analyzing Social Networks. Thousand Oaks, CA, USA: Sage Publications, 2018, ch. 1–3.

[2] S. Gurukar and B. Ravindran, “Temporal analysis of telecom call graphs,” in Proc. 6th Int. Conf. Communication Systems and Networks (COMSNETS), Bangalore, India, Jan. 2014, pp. 1–8.

[3] H. Ma, I. King, and M. R. Lyu, “Learning to recommend with social trust ensemble,” in Proc. 32nd Int. ACM SIGIR Conf., Boston, MA, USA, Jul. 2009, pp. 203–210.

[4] X. Su and T. M. Khoshgoftaar, “A survey of collaborative filtering-based recommender systems,” Artificial Intelligence Review, vol. 47, no. 4, pp. 427–451, Apr. 2017.

[5] F. Ricci, L. Rokach, and B. Shapira, Recommender Systems Handbook, 2nd ed. New York, NY, USA: Springer, 2015, ch. 1, pp. 1–35.

[6] X. Wang, X. He, M. Wang, F. Feng, and T.-S. Chua, “Neural graph collaborative filtering,” in Proc. 42nd Int. ACM SIGIR Conf., Paris, France, Jul. 2019, pp. 165–174

[7] X. Qian, H. Feng, G. Zhao, and T. Mei, “Personalized recommendation combining user interest and social circle,” IEEE Trans. Knowledge and Data Engineering, vol. 26, no. 7, pp. 1763–1777, Jul. 2014.

[8] A. Géron, Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow, 2nd ed. Sebastopol, CA, USA: O’Reilly, 2019, ch. 6–7.

[9] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in Proc. 22nd ACM SIGKDD, San Francisco, CA, USA, Aug. 2016, pp. 785–794

[10] G. Ke et al., “LightGBM: A highly efficient gradient boosting decision tree,” in Advances in Neural Information Processing Systems (NeurIPS), Long Beach, CA, USA, 2017, pp. 3146–3154.

[11] F. Ricci, L. Rokach, and B. Shapira, Recommender Systems Handbook. New York, NY, USA: Springer, 2011, ch. 1–2.

[12] T. Mahmood and F. Ricci, “Improving recommender systems with adaptive conversational strategies,” in Proc. ACM Conf. Hypertext and Hypermedia, Turin, Italy, 2009, pp. 73–82.

[13] F. Ortega, A. Hernando, J. Bobadilla, and J. H. Kang, “Recommending items to groups of users using matrix factorization-based collaborative filtering,” Information Sciences, vol. 345, pp. 313–324, May 2016.

[14] X. Su and T. M. Khoshgoftaar, “A survey of collaborative filtering-based recommender systems: From traditional methods to hybrid methods based on social networks,” Artificial Intelligence Review, vol. 47, no. 4, pp. 427–451, Apr. 2017.

[15] H. Ma, I. King, and M. R. Lyu, “Learning to recommend with social trust ensemble,” in Proc. 32nd Int. ACM SIGIR Conf. Research and Development in Information Retrieval, Boston, MA, USA, Jul. 2009, pp. 203–210.

[16] X. Qian, H. Feng, G. Zhao, and T. Mei, “Personalized recommendation combining user interest and social circle,” IEEE Transactions on Knowledge and Data Engineering, vol. 26, no. 7, pp. 1763–1777, Jul. 2014.

[17] J. Zhang, B. Chen, Y. Xie, H. Cui, and B. He, “A survey on graph neural networks for recommender systems,” ACM Computing Surveys, vol. 54, no. 5, pp. 1–37, 2021.

[18] X. Wang, X. He, M. Wang, F. Feng, and T.-S. Chua, “Neural graph collaborative filtering,” in Proc. 42nd Int. ACM SIGIR Conf., Paris, France, Jul. 2019, pp. 165–174.

[19] M. Al-Molhem et al., “Social network analysis in telecom data: A review,” Journal of Big Data, vol. 6, no. 1, pp. 1–27, 2019.

[20] S. Bae and S. Kim, “Identifying influential spreaders using k-shell decomposition in complex networks,” Physica A: Statistical Mechanics and its Applications, vol. 395, pp. 549–559, 2014.

5G, WiFi6 сүлжээнд зориулсан давтамжийн дахин тохируулгатай антенны загварчлал

Баянбат Урангуа
Холбооны салбар
МХТС, ШУТИС
Улаанбаатар, Монгол улс
B232300001@must.edu.mn

Энхболд Батжаргал
Холбооны салбар
МХТС, ШУТИС
Улаанбаатар, Монгол улс
B232300008@must.edu.mn

Азхүү Саруулбуян
Холбооны салбар
МХТС, ШУТИС
Улаанбаатар, Монгол улс
B232305005@must.edu.mn

Баярсайхан Пүрэвцэрэн
Холбооны салбар
МХТС, ШУТИС
Улаанбаатар, Монгол улс
purevtseren_b@must.edu.mn

Пүрэвдорж Ууганбаяр
Холбооны салбар
МХТС, ШУТИС
Улаанбаатар, Монгол улс
p.uuganbayar@must.edu.mn

***Хураангуй*—Өнөө үед 5G болон Wi-Fi6 зэрэг утасгүй холбооны өндөр хурдны системүүдэд олон давтамжийн зурваст ажиллах чадвартай, авсаархан антены шийдэл нэн шаардлагатай байна. Энэхүү өгүүлэлд варактор диодыг идэвхтэй элемент болгон ашиглан ажлын давтамжийг, оролтын хүчдэлийг ихэсгэж багасгах аргаар дахин тохируулдаг бичил туузан антены загварыг санал болгож байна. Судалгааны зорилго нь варактор диодын багтаамжийг өөрчлөх замаар антены цахилгаан уртыг зохицуулж, нэг антеныг олон өөр давтамжийн мужид үр ашигтай ажиллуулахад оршино. Санал болгож буй антен нь FR-4 цахилгаан тусгаарлагч материал дээр суурилсан,** 30×26 **мм² хэмжээтэй бөгөөд 50 Ω-ийн бичил туузан шугамаар тэжээгдэнэ. Антены бүтцэд нэгтгэсэн диодын тохируулгын горимын тусламжтайгаар 5 ГГц (5G NR n79), 6 ГГц (Wi-Fi6), болон 3.6 ГГц (5G NR n77) давтамжийн мужуудад шилжин ажиллах боломжийг бүрдүүлсэн. Загварчлалыг CST Studio Suite программ дээр гүйцэтгэж, ойлтын коэффициент (**S_{11}**), цацаргалтын диаграмм болон өсгөлтийн үзүүлэлтүүдийг тодорхойлов. Симуляцийн үр дүнгээс харахад бүх ажлын горимд** S_{11}**−10-дБ-ээс бага буюу зохицол сайн байгаа бөгөөд хамгийн их өсгөлт 5–6 дБ байна. Цаашид бодит загвартай программын тооцооллыг харьцуулан шинжилснээр загварчлалын найдвартай байдлыг нотлох бөгөөд энэхүү антен нь орчин үеийн утасгүй холбооны болон IoT төхөөрөмжүүдэд ашиглагдах бүрэн боломжтой юм.**

***Түлхүүр үг*—Горим, багтаамж, давтамжийн дахин тохируулга, 5G NR, WiFi6**

I. ОРШИЛ

Сүүлийн жилүүдэд утасгүй холбооны технологи хурдацтай хөгжиж, өгөгдөл дамжуулах хурд, сүлжээний багтаамж болон олон төрлийн үйлчилгээний хэрэгцээ тасралтгүй нэмэгдэж байна. Ялангуяа 5G болон WiFi6 зэрэг шинэ үеийн утасгүй холбооны системүүд нь өндөр хурд, өргөн зурвас болон олон хэрэглэгчийг нэгэн зэрэг дэмжих шаардлагатай болж байна [6]–[8]. Ийм нөхцөлд холбооны төхөөрөмжүүдэд ашиглагдах антен нь олон давтамжийн мужид ажиллах, уян хатан тохируулгатай байх шаардлага улам нэмэгдэж байна [1], [2]. Давтамжийн дахин тохируулгыг хэрэгжүүлэхэд PIN диод, варактор диод болон бичил цахилгаан механик элементүүдийг өргөн ашигладаг [9]. Манай судалгаатай ижил боловч PIN диод ашигласан ”Design of Frequency Reconfigurable Patch Antenna Using PIN Diodes at S-Band” [17] гэдэг судалгаа байдаг. PIN диодын дутагдалтай тал нь олон давтамж шилжих тусам олон диод хэрэг болно. Диод болгоныг свитчээр удирдах хэрэгтэй болдог. Иймээс учраас олон диод ашиглах тусам олон свитч хэрэг болно.Тиймээс нэг төрлийн варактор диодоор хүчдэлийг өөрчлөх зарчмаар олон давтамжийн мужид ажиллах боломжийг олгоно. [3], [5].Иймд варактор диод ашигласан бичил туузан антен нь давтамжийн тохируулга хийхэд тохиромжтой шийдэл болдог.

II. СУДАЛГААНЫ АРГА ЗҮЙ

Антены цахилгаан уртыг удирдах зарчим микрострип антены ажлын давтамж нь цацаргагч хавтгайн цахилгаан урт (electrical length- L_{eff})-тай шууд хамааралтай. Цахилгаан урт гэдэг нь физик уртаас ялгаатай бөгөөд цахилгаан орны тархалтын бодит замыг илэрхийлнэ. Варактор диодыг нэмэлт идэвхгүй цацаргагч элементтэй цуваа холбосноор хагас дамжуулагч элментийн туламжтайгаар цацаргагч хавтгайн цахилгаан уртыг динамикаар өөрчлөх боломж бүрддэг. Хүчдэлийн өөрчлөлтөөс үүдэлтэй физик дараалал: (1) Энд нь диодын реактив эсэргүүцэл бөгөөд багтаамж буурахад нэмэгддэг. Ийнхүү диод нь нэмэлт элементийг антенны үндсэн цацаргагчтай холбох буюу тусгаарлах үүрэг гүйцэтгэж, антены цахилгаан уртыг "агшааж" эсвэл "томруулж" давтамжийн шилжилтийг бий болгодог [1], [11]. Хүчдэл (C_v) нэмэгдэх тусам давтамж өндөрсөх дараалал:

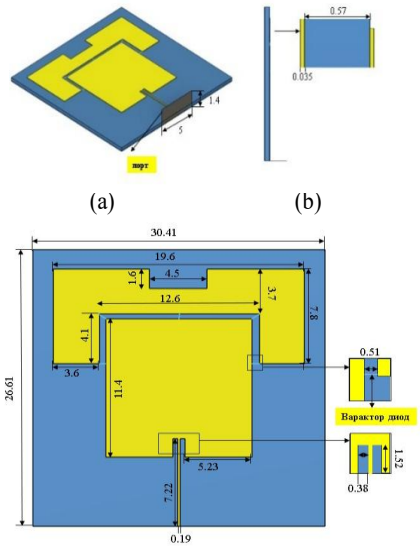
$$V_{\text{bias}} \uparrow \Rightarrow C_v \downarrow \Rightarrow L_{\text{eff}} \downarrow \Rightarrow f \uparrow \qquad (1)$$

Энд:
C_v -Варактор диодын багтаамж.
L_{eff}- Антенны эффектив цахилгаан урт.
Z_0- 50 Om тэжээлийн шугмийн эсэргүүцэл.
V_{bias}- Диодод өгөх удирдлагын хүчдэл.

III. АНТЕНЫ ЗАГВАРЧЛАЛ

A. Цахилгаан тусгаарлагч материал болон геометрийн бүтэц

Санал болгож буй антен нь FR-4 цахилгаан тусгаарлагч материал дээр хийгдсэн бичил туузан цацруулагч хавтгай бөгөөд харьцангуй диэлектрик тогтмол ε_r = 4.4, цахилгаан тусгаарлагч материал-ийн зузаан h = 0.57 мм байна. Антены нийт хэмжээ 30мм × 26мм бөгөөд 50Ω бичил туузан шугамаар тэжээгдэнэ [11].



(a) (b)

(c)

Зураг 1: Бичил туузан антены загвар — a. 3D загвар , b хажуугаас харсан загвар, c. Эгц урдаас харсан загвар

B. Патч антенны хэмжээний тооцоолол

Патч антенны өргөн W болон урт L-ийг дараах стандарт томьёогоор тооцоолно [1], [11]:

Өргөн (Width):

$$W = \frac{c}{2f_r}\sqrt{\frac{2}{\varepsilon_r + 1}} \quad (2)$$

Үүнд:

- c – Гэрлийн хурд (3×10^8 м/с);
- f –давтамж;
- ε_r – Тусгаарлагч суурь материалын диэлектрик нэвтрүүлэлт.

Үр дүнтэй диэлектрик тогтмол:

$$\varepsilon_{eff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left(1 + \frac{12h}{W}\right)^{-1/2} \quad (3)$$

Үүнд:

- h– Суурь материалын зузаан;
- W – Цацруулагчийн өргөн.

Урт (Length):

$$L = \frac{c}{2f_r\sqrt{\varepsilon_{eff}}} - 2\Delta L \quad (4)$$

B. Нэмэлт цацаргагчийн тооцоолол

Варактор диодоор холбон байрлуулах нэмэлт цацаргагчийн хэмжээний тооцоолол [4]:

$$L_s = \frac{\lambda_g}{2} \quad W_s = 0.05\lambda_g \text{ мм} \quad (5)$$

Тооцооллын үр дүнг Хүснэгт I-д харуулав.

Хүснэгт I: Патч Антенны Загварын Хэмжээний Үр Дүн

Параметр	Тэмдэглэгээ	Утга (мм)
Патч өргөн	W	10.7
Патч урт	L	11.4
Тэжээлийн өргөн	W_f	0.19
Нийт хавтгайн урт	L_{total}	30.41
Нийт хавтгайн өргөн	W_{total}	26.61

D. Варактор диодоор холболт нэмэх

Варактор диодыг ховилд хэрхэн холбох нь антенны үр ашигт нөлөөлдөг. Диод холбосон үед нэмэгдэх нийт цахилгаан урт [5]:

$$L_{eff} = L + \frac{c}{2\pi f_r}\arctan\frac{1}{2\pi f_r C_v Z_0} \quad (6)$$

энд $Z_0 = 50\Omega$ холбооны эсэргүүцэл, C_v нь варактор диодын багтаамж юм.

E. Давтамжийн тохируулгын горимууд ба варактор диодын багтаамж

Варактор диодын багтаамж C_v нь хүчдэлийн өөрчлөлтөөр тодорхойлогдоно [3]:

$$C_v(V) = \frac{C_{j0}}{\left(1 + \frac{V}{V_j}\right)^m} \quad (7)$$

Оролтын хүчдэл болон багтаамжийн утгаас хамааран антенн өөр өөр горимд ажиллана [6], [7]. Хүснэгт IIд хүчдэл–багтаамжийн харьцааг, Хүснэгт II-д горим тус бүрийн тохиргоог харуулав.

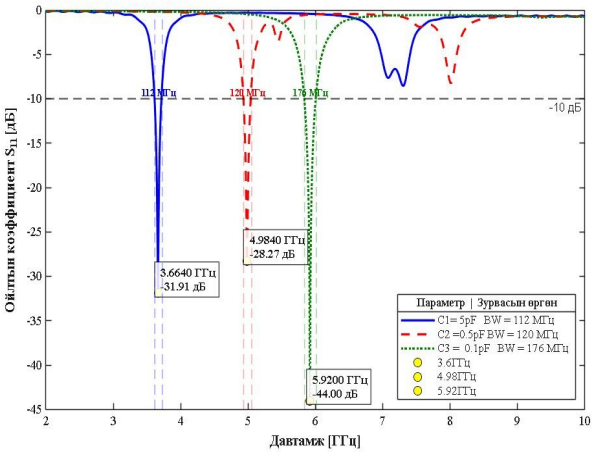
Хүснэгт II: Антенны Ажлын Горимуудын Тохиргоо

Горим	Багтаамж (пФ)	Давтамж (ГГц)	Систем
Горим1	0.1	6.0	WiFi6
Горим2	0.5	5.0	5G NR n79
Горим3	5	3.6	5G NR n78

IV. ЗАГВАРЧЛАЛЫН ҮР ДҮН

A. Ойлтын коэффициент

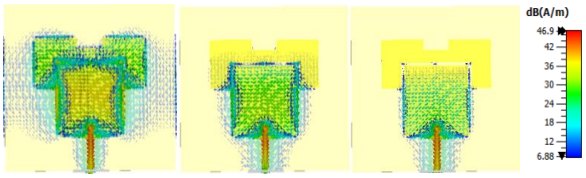
CST программаар хийгдсэн загварчлалын S_{11}ойлтын коэффициентийн үр дүнг Fig. 2-д харуулав. Гурван горимд ч $S_{11} < -10$-дБ байгаа нь антенн тухайн давтамжид сайн зохицсныг харуулж байна.



Зураг 2: 3 хос горимын S_{11} ойлтын коэффициентийн хамаарал.

B. Гадаргуугийн гүйдлийн тархалтын шинжилгээ

Антенны цацаргалтын механизмыг тодорхойлоход гадаргуугийн гүйдлийн (surface current) тархалтын шинжилгээ чухал үүрэгтэй. CST загварчлалаар 5.92 ГГц, 5 ГГц болон 3.6 ГГц давтамжуудад гадаргуугийн гүйдлийн тархалтыг үзүүлэв.



(a) (b) (c)

Зураг 3: a. 3.6 ГГц b. 5 ГГц болон c. 5.92 ГГц давтамжууд дахь гадаргуугийн гүйдлийн тархалт.

1) 5.92 ГГц дээрх гадаргуугийн гүйдэл (Mode1 — WiFi6): Зураг a-аас харахад 5.92 ГГц давтамжид гүйдлийн нягтрал Патчын дунд хэсэгт болон диодоор холбогдсон нэмэлт хэсгийн эргэн тойронд хамгийн өндөр байна. Варактор диодын багтаамж C_v = 0.1пФ үед цахилгаан орны дамжуулал нэмэлт хэсэгрүү хаагдаж, цацруулагч хавтгайн талбай хязгарлагдан, ажлын давтамж 5.92ГГц-д тогтдог. Гүйдлийн тархалт нэг чиглэлтэй бөгөөд тэгш хэмтэй хэлбэртэй байгаа нь антены цацаргалтын хэлбэрийн хэвийн шинжийг үзүүлэх юм.

2) 5 ГГц дээрх гадаргуугийн гүйдэл (Mode2 — 5G NR n79): Зураг b-ээс харахад 5 ГГц давтамжид (C_v = 0.5 пФ) гүйдлийн нягтрал цацруулагч хавтгайн талбай ирмэг болон тэжээлийн шугамтай холбогдох хэсэгт тархаж байна. Багтаамж нэмэгдэхийн хэрээр антены цацаргагчын талбай томрон, 5 ГГц-д шилжсэнийг гүйдлийн тархалтын хэлбэр тодорхой харуулж байна.

3) 3.6 ГГц дээрх гадаргуугийн гүйдэл (Mode3 — 5G NR n77): Зураг c-ээс харахад 3.66 ГГц давтамжид (C_v = 5.0 пФ) гүйдлийн нягтрал цацаргагчийн ирмэг болон тэжээлийн шугамтай холбогдох хэсэгт тархаж байна.

Гурван горимын гадаргуугийн гүйдлийн харьцуулалтаас дараах дүгнэлт гарч байна: C_vнэмэгдэх тусам гүйдлийн нягтрал Патчын захруу шилжиж, антены цацаргагчын талбай томрон давтамж буурч байна (Зураг 4.). Энэ нь (1) томьёонд тодорхойлсон тохируулгын зарчимтай бүрэн нийцэж байна.

C. Цацралтын хэлбэр ба өсгөлтийн шинжилгээ



(a) (b) (c)

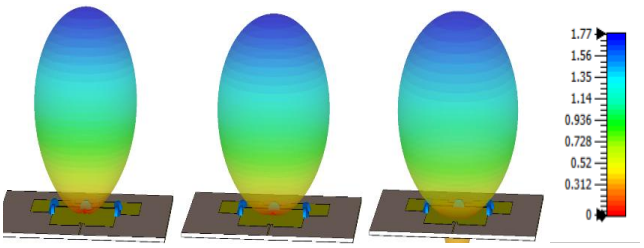
Зураг 4: a. 5.92 ГГц b. 5 ГГц болон c. 3.6 ГГц давтамж дахь цацаргалтын хэлбэр ба өсгөлт.

Антенны ажиллагааны үндсэн давтамжууд дээрх цацрагийн шинж чанарыг E-хавтгай (Phi=0°) болон H-хавтгай (Phi=90°)-н зүсэлтээр Зураг 4-д харуулав. Диаграммд улаан үргэлжилсэн шугамаар цахилгаан орны хүчлэг (E-plane), хөх тасархай шугамаар соронзон орны хүчлэгийг (H-plane) тус тус тэмдэглэсэн болно.

1) 5.92 ГГц (Mode1 — WiFi6): Зураг a-аас харахад 5.92ГГц давтамжид E-plane (Phi=0°): Цахилгаан орны хүчлэгийн үндсэн дэлбээ нь 0° градуст хамгийн их утгатай буюу 7.09 dBi байна. Харин -150° градуст хүчлэг хамгийн бага буюу -22.41 dBi хүртэл унаж байна. H-plane (Phi=90°): Соронзон орны хүчлэг нь мөн 0° градусын чиглэлд хамгийн их буюу 5.17 dBi утгатай байна. Энэ давтамжид антенн нь илүү нарийн чиглэлтэй цацрагийг үүсгэж байна.

2) 5 ГГц (Mode2 — 5G NR n79): Зураг b-ээс харахад 5ГГц давтамжид 1E-plane (Phi=0°): Цахилгаан орны хүчлэг 0° градусын чиглэлд 6.59 dBi хүртэл өсөж хамгийн их утгаа авч байна. Тэгш хэмт байдал алдагдахгүйгээр -130° градус орчимд хүчлэг хамгийн бага буюу -12.92 dBi байна. H-plane (Phi=90°): Соронзон орны хүчлэгийн хамгийн их утга 0° градуст 5.17 dBi байгаа бөгөөд -110° градусын орчимд хүчлэг хамгийн бага буюу -5.03 dBi болж байна.

3) 3.6 ГГц дээрх гадаргуугийн гүйдэл (Mode3 — 5G NR n77): Зураг c-ээс харахад 3.66 ГГц давтамжид E-plane (Phi=0°): Цахилгаан орны хүчлэгийн хамгийн их утга 0° градуст 5.44 dBi байна. Хамгийн бага хүчлэг нь -125° градуст -6.48 dBi орчимд ажиглагдаж байна.H-plane (Phi=90°): Соронзон орны хүчлэг 0° градусын чиглэлд 5.44 dBi утгатай байна. Энэ зурваст E болон H хавтгайн диаграммууд илүү өргөн (omnidirectional тал руугаа) хэлбэртэй байна.



(a) (b) (c)

Зураг 5: a. 5.92 ГГц b. 5 ГГц болон c. 3.6 ГГц давтамж дахь 3D цацаргалтын хэлбэр.

Цацралтын хэлбэрийн загварчлалаас харахад антенн бүх горимд нэг чиглэлтэй цацралтын хэв маягтай бөгөөд хамгийн их өсгөлт нь 5–7 дБ орчим байна [15]. Зураг 4-өөс Mode1 (WiFi6) буюу 6 ГГц дээрх, Mode2 (5G NR n79) буюу 5 ГГц дээрх Mode3 (5G NR n77) буюу 3.6 ГГц дээрх 2D болон 3D диаграммыг харуулж байна.Варактор диодын багтаамжийг өөрчлөх замаар ажлын давтамжийг шилжүүлэхэд антенны үндсэн цацаргалтын хэлбэр өөрчлөгдөлгүй, бүх горимд нэг чиглэлт (unidirectional) шинж чанараа хадгалж байна. Энэ нь антенны бүтэц давтамжийн өргөн мужид тогтвортой ажиллах чадвартайг нотолж байна.3D диаграммаас харахад цахилгаан соронзон энерги нь цацаргагчийн эгц дээрх чиглэлд (broadside radiation) үр ашигтай төвлөрч байгаа бөгөөд хажуугийн болон ар талын дэлбэ (side and back lobes) бага байгаа нь антены

цацаргалтын үр ашиг сайн байгааг илтгэнэ.Эдгээр үр дүн нь санал болгож буй антены загвар нь 5G болон WiFi6-ийн өндөр хурдны өгөгдөл дамжуулах шаардлагыг хангахүйц өндөр өсгөлттэй, чиглэлт шинж чанар сайтай болохыг техникийн талаас баталж байна.

D. CST загварчлалын үндсэн үзүүлэлтүүд

Хүснэгт III: CST Загварчилсан Антенны Үндсэн Үзүүлэлтүүд

Горим	Давтамж (ГГц)	S_{11} (dB)	Өсгөлт (dBi)	Зурвасын Өргөн (МГц)
Горим1	5.92	-44	6.95	176
Горим2	4.98	-28.27	6.47	120
Горим3	3.66	-31.91	5.37	112

V. ДҮГНЭЛТ

Энэхүү өгүүлэлд варактор диодыг идэвхтэй тохируулгын элемент болгон ашиглан 5G NR болон WiFi6 системүүдэд зориулсан давтамжийн дахин тохируулгатай бичил туузан антены загварыг боловсруулж, онолын үндэслэл болон CST Studio Suite программ дахь нарийвчилсан электромагнит загварчлалаар үр ашгийг нь баталгаажуулав.

Онолын үндэслэл. Варактор диодын тохируулгын зарчим нь (1) томьёогоор тодорхойлогддог: оролтын хүчдэл V_{bias} нэмэгдэх тусам диодын багтаамж C_v буурч, антены эффектив цахилгаан урт L_{eff} (6) томьёоны дагуу богиносч, улмаар ажлын давтамж f өндөрсдөг. Харин V_{bias} буурахад эсрэг үйл явц өрнөж давтамж буурна. Бичил туузан антены өргөн W (2), үр дүнтэй диэлектрик тогтмол ε_{eff} (3) болон ажлын урт L (4) зэрэг стандарт параметрүүдийг Balanis [1] болон Pozar [11]-ийн онолын аргачлалаар тооцоолж, Хүснэгт I-д үзүүлсэн загварын хэмжээсийг тодорхойлов. Варактор диодын багтаамжийн давтамжаас хамаарах шинж чанарыг (7) томьёогоор дүрслэн, C_v = 0.1 пФ, 0.5 пФ, 5.0 пФ гэсэн гурван тохируулгын цэгийг онолын тооцооллоор урьдчилан тодорхойлсон [3], [5].

Онолын тооцооллоор урьдчилан таамагласан гурван давтамжийн горимыг CST загварчлалын тоон үр дүн дараах байдлаар бүрэн баталгаажууллаа:

- Горим1 (WiFi6, 5.92 ГГц): Онолоор 6 ГГц давтамжид ажиллана гэж тооцоолсон бөгөөд CST симуляциар S_{11} = −44 дБ — стандарт шаардлага болох −10 дБ-ийн босгоос 34 дБ-ээр доогуур байгаа нь онолын таамаглалыг тоон утгаар нотолж байна. Өсгөлт 6.95 дБи, зурвасын өргөн 176 МГц байв .
- Горим2 (5G NR n79, 4.98 ГГц): Онолоор 5 ГГц давтамжид шилжинэ гэж тооцоолсон бөгөөд CST-ийн S_{11} = −28.27 дБ утга нь шаардлагат босгоос 18 дБээр доогуур байж, 3GPP TS38.104 стандартын n79 зурвасын шаардлагыг хангасан болохыг баталж байна [6]. Өсгөлт 6.47 дБи, зурвасын өргөн 120 МГц байв.
- Горим3 (5G NR n77, 3.66 ГГц): Онолоор 3.6 ГГц давтамжид шилжинэ гэж тооцоолсон бөгөөд CST-ийн S_{11} = −31.91 дБ утга нь шаардлагат босгоос 21 дБ-ээр бага байж онолын таамаглалыг баталгаажуулав [6]. Өсгөлт 5.37 дБи, зурвасын өргөн 112 МГц байв.

Онолоор (1) томьёонд тодорхойлсон $C_v \uparrow \Rightarrow L_{\text{eff}} \uparrow \Rightarrow f_r \downarrow$ тохируулгын механизмыг гадаргуугийн гүйдлийн шинжилгээ дүрслэлийн хэлбэрээр тодорхой нотолсон: C_v буюу оролтонд өгч буй хүчдэл нэмэгдэх тусам гүйдлийн нягтрал цацаргагчын захруу шилжиж, цацаргагч нийт талбай нэмэгдэн давтамж буурсан нь гурван горимын харьцуулалтаас тодорхой харагдав (Зураг 3).

Цацаргалтын хэлбэрийн загварчлалаас харахад антенн бүх горимд broadside чиглэлтэй, нэг чиглэлт (unidirectional) шинж чанараа хадгалж байгаа бөгөөд хажуу болон ар талын дэлбээний түвшин бага байгаа нь цацаргалтын үр ашиг сайн болохыг баталж байна [15]. Энэ нь бичил туузан антены broadside цацаргалтын онолтой бүрэн нийцэж байна [1].

Дээрх онолын тооцоолол болон симуляцийн тоон үр дүнгүүдийн нийцэл нь 30×26мм² хэмжээтэй, FR-4 цахилгаан тусгаарлагч суурь дээр суурилсан энэхүү авсаархан загвар нь 3.6–6 ГГц давтамжийн мужид, тухайлбал 5G NR n77/n79 болон WiFi6 стандартуудад нэгэн зэрэг үр ашигтай ажиллах чадвартайг онол болон тооцооллын хоёр талаас нотолж байна [6]–[8]. Ирээдүйд энэхүү загварыг бодитоор үйлдвэрлэж, векторын сүлжээний анализатор (VNA)-аар хэмжилт хийж, симуляцийн үр дүнтэй харьцуулан загварчлалын нарийвчлалыг улам нотлох ажлыг үргэлжлүүлэн хийхээр төлөвлөж байна.

VI. ИЖИЛ ТӨСТЭЙ СУДАЛГААНЫ АЖЛЫН ТОЙМ

Хүснэгт IV: Ижил төстэй судалгааны харьцуулалт

Судалгаа	**Шинж чанар**	**Энэхүү өгүүлэлийн давуу тал**
Design and fabrication of a novel frequency-reconfigurable patch antenna for WiFi and 5G application	Нам давтамж 2.4–3.3 ГГц хооронд шилждэг	3.6–6 ГГц хооронд шилждэг, WiFi6 стандартийн давтамж хүртэл хамардаг
Frequency reconfigurable antenna based on epsilon negative transmission lines	Антенны хийцлэл төвөгтэй, 1.77–4.18 ГГц давтамжид ажилддаг	Үйлдвэрлэхэд хийцлэл энгийн, өргөн давтамжийн мужид ажилддаг
Reconfigurable Antennas for Modern Wireless Communication Systems	PIN диод ашигласан, зөвхөн on/off горимтой, диодын тоогоор свитч хэрэгтэй	Варактор диод ашигласан, оролтын хүчдлээр горимуудыг тасралтгүй өөрчилнө

Ном зүй

[1] C. A. Balanis, *Antenna Theory: Analysis and Design*, 4th ed. Hoboken, NJ, USA: Wiley, 2016.

[2] J. T. Bernhard, *Reconfigurable Antennas*, ser. Synthesis Lectures on Antennas, vol. 2. San Rafael, CA, USA: Morgan & Claypool, 2007.

[3] A. Khidre, K.-F. Lee, F. Yang, and A. Z. Elsherbeni, “Circular polarization reconfigurable wideband E-shaped patch antenna for wireless applications,” *IEEE Trans. Antennas Propag.*, vol. 61, no. 2, pp. 960– 964, Feb. 2013, doi: 10.1109/TAP.2012.2220319.

[4] D. Peroulis, K. Sarabandi, and L. P. B. Katehi, “Design of reconfigurable slot antennas,” *IEEE Trans. Antennas Propag.*, vol. 53, no. 2, pp. 645– 654, Feb. 2005, doi: 10.1109/TAP.2004.841339.

[5] C. Won, I. Kwon, and J. Choi, “Design and analysis of a varactor-tuned dual-band reconfigurable antenna for 5G NR sub-6GHz applications,” *IEEE Access*, vol. 9, pp. 120243–120253, 2021, doi: 10.1109/ACCESS.2021.3108658.

[6] 3GPP, “TS 38.104: NR; Base station (BS) radio transmission and reception,” Tech. Spec. 38.104 v17.6.0, 2022. [Online]. Available: https://www.3gpp.org/DynaReport/38104.htm

[7] IEEE, “IEEE Std 802.11ax-2021: Enhancements for high efficiency WLAN,” New York, NY, USA: IEEE, 2021, doi: 10.1109/IEEESTD.2021.9442429.

[8] IEEE, “IEEE Std 802.11be-2024: Extremely high throughput (EHT),” New York, NY, USA: IEEE, 2024, doi: 10.1109/IEEESTD.2024.10590669.

[9] S. Nikolaou *et al.*, “Pattern and frequency reconfigurable annular slot antenna using PIN diodes,” *IEEE Trans. Antennas Propag.*, vol. 54, no. 2, pp. 439–448, Feb. 2006, doi: 10.1109/TAP.2005.863398.

[10] S. Zhang, G. H. Huff, J. Feng, and J. T. Bernhard, “A pattern reconfigurable microstrip parasitic array,” *IEEE Trans. Antennas Propag.*, vol. 52, no. 10, pp. 2773–2776, Oct. 2004, doi: 10.1109/TAP.2004.834440.

[11] D. M. Pozar, *Microwave Engineering*, 4th ed. Hoboken, NJ, USA: Wiley, 2011.

[12] Dassault Systèmes, *CST Studio Suite 2022*. Vélizy-Villacoublay, France, 2022. [Online]. Available: https://www.3ds.com/products/ simulia/cst-studio-suite

[13] ANSYS Inc., *ANSYS HFSS 2022*. Canonsburg, PA, USA, 2022. [Online]. Available: https://www.ansys.com/products/electronics/ansys-hfss

[14] N. Nguyen-Trong, A. Piotrowski, and C. Fumeaux, “A frequency-reconfigurable dual-band low-profile antenna for 2.4/5.2/5.8GHz WLAN applications,” *IEEE Trans. Antennas Propag.*, vol. 65, no. 7, pp. 3336–3343, Jul. 2017, doi: 10.1109/TAP.2017.2700087.

[15] H. A. Majid, M. K. A. Rahim, M. R. Hamid, and M. F. Ismail, “Frequency reconfigurable microstrip patch-slot antenna with directional radiation pattern,” *Prog. Electromagn. Res.*, vol. 144, pp. 319–328, 2014, doi: 10.2528/PIER13111807.

[16] ITU-R, “Technical characteristics of IMT-2020 radio interfaces,” Recommendation ITU-R M.2150-0, Geneva, Switzerland, 2022. [Online]. Available: https://www.itu.int/rec/R-REC-M.2150/en

[17] “T Design of Frequency Reconfigurable Patch Antenna Using Pin Diodes at S-Band”

FSO холбооны системийн бит алдааны магадлалын үнэлгээ

1rd Өлзий-Орших Манжаа
Электроникийн тэнхим
ШУТИС, МХТС
Улаанбаатар, Монгол
ulziiorshikh@must.edu.mn

2nd Хүрэлбаатар Цэвээнжав
Электроникийн тэнхим)
ШУТИС, МХТС
Улаанбаатар, Монгол
khurelbaatarts77@gmail.com

3st Мөнхбаяр Адъяа
Холбооны инженерчлэлийн тэнхим
ШУТИС, МХТС
Улаанбаатар, Монгол
*a_munkhbayar@must.edu.mn

Хураангуй—Энэхүү судалгаагаар Чөлөөт орон зайн (FSO) оптик холбооны системийн бит алдааны магадлалыг шинжилсэн. Үр дүнгээс үзвэл, манай орны нөхцөлд агаар мандлын унтралт ойролцоогоор 18дБ/км бөгөөд дунд хэмжээний турбулент (1×10^{-14}) нөхцөлд Улаанбаатар хотод FSO системийг амжилттай нэвтрүүлэх боломжтой байна. Судалгаагаар турбулентын хэмжээнээс үүсэх сцинтилляцийн эффектийг анхаарах шаардлагатайг тогтоосон бөгөөд энэ чиглэлээр судалгаа ховор байгаа тул цаашид нарийвчилсан судалгаа хийх нь зүйтэй. Дундаж турбуленттай нөхцөлд зөвшөөрөгдөх унтралт ойролцоогоор 27дБ/км байгаа нь FSO системийг манай оронд хэрэгжүүлэх боломжтойг харуулж байна.

Түлхүүр үг—FSO-Чөлөөт орон зайн оптик холбоо, турбуленс, сцинтилляци, агаар мандлын унтралт.

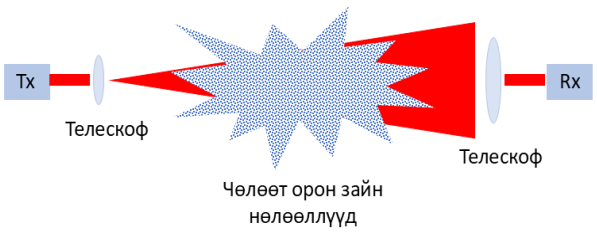
I. УДИРТГАЛ

Мэдээлэл харилцаа холбооны салбарын үсрэнгүй хөгжлийн явцад мэдээлэл дамжуулах найдвартай дамжуулах сүлжээний хэрэгцээтэй улам ихээр нэмэгдэх болсон. Энэхүү хэрэгцээг хангаж буй холбооны системийн нэг нь "Чөлөөт орон зайн оптик холбоо (FSO)"-ны систем [1] бөгөөд энэхүү систем нь өндөр хурд, найдвартай ажиллагаа, хялбар суурилуулалт зэрэг хэд хэдэн давуу талыг бий болгож байгаа. Энэ нь өндөр хурдны утасгүй технологийн хөгжилд томоохон хувьслыг бий болгож буй. Уламжлалт оптик холбооны системтэй харьцуулахад оптик кабелийг ашиглахгүй хэдий ч дамжуулах хурдны хувьд ойролцоо байдаг нь энэхүү технологийг орчин цагийн холбооны системд ашиглах байдлыг нэмэгдүүлж ирсэн [2].

Сүүлийн жилүүдэд мэдээллийн хэмжээ ихсэхийн хэрээр дамжуулах сүлжээний ачаалал ихсэх болсон. Хэрэглэгчийн сүлжээний өндөр хурд, өргөн зурвасыг хангах боломжтой гэдгээрээ судлаачид төдийгүй инженерүүд FSO системийг хэрэглэх, хөгжүүлэхэд анхаарлаа хандуулж байна. Энэхүү системийг нэвтрүүлэхэд тулгарч буй томоохон хүндрэлүүдийн нэг нь агаарын орчны нөлөөлөл бөгөөд судлаачид энэхүү нөлөөллийг багасгах аргуудыг өдөр бүр эрэлхийлэн хайж, анхаарлаа хандуулж байна. Манай орны хувьд энэхүү технологи нь шинэ орч ирж буй технологиудын нэг төдийгүй ашиглах хэрэгцээ шаардлага бий болж буйг өнөөдрийн өндөр хурдны хэрэгцээ шаардлага харуулж байна. Энэхүү судалгааны ажлаар манай орны байгаль цаг уурын нөхцөлд FSO холбооны системийн бит алдааны магадлалын үнэлгээг хийж товчхон дүгнэхийг зорьлоо.

II. FSO ХОЛБООНЫ СИСТЕМИЙН АГААР, ОРЧНЫ НӨЛӨӨЛӨЛ

FSO холбооны системийн ерөнхий бүтцийг зураг 1-т үзүүлэв. Зурагт үзүүлснээр уг системд мэдээлэл нь агаарын чөлөөт орон зайд гэрлэн дохиогоор дамжигдах бөгөөд нэвтрүүлэгч болон хүлээн авах тал LoS (Line of Sight) буюу бие биенээ харсан байрлалд байхыг шаарддаг. Энэхүү зурагт үзүүлсэнчлэн уг системийн



Зураг 1. FSO холбооны системийн ерөнхий бүтэц

дамжууллын чанарт агаар, орчны нөлөөлөл ихээхэн нөлөөллийг үзүүлэх бөгөөд энэхүү бүлэгт уг нөлөөллийн тухай ойлголтыг өгөх болно.

A. Атмосферын унтралт

Холбооны системийн чухал нэг параметрын нэг нь хүлээн авах дохионы чадал юм. Чөлөөт орон зайн оптик холбооны системд хүлээн авах дохионы чадал геометр алдагдал болон атмосферын унтралт зэргээс хамааран дараах томъёогоор тодорхойлно. [3]

$$P_r = P_t \left(\frac{D_r}{D_t + L\,\theta} \right)^2 10^{-\alpha\, L/10} \qquad (1)$$

Үүнд:

- P_r - Хүлээн авсан оптик чадал (Вт)
- P_t - Дамжуулсан оптик чадал (Вт)
- D_t - Дамжуулагчийн апертур диаметр (м)

- D_r - Хүлээн авагчийн апертур диаметр (м)
- L - Шугамын зай(км)
- θ - Дохионы дивергенцийн өнцөг (рад)
- α - Атмосферын унтралтын коэфициент (дБ/км)
- λ - Долгионы урт(нм)
- V - Үзэгдэц (км)

Геометр алдагдал: $\left(\frac{D_r}{D_t+L\,\theta}\right)^2$ нь шугамын дамжуулах зайнаас шалтгаалан үүсэж буй дохионы дивергенц болон нэвтрүүлэх, хүлээн авах апертуриас хамаарч илэрхийлэгдэж байна. Харин томъёо 1-ийн арын хэсэг болох атмосферын нөлөөллөөс үүсэх дохионы унтралтын чухал параметр болох α) нь доорхи томъёогоор илэрхийлэгдэнэ.

$$\alpha_{\text{нийт}} = \alpha_{\text{үзэгдэх}} + \alpha_{\text{цас}} + \alpha_{\text{бороо}} \tag{2}$$

Чөлөөт орон зайн оптик холбооны системд атмосферын нийт унтралтын коэфициент нь үзэгдэц (энэ нь манан будан, тоосонцор, агаарын боирдол зэргээс үүсэх унтралт), бороо, цас зэргээс үүсэх унтралтын коэфициентуудын нийлбэрээр тодорхойлно.

1) Үзэгдэцээс хамаарсан унтралтын коэфициент: Үзэгдэцээс хамаарсан унтралтын коэфициентыг тодорхойлох хамгийн түгээмэл хэрэглэгдэж байгаа 2 төрлийн загвар байна (Kruse model (1970)), (Kim model(2001)). Хоёр загварын үед уг параметрын ерөнхий илэрхийлэл нь [4], [5]:

$$\alpha(\lambda) = \frac{3.91}{V}\left(\frac{\lambda}{550\ \text{нм}}\right)^{-q(V)} \tag{3}$$

Үүнд:

- V - Үзэгдэц [км]
- λ -долгионы урт [нм],
- $q(V)$ - Үзэгдэх зайнаас хамаарах тархалтын үзүүлэлт

$q(V)$-г хоёр загварын доор үзүүлсэн утгаар авах бөгөөд Kim-н загвар нь үзэгдэх зайг бага хэмжээтэй буюу их хэмжээний манан, будан, агаарын бохирдолтой байх нөхцлийг авч үзсэнээрээ Kruse-н загварын сайжруулсан хувилбар болон хөгжсөн.

Kruse model:

$$q(V) = \begin{cases} 0.585V^{1/3}, & V < 6 \text{ км} \\ 1.3, & 6 \le V \le 50 \text{ км} \\ 1.6, & V > 50 \text{ км} \end{cases}$$

Kim model:

$$q(V) = \begin{cases} 0, & V < 0.5 \text{ км} \\ V - 0.5, & 0.5 < V < 1 \text{ км} \\ 0.16\,V + 0.34, & 1 < V < 6 \text{ км} \\ 1.3, & 6 \le V \le 50 \text{ км} \\ 1.6, & V > 50 \text{ км} \end{cases}$$

2) Бороо болон цаснаас хамаарсан унтралтын коэфициент: Бороо, цаснаас хамаарах унтралтын коэфициент нь хур тунадасны хэмжээ болон дамжуулах долгионы уртаас хамаарч дараах томъёогоор тодорхойлно. Борооны унтралтын коэфициент [6]:

$$\alpha_{\text{rain}} = k_R\, R^{\beta_R} \quad [\text{дБ/км }] \tag{4}$$

Үүнд:

- R - Хур тунадасны хэмжээ [мм/цаг]
- k_R,1550k$_R$,β_R-ийн хэмжээг дараах утгатайгаар сонгон авна.
  - $k_R \approx 1.076 \times 10^{-3}$
  - $\beta_R \approx 0.67$

  Цасны унтралтын коэфициент [?]:

$$\alpha_{\text{snow}} = a_s\, S^{b_s} \quad [\text{дБ/км}] \tag{5}$$

  Үүнд:
  - S - Хур тунадасны хэмжээ [мм/цаг]
  - a_s,b_s -долгионы уртаас хамаарсан тогтмол

  Оптик долгиоы хувьд a_s,b_s-ийн хэмжээг дараах утгатайгаар сонгон авна.
  - $a_s \approx 1 \times 10^{-4}$
  - $b_s \approx 2$

В. Сцинтилляцийн эффект

Сцинтилляцийн эффект нь гэрлийн эрчимжилтэд нөлөөлдөг томоохон нөлөөллүүдийн нэг бөгөөд энэхүү эффектийн тархалтын функцын хэлбэрээр нь хэд хэдэн ялгаатай загварууд хэрэглэгдэж байгаа. Эдгээрийн дотроос хамгийн түгээмэл хэрэглэгдэж байгаа тархалтын төрлүүд болох Log-Normal болон Gamma-Gamma тархалтын функцууд бүхий загварыг энэхүү судалгааны ажилд авч ашиглах болно. Сцинтилляцийн эффектийн нөлөөллийг дараах томъёогоор тодорхойлно.

$$\sigma_I^2 = \frac{\langle I^2\rangle - \langle I\rangle^2}{\langle I\rangle^2} \tag{6}$$

Үүнд: I хүлээн авч буй дохионы эрчимжилт,$\langle I\rangle$ дундаж эрчимжилт.

Хүлээн авч буй дохионы эрчимжилтийн магадалал тархалт нь турбуленсын хэмжээнээс хамааран өмнө дурдсан хоёр загварыг ашигладаг: Log-Normal тархалт бүхий загвар нь бага хэмжээний турбуленстэй нөхцөлд, Gamma-Gamma тархалт бүхий загвар нь дундаж, их хэмжээний турбуленстэй нөхцлүүдэд.

1) Log-Normal загвар : Сул турбуленстэй тохиолдолд хүлээн авсан дохионы эрчимжилт I нь log-normal тархалт бүхий загварыг ашиглавал тохиромжтой [7]:

$$f_I(I) = \frac{1}{2I\sqrt{2\pi\sigma_X^2}}\exp\left[-\frac{(\ln I - \mu_X)^2}{2\sigma_X^2}\right], \quad I > 0 \tag{7}$$

Үүнд: μ_X - дундаж лог-агууураг, σ_X^2 -агуургын хэлбэлзэл, ($\sigma_X^2 \approx \sigma_R^2$).

2) Gamma-Gamma загвар: Дунд - хүчтэй турбуленсын нөхцөлд Gamma-Gamma тархалтыг ашиглана [8]–[10]:

$$f_I(I) = \frac{2(\alpha\beta)^{(\alpha+\beta)/2}}{\Gamma(\alpha)\Gamma(\beta)} I^{(\alpha+\beta)/2-1} \times K_{\alpha-\beta}\left(2\sqrt{\alpha\beta I}\right), \quad I > 0 \tag{8}$$

Үүнд:

- α - Турбуленсын хамгийн жижиг хэсгийн хэмжээ [м]
- β -Турбуленсын хамгийн том хэсгийн хэмжээ [м]
- $\Gamma(\cdot)$ - Гамма функц
- $K_{\alpha-\beta}(\cdot)$ - Бесселийн функц

α болон β Rytov-н параметрээс хамааран дараах байдлаар тодорхойлогдоно:

$$\alpha = \left[\exp\left(\frac{0.49\,\sigma_R^2}{(1+1.11\,\sigma_R^{12/5})^{7/6}}\right) - 1\right]^{-1} \tag{9}$$

$$\beta = \left[\exp\left(\frac{0.51\,\sigma_R^2}{(1+0.69\,\sigma_R^{12/5})^{5/6}}\right) - 1\right]^{-1} \tag{10}$$

Үүнд: σ_R^2 нь Rytov-н параметр бөгөөд томъёо 11 -аар тодорхойлогдоно.

$$\sigma_R^2 = 1.23\,C_n^2\,k^{7/6}\,L^{11/6}, \quad k = \frac{2\pi}{\lambda} \tag{11}$$

Үүнд:

- C_n^2 хугарлын индекс (м$^{-2/3}$),
- $k = \frac{2\pi}{\lambda}$ оптик дохионы тоо (рад/м),
- L шугамын зай (м),
- λ долгионы урт (м).

III. УЛААНБААТАР ХОТЫН ЦАГ АГААРЫН НӨХЦЛИЙН ШИНЖИЛГЭЭ

Монгол орны хувьд эрс тэс уур амьсгалтай, жилийн дөрвөн улирал ээлжлэн байдаг нь FSO холбооны системийг хэрэгжүүлэхэд цаг агаарын олон нөхцлийг судлан авч үзэх шаардлагыг бий болгож байдаг. Цас, бороо, манан, агаарын бохирдол зэрэг олон нөхцөл байдлыг судлан үзэн тэдгээрийн өргөн утганд тооцоолол хийхийг шаардаж байна. Улаанбаатар хотын цаг агаарын хэмжилтийн үр дүнг хүснэгт I-т харуулав.

Хүснэгт I-т үзүүлсэн цаг агаарын код гэдэг нь олон улсын цаг агаарын датаг кодчилон харуулсан утга бөгөөд манай орны хувьд дунджаар 40 буюу үүлэрхэг төлвийг харуулж байгаа болно. Уг хүснэгтэд тоон утгуудад харгалзах унтралтын томъёо 3-5-г ашиглан тооцоолсныг үснэгт II-т үзүүлэв.

Хүснэгт I
Цаг агаарын өгөгдөл (2022–2025)

Нөхцөл	Мин.	Макс.	Дун.	тайлбар
Үзэгдэц. (км)	1.0	20.0	10.2	Өвөл бага (цас/үзэгдэц), зун их
Бороо (мм)	0	20	4.3	Зун хамгийн өндөр
Цас (мм)	0	15	3.8	Өвөл хамгийн өндөр
Цаг агаарын код	0	86	~ 40	Өвөл бага (цас/үзэгдэц), зун их

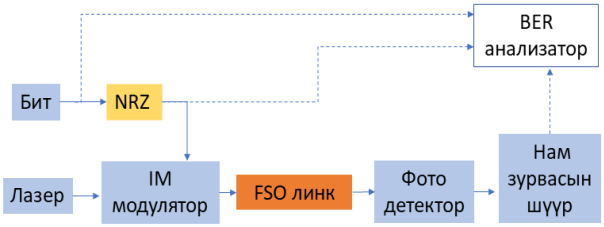
Хүснэгт II
$\lambda = 1550$ нм дээрх атмосферын унтралт

Нөхцөл	Мин. (дБ/км)	Макс. (дБ/км)	Дун. (дБ/км)
Манан/Тоосонцор/ Агаарын бохирдол (Kim - загвар)	0.2	15.0	1.5
Бороо	0.0	3.0	0.8
Цас	0.0	2.5	0.6
Нийт	0.2	18.0	2.3

Хүснэгт II-т үзүүлсэн нийт унтралтыг томъёо 2-оор тооцоолсон бөгөөд бороо болон цасны унтралтын аль нэг байна гэж үзэн аль их утгаар нь авсан болно.

IV. Симуляцийн хэсэг

FSO холбооны системийн бит алдааны магадлалын үнэлгээг хийхийн тулд зураг 2-т үзүүлсний дагуу симуляцийн хэлхээг угсарч гүйцэтгэсэн бөгөөд симуляцийг оптик холбооны системийн симуляцийн програм болох optiSystem 18 програм дээр хийсэн. Симуляцийн үндсэн тохиргоог хүснэгт III-т үзүүлэв.



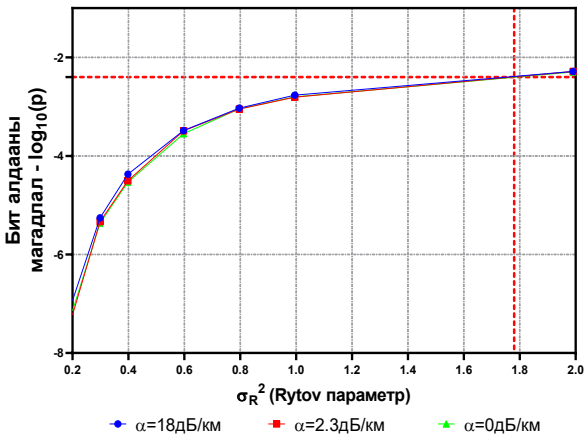
Зураг 2. Симуляцийн ерөнхий схем

Симуляцийн үр дүнг зураг 3 болон 4-т үзүүлэв. Зураг 3-т бит алдааны магадлал (BER) болон Rytov параметр σ_R^2 -н хамаарлыг харуулж байна. Энэхүү параметр нь агаарын турбулент байдлын хүчийг илэрхийлдэг. BER нь логарифм хэлбэрт үзүүлэгдсэн бөгөөд $\alpha = 0, 2.3, 18$ дБ/км гэсэн гурван алдагдлын коэффициентын хувьд харуулсан. σ_R^2 нэмэгдэх тусам BER эрс мууддаг нь, турбуленсын хэмжээ ихсэх тусм системийн гүйцэтгэл муудаж буйг үзүүлж байна. Харин алдагдлын нөлөө

Хүснэгт III
FSO системийн параметрүүд

№	Параметр	Утга
1	FSO шугамын урт	1000 м
2	Нэвтрүүлэгчийн диаметр (D_t)	5 см
3	Хүлээн авагчийн диаметр (D_r)	20 см
4	Геометр алдагдал	-20.2145 дБ
5	Унтралт	0–30 дБ/км
6	Хугарлын индекс (C_n^2)	$10^{-15} - 4 \times 10^{-14}$
7	Бит хурд	10 Гбит/с
8	Модуляци	IM

харьцангуй бага бөгөөд бүх гурван муруй нь бүхий л σ_R^2 завсарт бие биедээ ойрхон байгаагаас харж болно. Хэвтээ тасархай шугамаар зөвшөөрөгдөх BER-ийн түвшинг ($\approx 10^{-2.4}$, FEC) дүрсэлсэн ба энэ үеийн Rytov параметрын утгыг босоо тасархай шугамаар дүрслэн харуулав. Зураг үзүүлснээр энэ түвшинд ойролцоогоор $\sigma_R^2 \approx 1.8$-д харагдах бөгөөд энэ түвшнээс дээш турбуленстэй нөхцөлд систем гүйцэтгэл чанарын шаардлагыг хангахгүй харуулж байна. Хүснэгт IV-т агаарын хугарлын индекс C_n^2-н утганд харгалзах σ_R^2-н утгын тооцоолын үр дүнг харуулав.

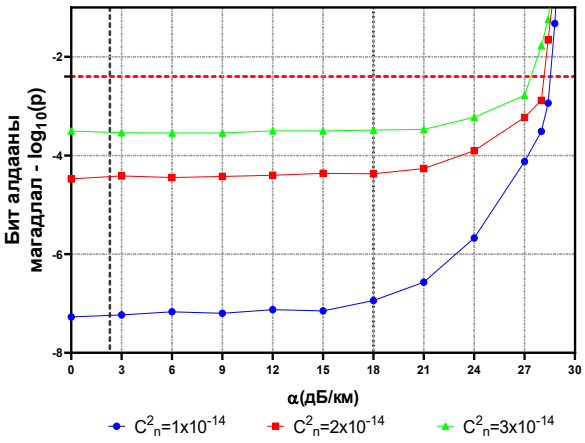


Зураг 3. Бит алдааны магадлал Rytov параметрээс хамаарсан хамаарал

Зураг 4-т бит алдааны магадлал унтралтын коэфициентээс хамаарсан хамаарлыг үзүүлэв. Үр дүнгээс харахад α ихсэх тусам бит алдааны магадлал өсч, системийн гүйцэтгэл буурч байна. Мөн хугарлын индексийн параметр C_n^2 ихсэх тусам (өөрөөр хэлбэл турбуленс хүчтэй болох үед) BER илүү муудаг нь ажиглагдаж байна. Ялангуяа $\alpha \approx 25$–30 дБ/км мужид системийн гүйцэтгэл огцом буурч, бит алдааны магадлал эрс нэмэгдэж байна. Энэ нь хүчтэй атмосферийн нөлөөн дор FSO системийн найдвартай ажиллагаа алдагдаж байгааг харуулж байна. Зураг 3 дээрхитэй ижлээр хэвтээ тасархай шугамаар зөвшөөрөгдөх BER-ийн

Хүснэгт IV
Rytov variance σ_R^2 for different turbulence strengths C

C ($m^{-2/3}$)	1E-15	5E-15	1E-14	1.5E-14
σ_R^2	0.0199	0.0995	0.1991	0.2986
C ($m^{-2/3}$)	2E-14	3E-14	4E-14	5E-14
σ_R^2	0.3982	0.5973	0.7964	0.9955
C ($m^{-2/3}$)	1E-13	2.5E-13	5E-13	7.5E-13
σ_R^2	1.9910	4.9774	9.9548	14.9322
C ($m^{-2/3}$)	1E-12	–	–	–
σ_R^2	19.9095	–	–	–



Зураг 4. Бит алдааны магадлал унтралтын коэфициентээс хамаарсан хамаарал

түвшинг илэрхийлсэн бөгөөд 3 ялгаатай турбуленстэй нөхцөлд ойролцоогоор 27дБ/км унтралттай байх үед системийн ажиллагааг найдвартай хангаж болохоор байна. Манай орны хувьд хүснэгт II-т үзүүлснээр хамгийн их (18дБ/км) болон дундаж (2.3дБ/км) унтралттай үеүүдэд дээрх 3 тохиолдолд систем найдвартай ажиллаж болох нь харагдаж байна.

V. Дүгнэлт

Чөлөөт орон зайн оптик холбооны системийн бит алдааны магадлалын шинжилгээг энэхүү судалгааны ажлаар хийж гүйцэтгэлээ. Судалгааны үр дүнгээс харахад манай орны хувьд агаар мандлын нөлөөллөөс үүсэх унтралтын хэмжээ 18дБ/км гэж үзсэн бөгөөд уг унтралттай үед дунд хэмжээний турбуленс (110^-14) байх нөхцөлд Улаанбаатар хотод FSO системийг нэвтрүүлэх бүрэн боломжтой гэж дүгнэсэн. Чөлөөт орон зайн оптик холбоог нэвтрүүлэхэд турбуленсын хэмжээнээс хамаарсан сцинтилляцийн эффектийг чухлаар авч үзэх шаардлагатай нь бит алдааны магадлалын шинжилгээнээс харагдаж байна. Манай

орны хувьд сцинтилляцийн эффектийг судалсан судалгаа ховор хийгдсэн нь цаашид судлах судлагдахуун болон үлдэж буйг хэлэх нь зүйтэй болов уу. Дундаж хэмжээний турбуленстэй нөхцөлд агаар мандлын нөлөөллөөс үүсэх унтралтын зөвшөөрөгдөх хэмжээ $\approx$ 27/ байгаа нь Манай орны хувьд FSO системийг нэвтрүүлэх, хэрэгжүүлэх бүрэн боломжтойг харуулж байна.

Ашигласан материал

[1] Khalighi, M. A., & Uysal, M. (2014). Survey on free space optical communication: A communication theory perspective. IEEE Communications Surveys & Tutorials, 16(4), 2231–2258.

[2] Kaushal, H., & Kaddoum, G. (2017). Optical communication in space: Challenges and mitigation techniques. IEEE Communications Surveys & Tutorials, 19(1), 57–96.

[3] Bloom, S., Korevaar, E., Schuster, J., & Willebrand, H. (2003). Understanding the performance of free-space optics. Journal of Optical Networking, 2(6), 178–200.

[4] Kruse, P. W. (1970). Visibility and Atmospheric Transmission. Applied Optics.

[5] Hemani, A., Sharma, S., & Kumar, D. (2019). Performance analysis of FSO communication system under atmospheric turbulence. Optik, 183, 489–499.

[6] V. K. Jain, S. Kar, and H. Kaushal, Free Space Optical Communication, 1st ed., Optical Networks. Springer India, 2017.

[7] H. Moradi, M. Falahpour, H. H. Refai, P. G. LoPresti and M. Atiquzzaman, "BER analysis of optical wireless signals through log-normal fading channels with perfect CSI," Telecommunications (ICT), 2010 IEEE 17th International Conference on, Doha, 2010, pp. 493-497.

[8] Larry C. Andrews and Ronald L. Phillips, Laser Beam Propagation Through Random Media, 2nd edition, SPIE Press Book, 2005.

[9] Larry C. Andrews, Ronald L. Phillips, and Cynthia Y. Hopen, Laser Beam Scintillation with Applications, SPIE Press Book, 2001.

[10] A. A. Farid and S. Hranilovic, "Outage Capacity Optimization for Free-Space Optical Links With Pointing Errors," Journal of Lightwave Technology, vol. 25, no. 7, pp. 1702–1710, July 2007.

IPv6 шилжилтийн механизмын гүйцэтгэлийн харьцуулалт: их сургуулийн кампус сүлжээнд

П.Баясгалан, Б.Энхбаяр, Ц.Давгасүрэн*, Д.Баттулга, Х.Золзаяа
Электроник, Холбооны Инженерчлэлийн Тэнхим
Мэдээллийн Технологи, Электроникийн Сургууль, Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
zolzayakhh@num.edu.mn

***Хураангуй*—Интернэт орчин IPv6-руу шилжих шаардлага бий болоод удаж байна. Гэсэн хэдий ч IPv4 -ийн хэрэглээ буурахгүй байгаа нь энэхүү шилжилтийг удаашруулж байгаа билээ. Энэхүү ажлаараа их сургуулийн кампус сүлжээг IPv6-only гэсэн төлөвт шилжүүлэхэд зорилготой хийгдэж буй алхамуудыг танилцуулж, тэдгээрийг нэвтрүүлэх үеийн гүйцэтгэлийг шинжилэх, харьцуулалт хийнэ. Бид эхлээд кампус сүлжээний хуулбар туршилтын орчин бүрдүүлж, хоёр хаяглалтын протокол болон шилжилтийн үе шатуудын хооронд round trip time (RTT) хоцрогдол, цэгээс цэгт нэвтрүүлэх чадамж гэсэн үзүүлэлтүүдээр харьцуулан үнэлсэн. Үр дүнгээс харахад дундаж хоцрогдол ойролцоо байсан ч IPv4 бага зэрэг илүү гарсан. RTT-ийн хувьд IPv6 нь IPv4-тэй (163.72 ms) харьцуулахад их (194.85 ms) байсан. Ялангуяа сүлжээний нэвтрүүлэх чадамжийн хувьд IPv6-only төлөв нь харьцангуй хурдан буюу хамгийн ихдээ 7571 KBps гэсэн үр дүн үзүүлсэн. IPv6 -руу шилжихэд сүлжээний дэд бүтцээс илүү системүүдийн шинэчлэлийг хурдан гүйцэтгэх хэрэгтэй болох нь ажиглагдсан. Төгсгөлд нь бодит сүлжээнд IPv4-only болон IPv6-only сүлжээнүүдийг байгуулж Интернэт байрлах сервертэй файл солилцох үеийн харьцуулалтыг хийсэн.**

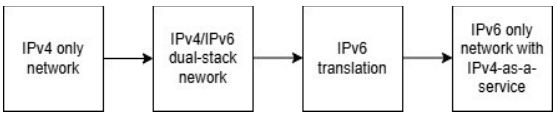
***Түлхүүр үг*—IPv6 transition, adoption, dual-stack, tunnelling, translation, Network architecture,**

I. Оршил

Өнөөдрийн интернэтийн хэрэглээ хэт их болсноос үүдэн сүлжээний орчин IPv4 хаягийн нөөцийн хомсдлоос IPv6 рүү шилжих нь техникийн зайлшгүй шаардлага болоод байна. Мөн нийтийн IPv4 (public) хаягийн хомсдол үүсээд удаж байгаа билээ [1]. Интернэтэд суурилсан үйлчилгээ, түүний хэрэглээ болон төхөөрөмжийн тоо эрчимтэй өсөн нэмэгдэж буй нь энэ асуудлыг улам хурцатгаж, IPv6-руу шилжихийг зайлшгүй болгож байна [2]. Мэдээж шилжилт хийснээр IPv6 -ийн олон давуу талыг ашиглах боломжтой болно.

Сүлжээний дэд бүтцээ IPv4 ба IPv6 зэрэгцэн ажилладаг хоёр стек бүхий (dual-stack) байхаар тохируулах нь энэхүү шилжилтийн энгийн ба эхний шат бөгөөд хуучин үйлчилгээний нийцтэй байдлыг хангах архитектур юм [3]. Dual-stack-аар сүлжээг ажлуулах нь зөвхөн хуучин системтэй нийцтэй байдлыг хангаад зогсохгүй, IPv6-ийн өргөтгөсөн боломжид тулгуурладаг шинэ үйлчилгээнүүдийг мөн дэмждэг [4]. Гэхдээ dual-stack нь шилжилтийн дундын шат учир байгуулгууд бүрэн IPv6 буюу IPv6-only архитектурруу чиглэсэн стратеги төлөвлөгөө боловсруулах хэрэгтэй.

Өөрөөр хэлбэл, IPv6 нэвтрүүлэх шилжилт нь зөвхөн шинэчлэлт биш, харин сүлжээний тасралтгүй ажиллагаа болон дэлхийг хамарсан хамгийн том сүлжээ болох Интернэттэй харилцан ажиллах чадвар, боломжийг баталгаажуулах зорилготой урт хугацаанд стратеги төлөвлөгөө юм [5]. Энэ үүднээс авч үзвэл сүлжээг зөвхөн IPv6 дангаар (IPv6-only) ажиллах архитектурт шилжүүлэхийн тулд dual-stack -ийн дараа хэд хэдэн алхам гүйцэтгэх шаардлагатай юм. Зураг 1 -т эдгээр dual-stack, Туннел, Хөрвүүлэлт гэсэн 3 үндсэн алхам болон IPv6-only сүлжээнд IPv4 үйлчилгээг үзүүлэх хувилбар гэсэн алхамуудыг тоймлон үзүүлэв. Эдгээр үндсэн алхамуудаас гадна дундын алхамууд нэмэгдэж болно. Бүгдийг нь нийлүүлбэл Dual-stack бүхий хэрэглэгчид, IPv6-only pilot буюу хаяглалтыг туслаа шийдсэн дэд бүтэц бүхий сүлжээ, IPv4-to-IPv6 туннел буюу IPv6 -тай дэд бүтэц дээгүүрээ IPv4 -ийн урсгалыг капсул хийн дамжуулах, IPv6 голчилсон буюу сүлжээний ихэнх үйлчилгээ болон траффик IPv6 -аар дамжих, тусгаарлалт буюу хуучин IPv6 дэмжихгүй төхөөрөмж болон үйлчилгээг тусгаарлаж дэд сүлжээ үүсгэх, IPv6-only буюу сүлжээний анхдагч бүх тохиргоо IPv6-аар хийгдэх (жишээ нь DHCP, DNS гэх мэт) юм.



Зураг 1 IPv4-only to IPv6-only transition

Энэхүү судалгааны ажлын үндсэн зорилго нь их сургуулийн кампус сүлжээнд IPv6 нэвтрүүлэх, шилжилт хийхдээ сүлжээний дэд бүтцэд IPv4 болон IPv6 хаяглалттай үеийн үр ашгийг шинжлэхэд оршино. Мөн туршилтын орчинд байгуулсан кампус сүлжээний загвартаа IPv6 шилжилтийн үе шатуудын хооронд харьцуулалт хийнэ. Ингэснээр тухайн үе шатыг нэвтрүүлэх, гүйцэтгэх, үргэжлэх хугацааг тогтоож шилжилийн стратеги төлөвлөгөөг бодитоор гаргах боломжтой болно.

Энэхүү өгүүллийн бүтэц нь хоёрдугаар бүлэгт IPv6 руу хийх шилжилтийн талаар суурь ойлголт болон судлагдсан байдал, гуравдугаар бүлэгт туршилтын сүлжээний бүтцийн талаар тайлбар, дөрөвдүгээр бүлэгт ба туршилтын үр дүн, төгсгөлийн тавдугаар бүлэгт дүгнэлтийг багтаасан.

II. IPv6 ШИЛЖИЛТ

A. IPv6 хаяглалт

IPv6 нь 32 битийн урттай байсан хаяглалтыг 128 бит болгон нэмэгдүүлж хаягийн нөөцийг ихээр нэмэгдүүлсэн бөгөөд ингэснээр IPv4 хувилбарынхаа хязгаарлагдмал байдлыг шийдвэрлэж, дижитал шилжилтийн үеийн хөгжлийн тасралтгүй байдлыг хангах боломжтой болгодог [4,5]. Өөрөөр хэлбэл IPv6 нь хялбаршуулсан толгойн бүтэц, хөдөлгөөнт төхөөрөмжийн илүү сайн дэмжэх зэргээр өнөөгийн Интернэтийн өөрчлөлтийн онцлогийг тусган гүйцэтгэл болон өргөтгөх чадварыг сайжруулахад чухал юм [6].

Мөн IPv6 нь сайжруулсан multicast чиглүүлэлт, төлөвгүй хаягийн автомат тохируулга (Stateless Address Autoconfiguration, SLAAC) зэрэг илүү боломжуудыг сүлжээнд нэвтрүүлдэг бөгөөд эдгээр нь QoS-ийг сайжруулах боломжийг бүрдүүлдэг [7].

B. Dual-stack

IPv4 хаягийн нөөц болон хэрэглээнд үүсгэх давуу тал бүрэн шавхагдсан тул IPv4-өөс IPv6 рүү шилжих нь техникийн зайлшгүй шаардлага болсон [3,8]. Сүлжээнд анхлан IPv6-ийг нэвтрүүлж эхлээд хаяглалтын чадавхыг өргөжүүлж, чиглүүлэлтийн үр ашгийг нэмэгдүүлэх зэргээр ахиц гарах боловч IPv4 ба IPv6 зэрэгцэн ажлуулах dual-stack орчинд үйл ажиллагаа болон төлөвлөлтийн сорилтууд бий болдог [9,10].

Энэхүү хоёр хаяглалт зэрэгцэн орших байдлыг үр дүнтэй удирдаж байж сүлжээний тасралтгүй холболтыг хангах, латенсийг бууруулах, үйлчилгээний чанарыг өндөр түвшинд хадгалах үндэс болдог. Мөн эдгээр нь хэрэглэгчийн сэтгэл ханамжид шийдвэрлэх нөлөөтэй хүчин зүйлс юм. Гэвч dual-stack байдлаар сүлжээг ажлуулах үед чиглүүлэлт, NAT, мониторинг гэх мэт олон сүлжээний функцүүдийн гүйцэтгэх үүрэг, ачаалал нэмэгдэх, зарим тохиолдолд функцүүдийг хоёр хаяглалтанд зориулж тусад нь ажлуулах хэрэгцээ гарна. Мөн шилжилтийн механизмын дэмжих хуучин төхөөрөмж нь төвөгтэй байдал нэмэгдүүлнэ.

C. Туннел / Encapsulation

Энэ арга нь нэг протоколыг нөгөөгийн дотор зөөвөрлөн ажиллана. Шилжилтийн үед хамгийн түгээмэл нь IPv4-ийг пакетийг IPv6 -ийн өгөгдөл зөөх талбарт агуулан дамжуулах юм [11]. Үүний нэг хэлбэр нь DS-Lite (Dual-Stack Lite) бөгөөд хэрэглэгчийн IPv4 пакетуудыг зөвхөн IPv6-тай хандалтын сүлжээгээр дамжуулахын тулд IPv6 дотор капсуллах, дараа нь гарц дээр carrier-side NAT хийдэг [12]. Өөрөөр хэлбэл хандалтын сүлжээ IPv6-тай ажиллаж, IPv4 траффикийг IPv6 дээгүүр туннеллэгдэнэ, Интернэт рүү гарах үед NAT хийнэ.

Энэхүү шилжилтийн шатыг сүлжээний дэд бүтцээ IPv6 суурьтай болгоод IPv4 суурьтай үйлчилгээг үргэлжлүүлэн үзүүлэх боломжтой байхаар шийдэх үед хэрэглэнэ. Давуу талууд нь хандалтын сүлжээнд IPv4 хэрэглээг буурна, дэд бүтэц IPv6 болно, IPv4 холболтыг хадгална. Харин сул тал нь капсуллах болон түүнээс үүдэлтэй нэмэлт ачаалал үүснэ, NAT-аас бүрэн хамаарна, хянахад төвөгтэй болно.

Өөр нэг хэлбэр бол MAP-E юм [13]. Энэ нь Mapping of Address and Port using Encapsulation гэсэн үгний товчлол бөгөөд IPv6 сүлжээгээр IPv4 үйлчилгээг хүргэхдээ капсуллах болон IPv4 хаяг/портыг алгоритмын дагуу IPv6-ийн хаяг болгох өөрчилдөг. Давуу тал нь NAT state-ийг бууруулж чадна, өргөтгөх боломжтой. Харин сул тал нь үйл ажиллагааны хувьд алгоритмаас хамааралтай болж илүү төвөгтэй байдал үүснэ, нийцтэй төхөөрөмж шаардана.

D. Хөрвүүлэлт (Translation)

Хөрвүүлэлт нь нэг протоколыг нөгөөгийнх нь дотор зөөх биш, харин IPv4 ба IPv6 хооронд шууд хөрвүүлэлт хийх арга юм. Үүний эхний хэлбэр бол NAT64, DNS64 сүлжээний функцуудийг ашиглах юм [14]. NAT64 нь сүлжээний талд IPv6 ба IPv4 хооронд хөрвүүлэлт хийдэг. DNS64 нь IPv4-ийн A record-оос хиймэл AAAA record үүсгэж, IPv6-only клиентүүд IPv4-only серверт хандах боломж олгодог.

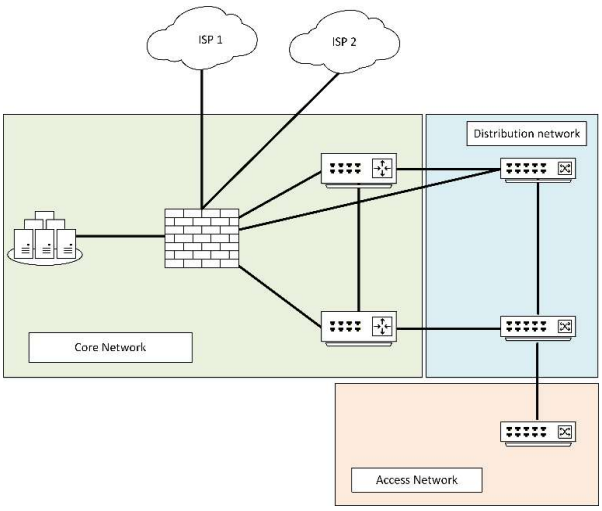
Ихэвчлэн IPv6-only хандалтын сүлжээ, Мобайл сүлжээ, Байгууллагын буюу campus-ийн IPv6-only байгуулахад ашиглана [15,16]. Давуу тал нь клиент талд IPv4 хэрэглэхгүй, IPv6-only сүлжээ болно, IPv6 хэрэглээ нэмэгдэх тусам хялбар болно. Сул тал нь аппликейшн “literal IPv4 address” ашиглавал асуудал гарна, DNS-ээс хамаардаггүй аппликейшнүүдийг зохицуулах төвөгтэй, зарим хуучин протокол, аппликейшн ийм хөрвүүлэлтийн үед зөв ажиллахгүй байх магадлалтай.

Өөр хувилбарууд нь 464XLAT, MAP-T юм [17]. NAT64 өргөтгөж, IPv6-only сүлжээнд байгаа IPv4-only аппликейшнүүдийг илүү сайн дэмжих зорилготой механизм юм. MAP-T нь MAP-E-ээс капсуллах үйлдлийн оронд stateless хөрвүүлэлт хийдэг. Үйлчилгээ болон аппликейш нь IPv6 биш ч IPv6-only хандалтын сүлжээ ажиллуулахыг хэрэгтэй болвол энэ механизм маш хэрэгтэй.

III. ТУРШИЛТЫН СҮЛЖЭЭНИЙ АРХИТЕКТУР

Туршилт гүйцэтгэх зорилгоор их сургуулийн кампус сүлжээг дууриалган симуляцийн орчинд байгуулсанг тоймлон Зураг 2-т үзүүлэв. Сүлжээний топологийг гадаад (ISPs), цөм, түгээх болон хандалтын сүлжээ гэсэн байдлаар байгуулсан. Хандалтын сүлжээг хэрэглээ болон үүргийн дагуу 8 тусдаа VLAN болгон хуваасан. Мөн VLAN тус бүрдээ NAT, DHCP, чиглүүлэлт, гадаад сүлжээний гарц өөр байхаар шийдсэн. Түгээх сүлжээний хувьд цөм болон хандалтын сүлжээ хоорондын холболтыг цэгээс цэгт холболтоор хангана. Цөм сүлжээ нь аюулгүй байдлыг хангах галтхана, чиглүүлэлт болон NAT -ийг гүйцэтгэх рүтерүүд, бүх дэд сүлжээний зохицуулагчийн үүрэг бүхий цөм свичүүдээс бүрдэнэ. Клиент төхөөрөмжийн талд хандалтын хэсгээс бусад хэсэгт шилэн кабелийн холболттой учир физик давхаргын хувьд асуудал багатай гэж үзээд туршилтын орчинд бүрэн тусгаагүй болно. IP хаяглалтыг IPv4 болон IPv6 -д ETSI-ийн зөвлөмжийн дагуу бүрэн тохируулсан [18].

Туршилт гүйцэтгэхдээ туннел хийх хувилбарыг оруулаагүй. Учир нь туннелийг хандалтын сүлжээ IPv6-only, сүлжээний цөм IPv4-only болон хандалтын сүлжээ IPv4-only, сүлжээний цөм IPv6-only гэсэн нөхцөлд



Зураг 2 Их сургуулийн кампус сүлжээний топологи (хураангуйлсан)

гүйцэтгэх бөгөөд ISP-тай холбогдож хэмжилт хийхэд нөлөөгүй гэж үзсэн болно.

IV. Үр дүн

Энэхүү ажлын гол асуулт нь IPv4-only, Dual-stack, IPv6-only, хөрвүүлэлт хийж ажиллаж буй гэсэн 4 төрлийн архитектурын хувилбарыг их сургуулийн кампус сүлжээний хувьд гүйцэтгэлээр харьцуулна. Цаашлаад тухайн архитектур бүрт холболт болон үйлчилгээний чанарыг сайжруулахын тулд QoS-ийн тодорхой стратеги, бодлогуудыг хэрхэн боловсруулж, хэрэгжүүлэх талаар төлөвлөгөө гаргах суурь болно. Тодруулбал харьцуулалт хийхдаа RTT хоцрогдол, цэгээс цэгт нэвтрүүлэх чадамж гэсэн үзүүлэлтүүдийг шалгуур болгоно. Ингэснээр сүлжээний гүйцэтгэлийн зөрүү, үйлчилгээний чанарыг тодорхойлох болон дэд бүтцийн хязгаарыг тодорхойлж, IPv6 рүү илүү үр ашигтай шилжихэд шаардлагатай өөрчлөлтийн талаар дүгнэх боломжтой болно.

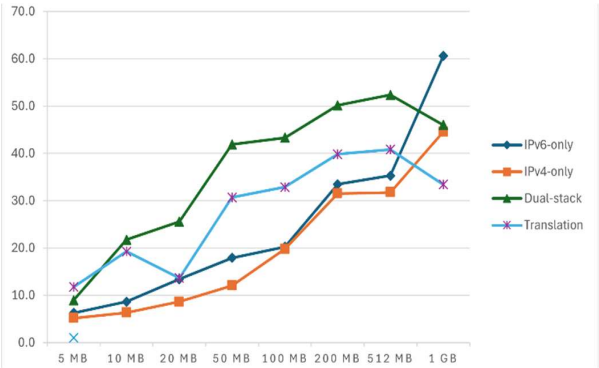
Туршилтын орчинд зориулан ажиллуулсан сүлжээний үйлчилгээнүүд нь IPv4 болон IPv6 тус тусдаа бөгөөд зарим функцийг хоёр хаяглалтын системийг зэрэг дэмждэг байхаар тусгасан. Энэ нь ялангуяа dual-stack тохиргоотой төхөөрөмжийн ачаалал, буфер зэрэгт чухал юм.

A. Нэвтрүүлэх чадамж

Цэгээс цэгт нэвтрүүлэх чадамжийг тодорхойлох зорилгоор дурын хэрэглэгч болон ISP талд байрлуулсан сервер хооронд файл дамжуулалт хийж туршилтаа гүйцэтгэсэн. Файл дамжуулахдаа 5 MB, 10 MB, 20 MB, 50 MB, 100 MB, 200 MB, 512 MB, 1 GB гэсэн хэмжээтэй байхаар өөрчилсөн бөгөөд тус бүрд нь үзүүлэлтийг дундажлаж авч үзсэн болно. Dual-stack -ийг туршихдаа IPv4 болон IPv6 -ийн хэрэглэгчидийг тус тусад зохион байгуулсан.

Зураг 3 -т дундаж нэвтрүүлэх чадамжийг 4 хувилбар тус бүрд харуулав. Зургийн Y тэнхлэгийг Mbps -ийн нэгжээр үзүүлсэн болно. Dual-stack болон хөрвүүлэлтийн гэсэн 2 арга өндөр хурд үзүүлсэн. Гэхдээ хөрвүүлэлтийн аргад төхөөрөмжийн хүчин чадалтай, IPv6 суурьтай үйлчилгээний хэмжээ шууд хамааралтайг анхаарах хэрэгтэй. Мөн өмнө дурьдсанчлан dual-stack үед нэг хэрэглэгчилд тус тусдаа байгаа нь нөлөөлнө. IPv4-only, IPv6-only аргууд ойролцоо үр үзүүлсэн нь нэг хаяглалтын технологи дангаараа ажиллаж байгаатай холбоотой юм.

Энд мөн дурьдах зайлшгүй зүйл бол сүлжээний функцүүд IPv4 дээр ажиллаж нэмэлтээр IPv6 -г дэмждэг байх, эсвэл эсэргээрээ зөвхөн IPv4 -ийг дэмждэг байх нь энэ сүлжээнд ихээхэн нөлөөлдөг.

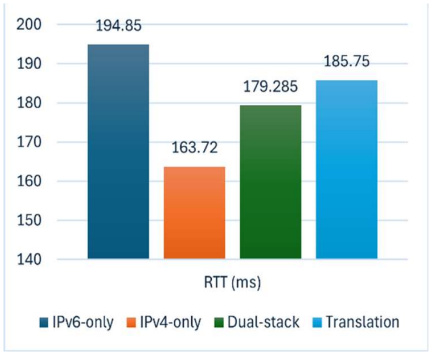


Зураг 3 цэгээс цэгт нэвтрүүлэх чадамж

B. RTT -ийн харьцуулалт

RTT -ээр сүлжээнд пакет эцсийн хүлээн авагч хүртэл явах ба хариу болох АСК -г ирэх хүртлэх хугацааг хэмжиж хоёр талын хоцрогдлыг хэмжсэн. Зураг 4-т нийт туршилтын оролдлогуудын дундажийг 4 хувилбар тус бүрээр үзүүлсэн болно. Зурагт Y тэнхлэгт дундаж хоцрогдол (ms) -ийг, X тэнхлэгт 4 хувилбарыг тус тус үзүүлэв.

Туршилтын үр дүнгээс харахад IPv4-only болон хөрвүүлэлтийн аргууд бусдаас бага зэрэг хоцрогдол бага үзүүлсэн байна. Үүнийг туршилтын сүлжээний онцлог болон сүлжээний функцүүдийн суурь IPv4 байгаатай холбон тайлбарлаж болох юм.



Зураг 4 RTT-ийн үр дүн

Хүснэгт 1 IPv4-only болон IPv6-only бодит сүлжээ дэх харьцуулалт

						IPv6			IPv4		
№	Server domain	IPv6 RTT average (ms)	IPv4 RTT average (ms)	IPv6 hop count	IPv4 hop count	Sender Bitrate (Mbits/s)	Receiver Bitrate (Mbits/s)	Transferred file size (Mbytes)	Sender Bitrate (Mbits/s)	Receiver Bitrate (Mbits/s)	Transferred file size (Mbytes)
1	tracert iperf.angolacables.co.ao	335	296	17	26	41.8	38.6	98.4	*	*	*
2	speedtest.uztelecom.uz	199	166	18	18	102	90.4	229	*	*	*
3	iperf.he.net	203	183	16	20	90.1	87	213	97.8	93.9	232
4	iperf.astra.in.ua	190	129	14	14	119	106	272	144	130	329
5	speedtest.serverius.net	348	197	20	14	37	33.6	86.1	108	104	254
6	paris.bbr.iperf.bytel.fr	184	115	16	17	115	111	812	83.9	79.4	585
7	speedtest.milkywan.fr	248	186	18	21	73	68	508	55.6	53.3	391

C. Бодит сүлжээний орчинд хийсэн харьцуулалт

Энэ хэсэгт симуляуцийн орчинд хийсэн тохиргоог бодит төхөөрөмжүүд бүхий туршилтын орчинд ажлуулсан үр дүнг танилцуулна. Туршилт хийхдээ IPv4-only болон IPv6-only сүлжээг тус тусад нь ажлуулсан болно. Интернэт байрласан серверүүдрүү файл дамжуулах, хүлээн авах үйлдлийг гүйцэтгэж хэмжсэн үр дүнг Хүснэгт 1-д үзүүлэв. Илгээх болон татах үйлдлийг зэрэг буюу хос чиглэлтэй урсгал үүсгэж буй учир илгээгч болон хүлээн авагч талд бит хурд өөр өөр гарч байгаа болно.

Үр дүнгээс харвал IPv6 -ийн hop -ийн тоо бага боловч дундаж RTT -ийн хэмжээ их байна. Энэ нь IP датаграмм хэмжээ болон холболтын төрлөөс хамаарч байгаа үзүүлэлт юм. Мөн эдгээр үзүүлэлт нь IPv4 датаграммын дамжууллын үеийн хуваалт хийгдэж байгаа эсэхээс хамаарна. Энд dual-stack болон хөрвүүлэлт хийж буй хоёр хувилбарыг оруулаагүй нь IPv4 болон IPv6 -уудыг хослуулан ашиглаж байгаа хувилбар учир төгсгөлөөс төгсгөл хүртлэх файл дамжуулалтын үр дүнг харуулах хүндрэлтэй юм.

V. Дүгнэлт

Их сургуулийн кампус сүлжээ нь хэрэглэгчийн нягтрал өндөр, траффикийн хэв шинж олон янз байдаг академик орчин учир IPv6-ийн гүйцэтгэлийг шалгах хамгийн чухал туршилтын талбар юм. Их сургуулийн сүлжээний дэд бүтцийг IPv6 ашиглан сайжруулахад тулгарч буй сорилтуудын нэг нь одоогоор давамгай байгаа IPv4-д суурилсан системүүд юм. Эдгээрэтй бүрэн нийцтэй, саадгүй интеграцчиллыг хангах шилжилтийн үеийн гол зорилго болно. Шилжилтийн хамгийн эхэнд авч хэрэгжүүлж буй dual-stack нь нийцтэй байдлыг хангах, системүүдийг шинэчлэх дундын хувилбар бөгөөд гүйцэтгэл хангалттай байна. Харин дараагийн хувилбар руу хэзээ шилжих нь системүүд IPv6 -г дэмждэг болохоос хамаарах бол сүлжээний дэд бүтэц талаас хамаарал бага байна. Хөрвүүлэлт хийх арга нь мэдээж одоогоор гүйцэтгэл хангалтгүй байгаа боловч төхөөрөмжийн хүчин чадал, системийн боломж нэмэгдэх тусам сайжрах болно.

IPv6 рүү шилжих нь мөн сургалт, техник хангамжийн шинэчлэлт, аюулгүй байдлын тохиргоо ихээхэн хөрөнгө оруулалт шаардана. Учир нь сүлжээний администраторууд шинэ хаяглалтын схем, чиглүүлэлтийн протоколууд болон хяналт-шинжилгээний системүүдийг нэвтрүүлэх шаардлагатай болдог.

Acknowledgment

This work was supported by the National University of Mongolia, grant number P2025-5010 and APNIC Information Society Innovation Fund (ISIF Asia) M-202503-03855.

МУИС P2025-5010, APNIC M-202503-03855 сангийн санхүүжилттэй төслийн хүрээнд хийж гүйцэтгэв.

Ашигласан материал

[1] Igulu, Kingsley, Friday Onuodu, and Thipendra P. Singh. "IPV6: Strengths and limitations." Communication Technologies and Security Challenges in IoT: Present and Future. Singapore: Springer Nature Singapore, 2024. 147-172.

[2] Hamarsheh, Ala. "Assessing the progress of transition to ipv6: a global perspective." IETE Journal of Research 71.8 (2025): 2578-2592.

[3] Zhang, Yanan, Yanfang Fu, and Qian Wang. "IPv4 to IPv6 Transition Strategy Based on Dual Stack Protocol." Intelligent Computing Technology and Automation (2024): 428-435.

[4] Li, Kwun-Hung, and Kin-Yeung Wong. "Empirical analysis of IPv4 and IPv6 networks through dual-stack sites." Information 12.6 (2021): 246.

[5] Hamarsheh, Ala. "Evaluating the Advantages and Disadvantages of Transition Technologies for IPv6 in the Context of IPv4-as-a-Service." IETE Journal of Research 70.11 (2024): 8031-8043.

[6] Islam, Shayla, et al. "Mobile networks toward 5G/6G: Network architecture, opportunities and challenges in smart city." IEEE Open Journal of the Communications Society 6 (2024): 3082-3093.

[7] Al-Shareeda, Mahmood A., et al. "Ipv6 link-local network slaac attack detection mechanisms: a review." 2022 Fifth College of Science International Conference of Recent Trends in Information Technology (CSCTIT). IEEE, 2022.

[8] Piraux, Maxime, et al. "The multiple roles that IPv6 addresses can play in today's internet." ACM SIGCOMM Computer Communication Review 52.3 (2022): 10-18.

[9] Jing, Shan, et al. "Research and deployment of IPv4/IPv6 dual stack network in large-scale campus network." 2021 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS). IEEE, 2021.

[10] Praptodiyono, Supriyanto, et al. "Mobile IPv6 vertical handover specifications, threats, and mitigation methods: A survey." Security and Communication Networks 2020.1 (2020): 5429630.

[11] Jain, Neha, and Ashish Payal. "Performance comparison between different tunneling techniques using different routing protocols." Wireless Personal Communications 123.2 (2022): 1395-1441.

[12] Al-Azzawi, Ameen, and Gábor Lencse. "Analysis of the security challenges facing the DS-Lite IPv6 transition technology." Electronics 12.10 (2023): 2335.

[13] Hamarsheh, Ala. "Evaluating the Advantages and Disadvantages of Transition Technologies for IPv6 in the Context of IPv4-as-a-Service." IETE Journal of Research 70.11 (2024): 8031-8043.

[14] Hsu, Amanda, et al. "A first look at NAT64 deployment in-the-wild." International Conference on Passive and Active Network Measurement. Cham: Springer Nature Switzerland, 2024.

[15] Streibelt, Florian, et al. "How ready is DNS for an IPv6-only world?." International Conference on Passive and Active Network Measurement. Cham: Springer Nature Switzerland, 2023.

[16] Robinson, Kate, Jason Zurawski, and Tom Costello. "Designing, Constructing, and Operating an IPv6 Network at SC23: A case study in implementing the IPv6 protocol on a heterogenous network that supports the SC23 conference." Practice and Experience in Advanced Research Computing 2024: Human Powered Computing. 2024. 1-8.

[17] Hamarsheh, Ala. "Evaluating the Advantages and Disadvantages of Transition Technologies for IPv6 in the Context of IPv4-as-a-Service." IETE Journal of Research 70.11 (2024): 8031-8043.

[18] ETSI. (2020). IPv6 Best Practices, Benefits, Transition Challenges and the Way Forward. Antipolis: ETSI. Retrieved January 28, 2023, from

Чөлөөт орон зайд ОӨМ бүхий дохиог дамжуулахад турбуленсийн нөлөөлөх нөлөөллийн судалгаа

Г. Намуундарь
Холбооны иженерчлэлийн тэнхим
ШУТИС-ийн харьяа Мэдээлэл холбооны технологийн сургууль
Монгол улс, Улаанбаатар хот
gannorovnamuundari@gmail.com

А. Мөнхбаяр
Холбооны иженерчлэлийн тэнхим
ШУТИС-ийн харьяа Мэдээлэл холбооны технологийн сургууль
Монгол улс, Улаанбаатар хот
a_munkhbayar@must.edu.mn

Хураангуй*— Чөлөөт орон зайн оптик холбоо нь лазер туяаг ашиглан агаар мандлаар өндөр хурдтай өгөгдөл дамжуулах боломжтой, өргөн зурвасын багтаамжтай, дэвшилтэт технологи бөгөөд орчин үеийн холбооны сүлжээний чадавхыг нэмэгдүүлэх чухал шийдэл болж байна. Гэвч атмосферийн турбуленс нь оптик долгионд нөлөөлж, дохионы эрчимшлийн хэлбэлзэл, фазын гажилт үүсгэснээр дамжууллын чанарыг бууруулах үндсэн саад болдог. Энэхүү судалгаанд сувгийн багтаамжийг нэмэгдүүлэх зорилгоор орбитал өнцгийн моментум (ООМ) бүхий Лагерр-Гауссын дохиог ашиглан чөлөөт орон зайд дамжуулах үед турбуленсийн нөлөөг симуляцийн аргаар судалсан. Судалгааны хүрээнд апертур диаметрын оновчтой утгыг тодорхойлох, турбуленс болон дамжуулах зайнаас хамаарсан тохиромжтой моодын тоог тодорхойлох, мөн өндөр моодтой дохионы тархалтын онцлогийг шинжилсэн. Үр дүнгээс харахад дундаж турбуленсийн нөхцөлд 3 mm апертур диаметер хамгийн тохиромжтой бөгөөд хүлээн авах талд бага моодтой дохионууд илүү өндөр магадлалтай танигдаж байна. Түүнчлэн турбуленсийн хэмжээ болон дамжуулах зай ихсэх тусам дохионы танигдах магадлал буурч байгааг тогтоосон. 500 м зайд 01–04 моодыг ашиглах нь оновчтой бөгөөд сувгийн багтаамжийг 8 дахин нэмэгдүүлэх боломжтой бол богино зайд илүү олон моод ашиглах боломжтой. Харин өндөр моодтой дохионууд их турбуленстэй орчинд хөрш моодууд руу шилжих хандлагатай байгааг туршилтаар дүгнэсэн.

Түлхүүр үг— турбуленс, сувгийн багтаамж, Лагерр-Гаусс , моод

I. Удиртгал

Сүүлийн жилүүдэд мэдээлэл, харилцаа холбооны технологийн хөгжил эрчимжиж, өндөр хурдтай, өргөн зурвасын багтаамжтай өгөгдөл дамжуулах сүлжээний хэрэгцээ эрс нэмэгдэж байна. Интернэт хэрэглээ, үүлэн тооцоолол, дата төвүүдийн хоорондын холболт, сансрын холбооны систем зэрэг орчин үеийн дэд бүтцүүд нь их хэмжээний мэдээллийг богино хугацаанд дамжуулах шаардлагыг бий болгож байгаа бөгөөд энэ нь уламжлалт холбооны технологийн хүчин чадал, боломжийг дахин үнэлэхэд хүргэж байна. Ялангуяа радио давтамжид суурилсан утасгүй холбооны системүүдийн спектрийн нөөц хязгаарлагдмал, интерференцийн асуудал ихсэх зэрэг шалтгаанууд нь илүү өндөр давтамжийн мужид ажиллах шинэ төрлийн дамжууллын технологийг судлах шаардлагыг бий болгож байна.

Ийм нөхцөлд чөлөөт орон зайн оптик холбоо нь өндөр хурдтай өгөгдөл дамжуулах, өргөн зурвасын багтаамжийг хангах, кабельгүй холболтыг хэрэгжүүлэх боломжтой дэвшилтэт технологийн нэг болж байна. Энэхүү технологи нь лазер эсвэл гэрлийн долгионыг ашиглан мэдээллийг агаар мандлаар дамжуулдаг бөгөөд оптик холбооны үндсэн зарчимд тулгуурласан утасгүй дамжууллын систем юм. Чөлөөт орон зайн оптик холбоо нь хот доторх өндөр хурдны цэгээс цэг холболт, дата төвүүдийн хоорондын өгөгдөл дамжуулалт, сансрын холбоо зэрэг олон салбарт өргөн хэрэглэгдэх боломжтой. Гэсэн хэдий ч энэхүү технологийн гүйцэтгэлд атмосферийн орчны нөлөөлөл нь дамжууллын чанарыг бууруулах гол хүчин зүйл болоод байна. Ялангуяа агаарын турбулент байдал нь оптик долгионы тархалтад нөлөөлж, дохионы сулрал, фазын гажилт, эрчмийн хэлбэлзлийг үүсгэдэг. Энэ нь чөлөөт орон зайн оптик холбооны системийн найдвартай ажиллагаанд сөргөөр нөлөөлөх хүчин зүйл болдог.

Нөгөө талаас, оптик холбооны системийн дамжуулах чадварыг нэмэгдүүлэх зорилгоор гэрлийн долгионы орбитал өнцөг импульсийн шинж чанарыг ашиглах судалгаа сүүлийн жилүүдэд эрчимтэй хөгжиж байна. Орбитал өнцөг импульс бүхий лазерын туяанууд нь фазын мушгиралттай бүтэцтэй бөгөөд хоорондоо ортогональ горимуудыг ашиглан нэг физик сувгаар олон өгөгдлийн урсгалыг зэрэг дамжуулах боломжийг бүрдүүлдэг. Ийм шинж чанар нь оптик холбооны системийн сувгийн багтаамжийг мэдэгдэхүйц нэмэгдүүлэх боломжийг олгодог бөгөөд ирээдүйн өндөр багтаамжтай холбооны технологийн чухал судалгааны чиглэл болж байна. Иймээс чөлөөт орон зайн оптик холбооны системд атмосферийн турбулент орчны нөлөөллийг онолын болон загварчлалын түвшинд судлах нь чухал ач холбогдолтой юм.

II. Чөлөөт орон зайн оптик дамжуулал

Чөлөөт орон зайн оптик холбоо нь харилцаа холбооны сүлжээнд өгөгдлийг агаараар буюу чөлөөт орон зайгаар лазер эсвэл гэрлийн долгион ашиглан дамжуулах утасгүй технологи юм. Энэхүү технологи нь шилэн кабелийн холбооны зарчимд тулгуурладаг боловч дамжуулах орчин нь шилэн кабель бус агаар мандал байдаг. Уг систем нь өндөр зурвасын өргөн, лиценз шаарддаггүй спектр ашиглах давуу талтай бөгөөд богино болон дунд зайн өндөр хурдны холболтод өргөн хэрэглэгддэг.

Оптик холбооны системтэй дүйцэхүйц өндөр гүйцэтгэлтэйгээс гадна шилэн кабель татах шаардлага гарахгүй гэх зэрэг онцлог талуудтай.

A. Турбуленсийн онол

Агаар мандлын турбулент байдал гэдэг нь салхины хурд, чиглэл тогтмол бус өөрчлөгдөж байдаг, жижиг хэмжээний эмх замбараагүй агаарын хөдөлгөөнүүдийг хэлнэ. Турбуленс нь агаар мандлыг холих, хутгах үүрэгтэй тул усны уур, утаа болон бусад бодисууд, мөн энерги нь босоо болон хэвтээ чиглэлд тархан жигд хуваарилагдахад чухал ач холбогдолтой [1].

Агаарын турбулент байдлыг хэмжихдээ C_n^2 параметрыг авч үздэг [2].

$$C_n^2 = b\frac{K_H}{\varepsilon^{\frac{1}{3}}}\left(\frac{dn}{dz}\right)^2 \mathrm{m}^{-2/3} \qquad (1)$$

Энд:
$b = 3.2$ - Обухов–Коррсиний тогтмол
K_H - дулааны диффузийн турбулент солилцооны коэффициент
$\boldsymbol{\varepsilon}$ **-** энергийн сарнилтын хурд
$\frac{dn}{dz}$ - хугарлын илтгэгч n-ийн босоо чиглэл дэх градиент

Турбуленсын параметр $\mathbf{C_n^2}$-ийн утга нь $1\cdot10^{-12}$ – $1\cdot10^{-16}$ $\mathrm{m}^{-2/3}$ орчим мужид хэлбэлздэг.

- $\mathbf{C_n}^2 = 1\cdot10^{-12} \sim 1\cdot10^{-15}\mathrm{m}^{-2/3}$ хүчтэй,
- $\mathbf{C_n}^2 = 1\cdot10^{-16} \sim 1\cdot10^{-15}\mathrm{m}^{-2/3}$ сул турбуленстэй гэж үзнэ.

B. Колмогоровын турбулентийн загвар

Колмогоров–Татарскийн загвар нь агаар мандлын турбулент байдлыг долгионы фронтын фазын гажилтаар тодорхойлж, одон орны дүрслэлийн чанарыг ойлгох, загварчлахад суурь онол болдог [3].
Энэхүү загварын дагуу агаар мандлын хугарлын илтгэгчийн хэлбэлзэл нь долгионы фронтыг гажуудуулдаг ба энэ нь шууд фазын хэлбэлзэл $\phi_a(\mathbf{r})$ үүсгэнэ. Харин амплитудын хэлбэлзэл $\chi_a(\mathbf{r})$ нь зөвхөн хоёрдогч нөлөөгөөр, өөрөөр хэлбэл долгион атмосфер дундуур тархах явцад үүсдэг.
Оптик болон инфра-улаан мужид:

- Фазын хэлбэлзэл $\phi_a(\mathrm{r})$ нь дүрсний чанарт голлон нөлөөлнө
- Амплитудын хэлбэлзэл $\chi_a(\mathrm{r})$ нь ихэвчлэн үл тоомсорлож болохуйц бага нөлөөтэй гэсэн дүнэлтэд хүрнэ.

III. Орбитал өнцгийн моментот (ООМ) дохио

Уг судалгааны ажлын гол ойлголт болох орбитал өнцгийн моментум (ООМ) нь гэрлийн долгионы фазын мушгиралттай бүтэцтэй холбоотой физик шинж чанартай дохио юм. Энгийн лазер туяа шулуун фазтай байдаг бол ООМ агуулсан долгионы фаз нь мушгиа (helical wavefront) хэлбэртэй тархана. Энэ нь тухайн фотон бүр тодорхой хэмжээний өнцгийн импульс агуулна гэсэн ойлголт юм [4].

ООМ-тай долгионы цахилгаан орны фаз дараах хэлбэртэй илэрхийлэгдэнэ:

$$E(r,\phi,z) = A(r,z)e^{il\phi}\,\frac{V}{m} \qquad (2)$$

Энд:
l- мушгиралтын зэрэг
ϕ- өнцөг

Фотон бүрийн орбитал өнцгийн момент нь :

$$L = l\hbar \;\; J\cdot s \qquad (3)$$

Энд:
$\hbar = 6.62607015 \times 10^{-34}\ J\cdot s$ - планкийн тогтмол

Чөлөөт орон зайн оптик холбоонд уламжлалт лазер дохио нэг сувгаар нэг өгөгдлийн урсгал дамжуулдаг. Тэгвэл орбитал өнцгийн моментум бүхий дохио нь:

- Мood бүр ортогональ буюу мood бүр нь хоорондоо хөндлөн харилцан нөлөөгүй, нэг моодны дохио нөгөө моодын дохионд нөлөөлдөггүй.
- Нэг физик сувгаар олон өгөгдлийн урсгал дамжуулах боломжтой.
- тэдгээрийг хооронд нь саадгүйгээр зэрэг ашиглах боломжтой.

Ортогональ шинжийг нотлох нь [5]:
Хоёр өөр ООМ мood ℓба ℓ'-ийн скаляр үржвэрийг авч үзье:

$$\langle\psi_\ell,\psi_{\ell'}\rangle = \int_0^{2\pi}\int_0^{\infty}\psi_\ell\,(r,\phi)\,\psi_{\ell'}^{*}(r,\phi)\,r\,dr\,d\phi \qquad (3)$$

Орлуулбал:

$$= \int_0^{2\pi}\int_0^{\infty} A(r)\,e^{i\ell\phi}\cdot A^{*}(r)e^{-i\ell'\phi}\,r\,dr\,d\phi$$

$$= \int_0^{\infty} |\,A(r)\,|^2\,r\,dr \;\cdot\; \int_0^{2\pi} e^{i(\ell-\ell')\phi}\,d\phi$$

$$\int_0^{2\pi} e^{i(\ell-\ell')\phi}\,d\phi = 2\pi\,\delta_{\ell\ell'} \qquad (4)$$

Энд $\delta_{\ell\ell'}$ нь Kronecker delta:

- $\ell = \ell'$үед → инте грал = 2π
- $\ell \neq \ell'$үед → интеграл = 0

Иймээс:

$$\langle\psi_\ell,\psi_{\ell'}\rangle \propto \delta_{\ell\ell'} \qquad (5)$$

Өөрөөр хэлбэл:

- Өөр өөр ℓ -тэй ООМ моодууд хоорондоо ортогональ

- Энэ нь тэдгээрийг хооронд нь ялгаж, давхар ашиглах боломжийг олгодог

Үүнийг ашиглахад сувгийн багтаамж хязгааргүй мэт нэмэгдэх буюу ОӨМ нь сувгийн багтаамжийг нэмэгдүүлэх физик механизм юм.

A. ОӨМ дохионы төлүүд

ОӨМ дохионы төрөл бүр нь өөр өөр физик болон математик шинж чанартай бөгөөд тухайн хэрэглээнээс хамааран сонгогддог. Лагерр-Гауссын дохио нь суурь загвар болж, харин Bessel, Airy зэрэг beam-үүд нь тархалтын өвөрмөц шинж чанараараа давуу талтай [5].

Лагерр-Гауссын дохионы математик илэрхийлэл [7]:

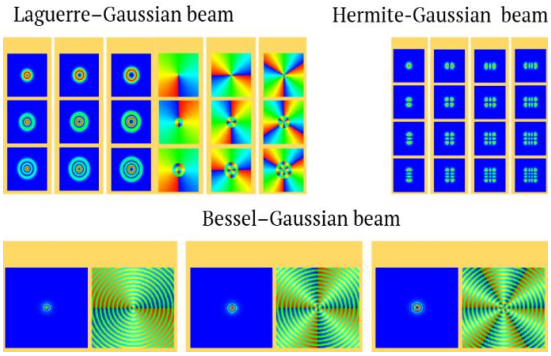
$$U_{LG} = (\rho, \phi, z) = R_{\rho m}(r, z)\exp(-im\phi)\frac{V}{m} \quad (6)$$

Hermite-Gaussian дохионы математик илэрхийлэл [7]:

$$U_{HG} = H_p(r, \phi) = H_p\left(\frac{\sqrt{2}}{w_{0x}}\right)\exp\left(\frac{x^2}{w_{0x}^2}\right)\exp\left(j\frac{\pi x^2}{\lambda R_{0x}}\right) \times H_m\left(\frac{\sqrt{2}}{w_{0y}}\right)\exp\left(\frac{y^2}{w_{0y}^2}\right)\exp\left(j\frac{\pi y^2}{\lambda R_{0y}}\right)\frac{V}{m} \quad (7)$$

Bessel-Gaussian дохионы математик илэрхийлэл [7] :

$$U_{BG}(\rho, \phi) = J_m(k_r\rho)\exp(im\phi)\exp\left(-\frac{\rho^2}{w_g^2}\right)\frac{V}{m} \quad (8)$$



Зураг 1. ОӨМ дохионы төрлүүд [7]

Зураг 1-т ОӨМ дохионы түгээмэл 3 төрөл болон тэдгээрийн эрчимшил, фазыг харууллаа.

IV. Үйцэтгэсэн ажил

Уг судалгааны ажил нь дараах 3 зорилгын хүрээнд туршиж, үр дүнд суурилан дүгнэх зорилготой.

1. Апертур диаметрыг оновчтой тодорхойлох.
2. Өндөр моодтой дохиог их турбуленстэй орчноор дамжуулах үед хүлээн авах талд хэрхэн таньж байгааг судлах.
3. Дараах тохиолдлуудад оновчтой моодын тоог тодорхойлох:

- Тогтмол зайд турбуленсээс хамааруулан.
- Тогтмол турбуленстэй үед зайнаас хамааруулан.

Ингэхдээ OptiSystem програмыг Matlab програмын хамт ашигласан болно.

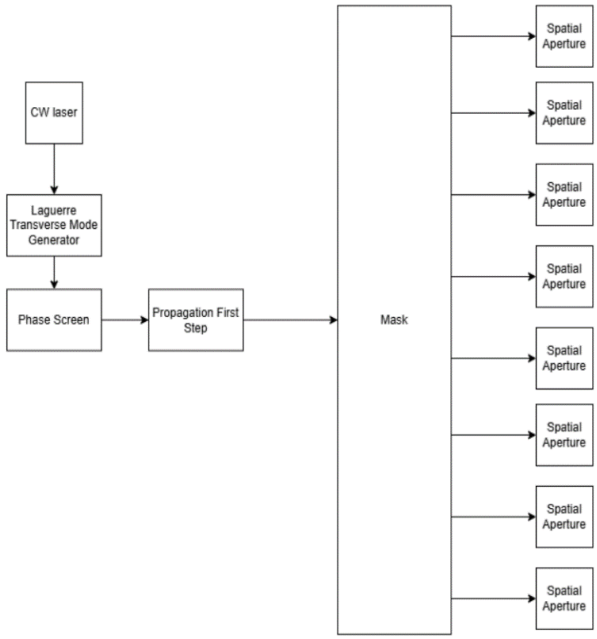
Симульяцын параметрүүд:

- Долгионы урт 850nm
- Давталтын тоо 10
- нэг фазын дэлгэц (phase screen)

SLM (Spatial Light Modulator) ашиглан phase screen-г үүсгэсэн.

Турбуленсийн параметер болон зайн мэдээллийг OptiSystem програмын глобал параметер дээр тусгасан болно.

Симульяын блок диаграмыг зураг 2-т харуулав.



Зураг 2. Блок диаграм

Зураг 2-т харуулснаар:

CW laser - Эхлэх долгионы фронт үүсгэнэ.

Laguerre Transverse Mode Generator - Лазерын гэрлийг Лагерр-Гауссын горимд хөрвүүлнэ.

Phase Screen - Хугарлын илтгэгчийн хэлбэлзэл C_n^2 -д тулгуурлан санамсаргүй фазын давхарга үүсгэнэ. SLM (Spatial Light Modulator) ашигласан.

Propagation First Step - Долгионы фронтыг фазын дэлгэцээр дамжуулж, цааш нь propagation operator ашиглан зайд тархана.

Mask – Өрнүүлсэн Лагерр-Гауссын дохиог эсрэг фазтай хологрaмаар үржүүлнэ.

Spatial Aperture – Дохиог анализлана.

V. Үр дүн

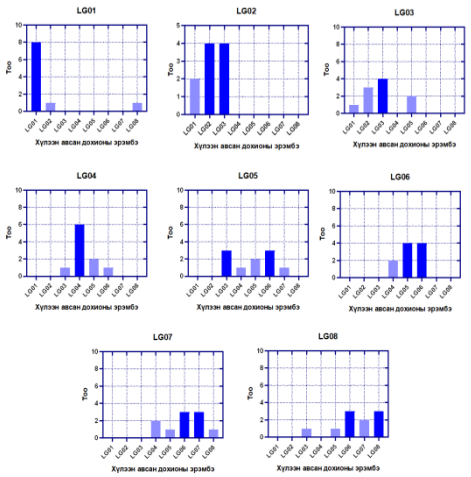
A. Апертур диаметрыг оновчтой тодорохойлох

$C = 1 \times 10^{-14}\mathrm{m}^{-2/3}$ турбуленсийн параметертэй буюу дундаж турбуленстэй үед апертур диаметрыг 2mm, 3mm, 4mm дээр тус бүр туршсан.

Турбуленсийн параметер болон зайн утгыг OptiSystem програмын глобал параметер дээр тусгасан болно.

1. $2000\mu m = 2mm$ *үед*

Зураг 3-т апертур диаметер 2mm үеийн хүлээн авах талд дохио хэрхэн танигдсан болохыг үзүүлэв.



Зураг 3. 2mm үед

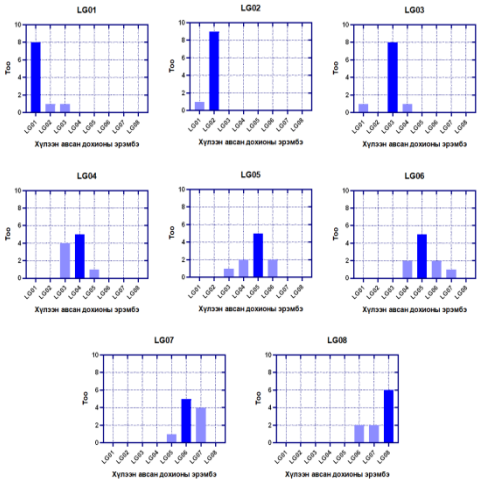
- Энэ үед 01 дамжуулахад хүлээн авах талд зөв таних магадлал 01→01 (80%)
- 02 дамжуулбал хүлээн авах талд 02 гэж зөв таних магадлал 40% гэх мэтчилэн:

03→03(40%) 04→04(60%)
05→05(20%) 06→06(40%)
07→07(30%) 08→08(30%)

Үүнийг хангалттүй үр дүн гэж үзнэ.

2. $3000\mu m = 3mm$ *үед*

Зураг 4-д апертур диаметер 3mm үеийн хүлээн авах талд дохио хэрхэн танигдсан болохыг үзүүлэв.



Зураг 4. 3mm үед

Зураг 4-с үзвэл

- 01 дамжуулахад хүлээн авах талд зөв таних магадлал 01→01 (80%)
- 02 дамжуулбал хүлээн авах талд 02 гэж зөв таних магадлал 90% гэх мэт:

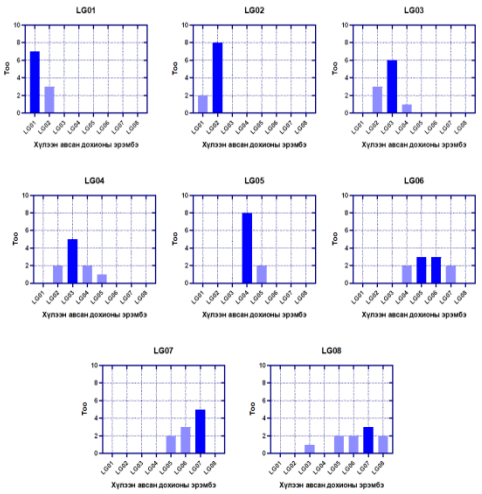
03→03(80%) 04→04(50%)
05→05(50%) 06→06(50%)
07→07(50%) 08→08(60%)

01, 02, 03 дохиог дамжуулах үед хүлээн авах талд амжилттай таньж чадаж байгаа учир 1,2,3 моодын дохиог хэрэглэхийг зөвлөж байна. Дан ганц 1,2,3 моодыг ашиглахын зэрэгцээ тэдгээрийн эсрэг фазын -1,-2,-3 моодуудыг хэрэглэснээр сувгийн багтаамжийг 6 дахин нэмэгдүүлнэ.

l=0,1 l =0,2 l =0,3
l =0,-1 l =0,-2 l =0,-3

3. $4000\mu m = 4mm$ *үед*

Зураг 5-д апертур диаметер 4mm үед дохио хүлээн авах талд хэрхэн танигдсан болохыг графикаар үзүүлэв.



Зураг 5. 4mm үед

- Энэ үед 01 дамжуулахад хүлээн авах талд зөв таних магадлал 01→01 (70%)
- 02 дамжуулбал хүлээн авах талд 02 гэж зөв таних магадлал 90% гэх мэтчилэн:

03→03(60%) 04→04(50%)
05→05(80%) 06→06(30%)
07→07(50%) 08→08(30%)

Үүнийг хангалттай үр дүн гэж үзэхэд дутагдалтай байна.

Дээрх 3 тохиолдлуудыг харьцуулахад 3mm-н апертур диаметер хамгийн тохиромжтой гэдгийг тодорхойллоо.

B. Тогтмол зайд турбулесийн хэмжээнээс хамааруулан тохиромжтой моодыг тодорхойлох

Өгөгдөл:

$l = const = 500m$

$C = 1 \times 10^{-15}\ \mathrm{m}^{-2/3}$

$C = 4 \times 10^{-15}\ \mathrm{m}^{-2/3}$

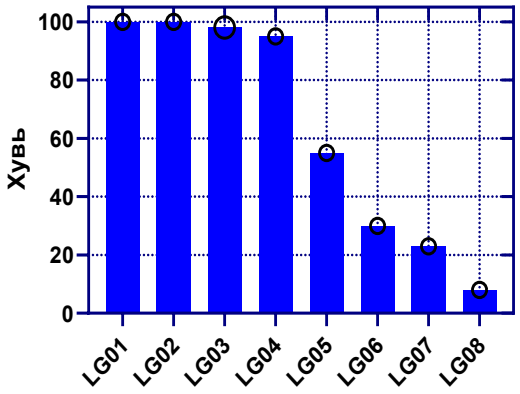
$C = 7 \times 10^{-15}\mathrm{m}^{-2/3}$

$C = 10 \times 10^{-15}\mathrm{m}^{-2/3}$

Хүснэгт 1 ба зураг 6-д тогтмол 500m-н зайтай, турбуленсийн хэмжээ өгөгдсөн утгуудын үед хүлээн авах талд дохиог хэдэн хувийн гүйцэтгэлтэй таньж байгааг харууллаа.

Хүснэгт 1. ЗАЙ ТУРБУЛЕНСЭЭС ХАМААРСАН ХАМААРЛЫН ХҮСНЭГТ

C/m	LG01	LG02	LG03	LG04	LG05	LG06	LG07	LG08
1	100%	100%	100%	100%	10%	0%	0%	0%
4	100%	100%	100%	100%	80%	30%	30%	0%
7	100%	100%	100%	100%	70%	30%	30%	0%
10	100%	100%	90%	80%	60%	60%	30%	30%



Зураг 6. Зай турбуленсээс хамаарсан хамаарлын график

Графикаас үзвэл:

- турбуленс ихсэх тусам хүлээн авах талд дохио амжилттай танигдах магадлал багасч байна.
- 0-с 8 хүртэлх моодын дохионууд нь хамгийн ихдээ хэд хүртэлх турбуленстэй орчинд амжилттай дамжих вэ гэдгийг харж болно:

01, 02, 03, 04 дохионууд $C = 1 \times 10^{-14}\mathrm{m}^{-2/3}$ хүртэл турбуленстэй орчинд 100 хувийн гүйцэтгэлтэй дамжиж байгаагаас дүгнэвэл:

- 500m –н дамжууллын тохиолдолд дээрх 4 моодыг ашиглахыг зөвлөж байна. Ингэснээр сувгийн багтаамжийг 8 дахин нэмэгдүүлнэ.

C. Тогтмол турбуленсийн утгатай үед зайнаас хамааруулан тохиромжтой моодыг тодорхойлох

Өгөгдөл:

$C = 1 \times 10^{-14}\mathrm{m}^{-2/3}$

$l = 250m$

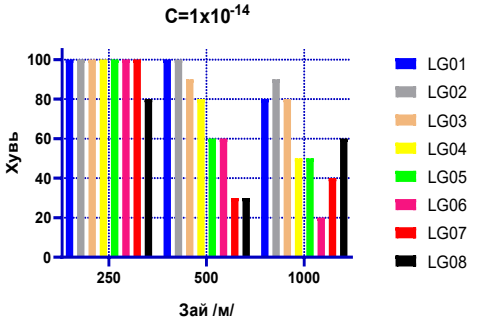
$l = 500m$

$l = 1000m$

2-р хүснэгт болон 7-р зураг турбуленсийн хэмжээ тогтмол $C = 1 \times 10^{-14}\mathrm{m}^{-2/3}$ үед 250m, 500m, 1000m-н зайд дохиог дамжуулахад хүлээн авах тал хэдэн хувийн гүйцэтгэлтэй таньж байгааг харууллаа.

Хүснэгт 2. ТУРБУЛЕНС ЗАЙНААС ХАМААРСАН ХАМААРЛЫН ХҮСНЭГТ

l/m	LG01	LG02	LG03	LG04	LG05	LG06	LG07	LG08
250	100%	100%	100%	100%	100%	100%	100%	80%
500	100%	100%	90%	80%	60%	60%	30%	30%
1000	80%	90%	80%	50%	50%	20%	40%	60%



Зураг 7. Турбуленс зайнаас хамаарсан хамаарлын график

Графикаас үзэхэд:

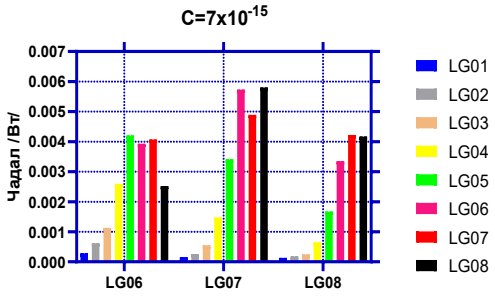
- Зай нь дохионы моодоос урвуу хамааралтай буюу зай ихсэх тусам хүлээн авах талд амжилттай танигдах магадлал буурч байна.
- 250m үед 0-с 8-р мод
- 500m үед 0-с 4-р мод
- 1000m үед 0-с 3-р моодын дохиог ашиглахыг зөвлөж байна. Мөн эдгээрийн эсрэг фазын дохиог ашиглах боломжтой ба ингэснээр сувгийн багтаамжийг 16, 8, 6 дахин нэмэгдүүлнэ.

D. Өндөр моодын дохиог их турбуленстэй орчноор дамжуулах үед хүлээн авах талд хэрхэн таньж байгааг судлах

Өгөгдөл:

$C = 7 \times 10^{-15}\mathrm{m}^{-2/3}$, $l = 500m$

Зураг 8-д 06, 07, 08 моодын дохиог дамжуулах үед хүлээн авах талд хэрхэн танигдаж байгааг харууллаа.



Зураг 8. Хүлээн авсан дохионы чадлын график

Графикаас үзэхэд өндөр моодын дохиог өндөр турбуленстэй орчноор дамжуулах үед дамжуулсан моодын дохио хүлээн авах талд 2 хажуугийн моодруугаа шилжиж байгаа буюу 2 хажуугийн дохиотой эндүүрэгдэж танигдах магадлал өндөр гэдгийг туршилтаар үнэллээ.

VI. Дүгнэлт

1. Турбуленс ихсэх тусам хүлээн авах талд дохио амжилттай танигдах магадлал багасч ба йна.
2. Зай нь ОӨМ дохионы моодоос урвуу хамааралтай буюу зай ихсэх тусам хүлээн авах талд амжилттай танигдах магадлал буурч байна.
3. Дунд хэмжээний турбуленстэй ($C = 1 \times 10^{-14} \mathrm{m}^{-2/3}$), 500m –н зайн дамжууллын үед 01, 02, 03, 04 моодуудыг ашиглахыг зөвлөж байна. Ингэснээр сувгийн багтаамжийг 8 дахин нэмэгдүүлнэ.
4. $C = 1 \times 10^{-14} \mathrm{m}^{-2/3}$ турбуленстэй үед
   - 250m-д 01-с 08-р моод
   - 500m-д 01-с 04-р моод
   - 1000m-д 01-с 03-р моодын дохио хэрэглэхийг зөвлөж байна. Мөн эдгээрийн эсрэг фазын дохиог ашиглах боломжтой ба ингэснээр сувгийн багтаамжийг 16, 8, 6 дахин нэмэгдүүлнэ.
5. Өндөр моодын дохиог өндөр турбуленстэй орчноор дамжуулах үед дамжуулсан моодын дохио хүлээн авах талд 2 хажуугийн моодруугаа шилжиж байгаа буюу 2 хажуугийн дохиотой эндүүрэгдэж танигдах магадлал өндөр.
6. 3mm-н апертур диаметер хамгийн тохиромжтой.
7. Дээрх судалгаа, дүгнэлтүүд цэвэр симульяцид суурилсан бөгөөд бодит туршилтад хэрэгжүүлэхэд тулгарч болох зөрүү нь турбуленсын бодит динамик, төхөөрөмжүүдийн идеаль шинж чанар, системийн элементүүдийн нарийвчлалын алдаа зэргээс үүдэлтэй байх боломжтой.

Ном зүй

[1] 1 April 2026. [Online]. Available: https://www.britannica.com/science/atmospheric-turbulence.

[2] A. Tunick, The Refractive Index Structure Parameter/Atmospheric Optical Turbulence Model: CN2, vol. 01, p. april.

[3] N. LAMBDA, "The Kolmogorov model of turbulence," [Online]. Available: https://lambda.gsfc.nasa.gov/product/websites/AMI/mrao.cam.ac.uk/telescopes/coast/theses/rnt/node5.html.

[4] M. W. B. R. J. C. S. J. P. W. L. Allen, Orbital angular momentum of light and the transformation of Laguerre–Gaussian laser modes.

[5] M. Beresna, Polarization engineering with ultrafast laser writing in transparent media, University of Southampton Research Repository, August 2012.

[6] M. B. David L. Andrews, The Angular Momentum of Light, Cambridge University Press.

[7] H. K. N. G. G. S. Munkhbayar Adiya, "8-ary orbital angular momentum shift keying for free-space optical communication system".

САЛБАР ХУРАЛДААН - III

Эх хэл боловсруулалт

№	**ҮБХИС-ийн хичээлийн I байр, 201 тоот, Эрдэм танхим** *Хуралдаан даргалагч: Ч.Алтангэрэл /МУИС/, А.Алтангэрэл /ШУТИС/*
1	“Бага нөөцтэй нутгийн аялгын ярианы синтезийн судалгаа” *Золзаяа Бямбадорж, Өлзийбаяр Соном-Очир, Батгэрэл Ариунзаяа*
2	“Латин галиглалтаас Кирилл рүү: Нөөц багатай хэлэнд хэлний том загвар ашигласан шийдэл” *Өлзийбаяр Соном-Очир, Бямбадорж Золзаяа, Батгэрэл Ариунзаяа*
3	“Munkhnote: A Modular AI-Assisted Meeting Minutes System for Mongolians with Improved Workflow Efficiency” *Magvan-Erdene Gantulga, Munkhtuya Uurtsaikh and Ganchimeg Ganbold*
4	“Дуу хоолой хувиргалтыг сайжруулахад авианы постериограмм (PPG)-д суурилсан аргыг хэрэгжүүлэх нь” *Lkhagvatseren Lkhagvasuren, Adiyabat Enkhjargal, Byambajav Dorj and Bayarpurev Mongol*
5	“Монгол хэлний ярианаас бичвэр буулгах системийн нууцлалын алгоритмуудын гүйцэтгэлийн жишиг судалгаа” *Дашренчин Эрдэнэболд, Эгшиглэн Наранбаатар, Нямжаргал Тэгшжаргал and Enkhtuya Tsogtbaatar*
6	“Монголын уртын дууны акустик онцлогт суурилсан хэв маягийн анализ” *Uranchimeg Norov, Amartuvshin Baatarkhuu, Adiyabat Enkhjargal, Sumiyakhand Dagdanpurev, Chuluuntsetseg Tseren and Byambajav Dorj*

Бага нөөцтэй нутгийн аялгын ярианы синтезийн судалгаа

Бямбадоржийн Золзаяа
Мэдээллийн технологийн тэнхим
ШУТИС, МХТС
Монгол улс, Улаанбаатар хот
b.zolzaya@must.edu.mn

Соном-Очирын Өлзийбаяр
Мэдээллийн технологийн тэнхим
ШУТИС, МХТС
Монгол улс, Улаанбаатар хот
ulziibayar@must.edu.mn

Батгэрэлийн Ариунзаяа
Мэдээллийн технологийн тэнхим
ШУТИС МХТС
Монгол улс, Улаанбаатар хот
b.ariunzaya@must.edu.mn

***Хураангуй*— Энэхүү судалгаанд хязгаарлагдмал өгөгдлийн нөхцөлд монгол хэлний буриад аялгад зориулсан текстээс яриа үүсгэх системийг VITS архитектурт суурилан хөгжүүлэх боломжийг судлав. Судалгаанд зорилтот өгөгдлөөр анхнаас нь сургах сургалт, ойролцоо аялгын өгөгдөлд суурилсан шилжүүлэн сургалт, хамтарсан сургалт, олон яригчийн сургалт зэрэг таван өөр сургалтын стратегийг харьцуулсан. Загваруудын гүйцэтгэлийг аудионы чанар, дуу хоолойн ижилсэл, үндсэн давтамжийн нийцэл гэсэн объектив хэмжүүрүүд болон тэдгээрийг нэгтгэсэн нийлмэл оноогоор үнэлэв. Туршилтын үр дүнгээр ойролцоо аялгын өгөгдлөөр урьдчилан сургалт хийж, дараа нь зорилтот буриад өгөгдлөөр нарийн тохируулсан загвар хамгийн өндөр нэгдсэн оноо (0.61) үзүүлсэн. Статистикийн шинжилгээгээр таван загварын хооронд нийт ялгаа статистикийн хувьд ач холбогдолтой болох нь тогтоогдсон (Friedman $\chi^2(4) = 32.72$, $p < 0.001$). Хос харьцуулалтын шинжилгээгээр энэхүү загвар нь хэд хэдэн загвараас статистикийн хувьд давуу байсан боловч олон яригчийн сургалтын дараах нарийн тохируулгатай загвартай харьцуулахад мэдэгдэхүйц ялгаа илрээгүй бөгөөд уг стратеги мөн өрсөлдөхүйц үр дүн үзүүлэх боломжтойг харуулсан. Судалгааны үр дүн нь бага нөөцтэй нөхцөлд урьдчилсан сургалт болон нарийн тохируулгад суурилсан шилжүүлэн сургалтын арга нь буриад аялгын ярианы синтезийн гүйцэтгэлийг сайжруулах үр дүнтэй арга болохыг харуулж байна.**

Түлхүүр үг*— ярианы синтез, шилжүүлэн сургалт, нөөц багатай хэл, аялга

I. Удиртгал

Хэл бол зөвхөн хүмүүсийн хоорондын харилцааны хэрэгсэл төдийгүй тухайн үндэстний түүх, соёл, сэтгэлгээ, ертөнцийг үзэх үзлийн тусгал болсон оюуны үнэт өв юм. Дэлхий дээр өнөөдөр хэдэн мянган хэл, аялгуу оршин байгаа боловч технологийн хурдацтай хөгжлийн нөхцөлд эдгээр хэлний зөвхөн тодорхой хэсэг нь л дижитал орчинд идэвхтэй ашиглагдаж байна. Ялангуяа цөөн тоотой угсаатны хэл болон нутгийн аялгуу нь цахим орчинд дэмжлэг багатай байдгаас шалтгаалан хэрэглээ нь буурч, улмаар үе дамжин уламжлагдах боломж нь багасах эрсдэлтэй тулгарч байна.

Сүүлийн жилүүдэд ярианы технологи хурдацтай хөгжиж, текстээс яриа үүсгэх (Text-to-Speech, ТЯҮ) системүүд нь мэдээлэл түгээх, боловсролын хүртээмжийг нэмэгдүүлэх, хөгжлийн бэрхшээлтэй иргэдэд зориулсан туслах технологи боловсруулах зэрэг олон салбарт өргөн хэрэглэгдэх болсон. Орчин үеийн гүн сургалтын аргуудад суурилсан ТЯҮ системүүд нь хүний яриатай ойролцоо, байгалийн мэт дуугаралт бүхий синтез үүсгэх боломжийг бүрдүүлж байгаа боловч ийм системийг үр дүнтэй сургахын тулд их хэмжээний тэмдэглэсэн ярианы өгөгдөл шаардлагатай байдаг.

Монгол хэл нь олон нутгийн аялгуутай бөгөөд эдгээр нь тухайн бүс нутгийн түүх, соёл, аман уламжлалтай салшгүй холбоотой байдаг. Монгол хэлний нутгийн аялгуунуудын дунд буриад аялга нь өөрийн гэсэн авиа зүй, үгийн сан, аялга өнгөний онцлог шинжүүдээрээ ялгардаг чухал нэгэн салбар юм. Буриад угсаатны аялгуу нь олон зууны туршид өөрийн онцлог шинжүүдийг хадгалсаар ирсэн. Гэвч орчин үеийн дижитал орчинд буриад аялгууг дэмжсэн хэлний технологийн шийдэл тун ховор, бараг байхгүй гэж хэлж болохоор байна.

Буриад аялгууны ярианы синтезийн системийг хөгжүүлэхэд тулгардаг гол бэрхшээл нь сургалтын өгөгдлийн хомс байдал юм. Орчин үеийн мэдрэлийн сүлжээнд суурилсан ярианы синтезийн загварууд нь ихэвчлэн хэдэн арван цагийн хэмжээтэй өндөр чанартай тэмдэглэсэн дуут өгөгдөл шаарддаг. Харин бага нөөцтэй хэл болон аялгууны хувьд ийм хэмжээний өгөгдөл цуглуулах нь цаг хугацаа, санхүү болон хүний нөөцийн хувьд хүндрэлтэй байдаг. Ийм нөхцөлд шилжүүлэн сургалт (transfer learning), өгөгдөл өсгөх (data augmentation) зэрэг аргачлалуудыг өргөн судлах болсон.

Энэхүү судалгааны ажилд монгол хэлний буриад аялгыг сонгон авч, хязгаарлагдмал өгөгдлийн нөхцөлд ярианы синтезийн загвар байгуулах, түүний чанарыг сайжруулах боломжийг судалсан.

II. Холбогдох судалгаанууд

Сүүлийн жилүүдэд ярианы синтезийн технологи гүн сургалтын аргуудын хөгжилтэй уялдан эрчимтэй хөгжиж байна. Анхны нейрон сүлжээнд суурилсан эхнээс нь төгсгөл хүртэл (end-to-end) суралцах ТЯҮ загваруудын нэг болох Tacotron нь текстийг шууд спектрограммд хөрвүүлэх боломжийг олгож, уламжлалт загваруудаас илүү байгалийн дуугаралттай синтез үүсгэх боломжийг бүрдүүлсэн [1]. Үүний дараа Tacotron2 [2] загвар танилцуулагдаж, мел-спектрограм үүсгэх нейрон сүлжээний загвар болон WaveNet [3] төрлийн аудио дохио үүсгэгчийг хослуулснаар синтезийн чанарыг мэдэгдэхүйц сайжруулсан.

Сүүлийн үеийн судалгаанд adversarial сургалт болон variational inference-ийг хослуулсан архитектур бүхий VITS (Variational Inference with adversarial learning for end-to-end Text-to-Speech) загвар өргөн хэрэглэгдэх болсон. Энэхүү загвар нь спектрограмм үүсгэх болон аудио дохио үүсгэх процессыг нэгтгэсэн эхнээс нь төгсгөл хүртэл суралцах бүтэцтэй бөгөөд бодит хүний яриатай ойролцоо чанартай синтез үүсгэх боломжийг олгодог [4]. Ийм архитектур нь олон хэлний болон бага өгөгдөлтэй нөхцөлд ярианы синтез хийх судалгаанд өргөн ашиглагдаж байна.

Гэсэн хэдий ч орчин үеийн ТЯҮ загваруудыг үр дүнтэй сургахын тулд их хэмжээний тэмдэглэсэн ярианы

өгөгдөл шаардлагатай байдаг тул ихэнх судалгаанууд англи, хятад зэрэг өгөгдөл ихтэй хэлнүүд дээр төвлөрөх хандлагатай байна. Энэ асуудлыг шийдвэрлэхийн тулд бага нөөцтэй (low-resource) хэл болон аялгуунуудад зориулсан судалгаанууд нэмэгдэж байгаа бөгөөд үүнд олон хэлний загвар (multilingual TTS), шилжүүлэн сургалт (transfer learning), болон өгөгдөл өсгөх (data augmentation) зэрэг аргачлалуудыг ашиглах болсон [5].

Олон хэлний ТЯУ судалгаанд олон хэлний өгөгдлийг нэгтгэн ашиглаж, хэл хоорондын нийтлэг төлөөллийг сургах замаар бага өгөгдөлтэй хэлний гүйцэтгэлийг сайжруулах боломжийг харуулсан байдаг [6, 7]. Yang болон He нарын судалгаанд 50 хэлний 1250 цагийн өгөгдлийг ашиглан нэгтгэсэн ТЯУ систем бүтээж, өгөгдлийн тэнцвэрт сургалтын стратегиар бага нөөцтэй хэлнүүдэд ч боломжийн чанартай синтез үүсгэх боломжтойг харуулсан [8]. Мөн хэл хооронд мэдлэг дамжуулах арга нь өгөгдөл ихтэй хэл дээр сурсан загварыг бага өгөгдөлтэй хэл дээр дахин тохируулан сургах замаар синтезийн чанарыг сайжруулах үр дүнтэй арга болохыг өмнөх судалгаанууд харуулсан [9].

Бага нөөцтэй нөхцөлд ТЯУ систем хөгжүүлэх өөр нэг чиглэл нь яригчид тааруулах (speaker adaptation) болон нарийн тохируулгын (fine-tuning) аргачлалуудыг ашиглах явдал юм. Jia нарын судалгаанд олон яригчийн өгөгдөл дээр сурсан загварыг шинэ яригчийн бага хэмжээний өгөгдлөөр тохируулан сургах боломжийг харуулсан бөгөөд энэ нь бага өгөгдөлтэй нөхцөлд ярианы синтезийн системийг хөгжүүлэхэд чухал ач холбогдолтой болохыг нотолсон [10].

Дээр дурдсан судалгаанууд нь бага нөөцтэй хэл болон аялгуунд зориулсан ТЯУ систем хөгжүүлэхэд олон хэлний сургалт, шилжүүлэн сургалт, тухайлбал нарийн тохируулга зэрэг аргачлалууд үр дүнтэй болохыг харуулж байна. Энэхүү судалгаанд мөн адил ийм аргачлалуудыг ашиглан монгол хэлний буриад аялгууны ярианы синтезийг хязгаарлагдмал өгөгдлийн нөхцөлд бүтээх боломжийг судалж, сургалтын өөр өөр стратегиудын гүйцэтгэлийг харьцуулан үнэлсэн болно.

III. Арга зүй

Энэхүү судалгаанд хязгаарлагдмал өгөгдөлтэй нөхцөлд буриад аялгын ярианы синтезийн загвар байгуулах зорилгоор гүн сургалтын аргад суурилсан ТЯУ архитектурыг ашигласан. Тодруулбал, орчин үеийн ТЯУ судалгаанд өргөн хэрэглэгдэж буй VITS загварыг ашиглан туршилтуудыг гүйцэтгэсэн [4]. Энэхүү бүлэгт судалгаанд ашигласан өгөгдөл, загвар болон сургалтын аргачлалыг тайлбарлана.

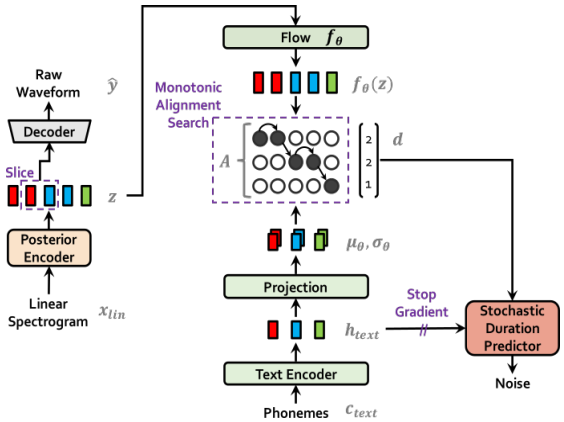
A. Өгөгдөл

Судалгаанд монгол хэлний халх аялгын болон буриад аялгын ярианы өгөгдлийг ашигласан. Халх аялгын өгөгдөл нь харьцангуй их хэмжээтэй бөгөөд студийн орчинд бэлтгэсэн 1 эрэгтэй яригчийн нийт 10 цаг орчим тэмдэглэсэн ярианы бичлэгээс бүрдэнэ. Харин буриад аялгын хувьд бэлтгэсэн өгөгдлийн хэмжээ харьцангуй бага бөгөөд 1 эмэгтэй яригчийн нийт 1 цаг орчим тэмдэглэсэн ярианы бичлэгийг ашигласан. Аудио файл бүр нь 22050 Гц дээжлэлтийн давтамжтай, нэг сувгийн, 16-бит PCM кодлолтой WAV форматтай байна Бага нөөцтэй хэл болон аялгууны хувьд ийм хэмжээний өгөгдөл нь орчин үеийн ТЯУ системийг шууд сургахад хангалтгүй байдаг. Иймээс судалгаанд их өгөгдөлтэй хэлний мэдээллийг ашиглан суурь загвар сургах, дараа нь зорилтот аялгын бага хэмжээний өгөгдлөөр дахин тохируулан сургах шилжүүлэн сургалтын стратегийг ашигласан.

Текст болон аудио өгөгдлийг урьдчилан боловсруулж бэлтгэх явцад текстийг хэвийн хэлбэрт оруулах (text normalization), тухайлбал тоо, тэмдэгт болон товчлолуудыг уншигдах хэлбэрт хөрвүүлэх ажил хийгдсэн. Харин аудио өгөгдлийг нэгэн жигд чанартай болгохын тулд дууны бичлэгүүдийг тодорхой давтамжид (22050 Гц) хөрвүүлэх, илүүдэл чимээг багасгах, мөн 10 секунд хүртэлх урттай хэсгүүдэд салгах зэрэг боловсруулалтыг хийсэн. Эдгээр алхмууд нь сургалтын өгөгдлийн чанарыг сайжруулж, загварын сургалтын тогтвортой байдлыг хангахад чухал үүрэг гүйцэтгэдэг. Мөн бид халх, буриад болон нэгтгэсэн өгөгдөл тус бүр дээр өгөгдлийг 90:10 харьцаатай хуваж сургалт болон баталгаажуулалтын багцыг үүсгэсэн.

B. VITS загвар

VITS загвар нь variational autoencoder (VAE), normalizing flow болон adversarial сургалтын аргуудыг нэгтгэсэн ТЯУ архитектур бөгөөд текстээс шууд аудио дохио үүсгэх чадвартайгаараа онцлог юм. Уламжлалт ТЯУ системүүдэд акустик загвар болон аудио дохио үүсгэх загварыг тусад нь сургах шаардлагатай байдаг бол VITS нь эдгээрийг нэг архитектурт нэгтгэснээр сургалтын үр ашиг болон синтезийн чанарыг сайжруулдаг. Зураг 1-т загварын архитектурыг харуулав. VITS загварын text encoder нь оролтын текстийг тэмдэгтийн дараалал хэлбэрээр авч контекстийн мэдээлэл бүхий далд төлөөлөл үүсгэнэ. Posterior encoder нь сургалтын үед бодит аудио дохионоос шугаман спектрограмм (linear spectrogram) гарган авч convolutional сүлжээ ашиглан латент төлөөлөл үүсгэдэг. Normalizing flow давхаргууд нь posterior encoder-ийн гаралтыг өмнөх (prior) тархалтын орон зайд хөрвүүлэх замаар variational inference-ийн сургалтыг хэрэгжүүлдэг. Decoder нь HiFi-GAN архитектурт суурилсан бөгөөд далд төлөөллөөс өндөр чанартай аудио долгион үүсгэнэ. Discriminator нь бодит болон синтезлэсэн аудиог ялган adversarial loss тооцоолсноор синтезийн дууны чанарыг сайжруулахад чиглэнэ.



Зураг 1. VITS загварын архитектур

C. Сургалтын аргачлал

Бага нөөцтэй нөхцөлд загварын гүйцэтгэлийг сайжруулах зорилгоор бид шилжүүлэн сургалтад суурилсан стратегийг хэрэгжүүлсэн. Өөрөөр хэлбэл хязгаарлагдмал өгөгдлийн нөхцөлд буриад аялгын ярианы синтезийн чанарт сургалтын стратегийн нөлөөг үнэлэх зорилгоор таван өөр аргачлалыг туршиж харьцуулсан. Нэгдүгээрт, зөвхөн 1 цагийн буриад ярианы өгөгдлөөр VITS загварыг анхнаас нь (from scratch) сургаж туршилт хийсэн (Model-A). Хоёрдугаарт, эхлээд 10 цагийн халх аялгын өгөгдлөөр суурь загварыг сургаж, дараа нь уг загварыг 1 цагийн буриад өгөгдлөөр нарийн тохируулга хийсэн (model-B). Гуравдугаарт, халх аялгын 10 цаг болон буриад аялгын 1 цагийн өгөгдлийг хамтад нь ашиглан суурь загварыг сургаж, дараа нь зөвхөн буриад өгөгдлөөр нарийн тохируулга хийсэн (model-C). Дөрөвдүгээрт, халх (10 цаг) болон буриад (1 цаг) аялгын өгөгдлийг хамтад нь ашиглан олон яригчийн (multispeaker) сургалтын аргачлал хэрэглэж, яригчийн вектор дүрслэл (speaker embedding) ашигласан хоёр яригчтай загварыг сургаж туршсан (Model-D). Тавдугаарт, дээрх олон яригчийн загварыг (Model-D) суурь болгон авч, зөвхөн буриад өгөгдлийг ашиглан нэг яригчид (буриад) чиглэсэн байдлаар үргэлжлүүлэн сургасан (Model-E).

Бүх туршилтад VITS загварын анхны архитектур болон бусад оновчлолын үндсэн тохиргоог өөрчлөлтгүй ашигласан. Урьдчилан сургасан суурь загваруудын хувьд epochs = 1000, learning rate = 5×10^{-5}, batch size = 32 гэж тохируулсан бол нарийн тохируулгын туршилтуудад epochs = 1000, learning rate = 1×10^{-5}, batch size = 16 гэж тус тус ашигласан. Мөн олон яригчийн загварын туршилтад зөвхөн n_speakers = 2 гэж тохируулсан. Туршилтад бүх загварыг урьдчилан тогтоосон epoch хүртэл сургасан. Загваруудын сургалтын аргачлалуудын харьцуулалтыг Хүснэгт 1-д үзүүлэв.

ХҮСНЭГТ I. СУРГАЛТЫН АРГАЧЛАЛУУДЫН ХАРЬЦУУЛАЛТ

Загвар	Сургалтын өгөгдөл	Сургалтын стратеги
Model-A	1 цаг буриад	Анхнаас нь буриад өгөгдлөөр сургасан сургалт (scratch)
Model-B	10ц халх → 1ц буриад	Халх суурь загвар, буриад өгөгдлөөр нарийн тохируулгатай
Model-C	10ц халх + 1ц буриад → 1ц буриад	Хамтарсан сургалтын суурь загвар, буриад өгөгдлөөр нарийн тохируулгатай
Model-D	10ц халх + 1ц буриад	Олон яригчийн загвар (яригчийн вектор ашигласан)
Model-E	Model-D → 1ц буриад	Олон яригчийн суурь загвар, буриад өгөгдлөөр нарийн тохируулгатай

IV. ҮНЭЛГЭЭНИЙ ХЭМЖҮҮР

Загваруудын гүйцэтгэлийн үнэлгээг хийхдээ загвар тус бүрээс үүсгэсэн 10 аудио дээжийг ашигласан. Үнэлгээг илүү найдвартай, олон талт байдлаар гүйцэтгэх зорилгоор загваруудын гүйцэтгэлийг дараах гурван үндсэн хэмжүүрээр тодорхойлов:

- *Аудионы чанар (UTMOSv2 Mean Opinion Score (MOS)*): Ярианы ерөнхий чанарыг автоматаар үнэлэх хэмжүүр бөгөөд 1–5 хооронд утга авна. Утга өндөр байх тусам ярианы чанар сайжирч, илүү байгалийн сонсогдоно.
- *Дуу хоолойн ижилсэл (Speaker Similarity)*: Үүсгэсэн ярианы дуу хоолой нь зорилтот яригчийн дуу хоолойтой хэр зэрэг ижил байгааг хэмжих үзүүлэлт бөгөөд 0-1 хооронд утга авна. Утга 1-д ойртох тусам дуу хоолойн ижилсэл өндөр байна.
- *Үндсэн давтамжийн нийцэл (Pitch Correlation)*: Үүсгэсэн ярианы аялгуу (pitch) нь бодит ярианы аялгуутай хэр зэрэг таарч байгааг хэмжих корреляцийн үзүүлэлт бөгөөд [−1, 1] хооронд утга авдаг. Утга 1 нь үүсгэсэн аялгуу бодит аялгуутай төгс таарч байгааг, 0 нь ямар ч хамааралгүй байгааг, харин −1 нь урвуу буюу эсрэг хамааралтай байгааг илтгэнэ. Ингэснээр утга 1-д ойртох тусам үүсгэсэн аялгуу бодит аялгууны хэлбэр, давтамж зэрэг шинж чанарт илүү үнэн зөв таарч байгааг илтгэнэ.
- *Нэгдсэн оноо:* Загваруудын нийт гүйцэтгэлийг үнэлэхдээ дээрх гурван хэмжүүрийг нэгтгэх ашигласан. Тодруулбал, аудионы чанар болон дуу хоолойн ижилсэлд тус бүр 40%, харин үндсэн давтамжийн нийцэлд 20%-ийн жин өгч, нийт оноог тооцоолон загваруудын цогц гүйцэтгэлийг харьцуулсан. Энэхүү жингийн хуваарилалтыг ТЯУ чанарын үнэлгээнд аудио чанар болон дуу хоолойн ижилсэл нь аялгууны нийцлээс илүү чухал гэж үзсэн эмпирик ажиглалтад тулгуурлан тогтоосон. Гэсэн хэдий ч энэхүү жингийн сонголт нь тодорхой хэмжээнд субъектив шинжтэй бөгөөд өөр жингийн тохиргоо ашигласан тохиолдолд загваруудын эрэмбэ хэсэгчлэн өөрчлөгдөх боломжтой. Иймд энэхүү нэгдсэн оноо нь абсолют үнэлгээ бус, харин олон хэмжүүрийг нэгтгэсэн харьцуулалтын үзүүлэлт болохыг тэмдэглэж байна. Гурван хэмжүүр өөр өөр хэмжээстэй (MOS: 1–5, Speaker Similarity: 0–1, Pitch Correlation: [−1, 1]) тул нормчлол хийж [0, 1] мужид хөрвүүлсэн. Нэгдсэн оноог дараах (1)-р тооцно:

$$\text{Нийт оноо} = MOS_{norm} * 0.4 + \text{Speaker Similarity} * 0.4 + \text{PitchCorr}_{norm} * 0.2 \quad (1)$$

V. ТУРШИЛТ, ҮР ДҮН

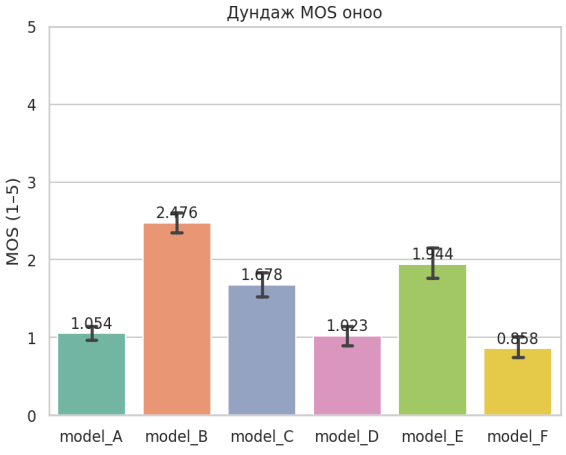
A. Аудионы чанар (MOS үнэлгээ)

Энэхүү судалгаанд ярианы синтезийн чанарыг үнэлэх зорилгоор UTMOSv2 загварыг ашигласан [11]. Уг загвар нь VoiceMOS Challenge 2022-д танилцуулагдсан MOS (Mean Opinion Score) таамаглах системүүдийн нэг бөгөөд орчин үеийн (state-of-the-art) арга юм. Тус загвар нь ярианы дохионоос хүний субъектив үнэлгээг (MOS) автоматаар таамаглах зорилготой бөгөөд лавлагаа (reference) аудио шаардалгүйгээр зөвхөн синтезлэсэн яриаг оролт болгон ашигладаг. Иймд хүний сонсголын үнэлгээг орлуулах боломжтой объектив хэмжүүрийн нэг хэлбэр гэж үзэж болно. UTMOSv2 нь хэлнээс үл хамааран ярианы чанарыг үнэлэх чадвартай тул уг загварыг ашигласнаар синтезлэсэн ярианы байгалийн

байдал, ойлгомжтой байдал болон дууны чанарыг нэгтгэн үнэлэх боломж бүрдсэн.

Туршилтад дээрх загваруудаас гадна Model-F-ийг нэмэлтээр харьцуулан үнэлэв. Model-F нь халх аялгын 10 цаг болон буриад аялгын 1 цагийн өгөгдлийг ялгалгүйгээр хамтад нь ашиглан сургасан, өөрөөр хэлбэл Model-C загварын суурь (урьдчилан сургасан) хувилбар юм. Энэхүү загвар нь олон яригчийн өгөгдлийг нэгтгэн ашигласан боловч яригчийн ялгарал буюу яригчийн вектор дүрслэл ашиглаагүй. Model-F загварыг энэхүү судалгаанд зөвхөн аудионы чанарын хэмжүүрт үнэлж, олон яригчийн өгөгдлийг шууд нэгтгэн сургасан стратегийн нөлөөг тодорхойлох зорилгоор ашиглав.

Зураг 2-н үр дүнгээс харахад Model-A, Model-D болон Model-F загварууд харьцангуй бага гүйцэтгэл үзүүлж байна. Эдгээр загваруудын ялгаг авч үзвэл сургалтын стратегийн нөлөө тодорхой ажиглагдана. Юуны өмнө, Model-F нь халх болон буриад яригчийн өгөгдлийг ялгалгүйгээр шууд нэгтгэн сургасан загвар бөгөөд хамгийн бага MOS үзүүлэлттэй байна. Энэ нь олон яригчийн өгөгдлийг ямар нэг ялгаралгүйгээр хамтад нь сургах үед загвар нь яригч бүрийн онцлогийг ялган сурах чадваргүй болж, улмаар ярианы чанар буурдаг болохыг харуулж байна. Харин Model-D нь олон яригчийн сургалтыг яригчийн вектор дүрслэл ашиглан хийсэн бөгөөд Model-F-ээс илүү гүйцэтгэл үзүүлсэн байна. Хэдийгээр энэ нь өндөр үзүүлэлт биш боловч яригчдыг ялгах механизм нэвтрүүлснээр загвар нь илүү тогтвортой, ойлгомжтой аудио үүсгэх боломжтойг илтгэнэ. Өөрөөр хэлбэл, олон яригчийн өгөгдлийг ашиглахдаа тэдгээрийн ялгааг загварт тусгах нь чухал болохыг баталж байна. Гэвч 1 цагийн буриад өгөгдөл нь маш бага тул буриад яригчийн синтезийн чанарт мэдэгдэхүйц сайжрал гарч чадсангүй. Мөн Model-A нь зөвхөн буриад өгөгдөл дээр анхнаас нь сурсан загвар бөгөөд Model-D-тэй ойролцоо түвшний гүйцэтгэлтэй байна. Энэ нь бага өгөгдөлтэй нөхцөлд эхнээс нь сурах сургалт аудионы чанарын хувьд хангалтгүй болохыг харуулж байна.

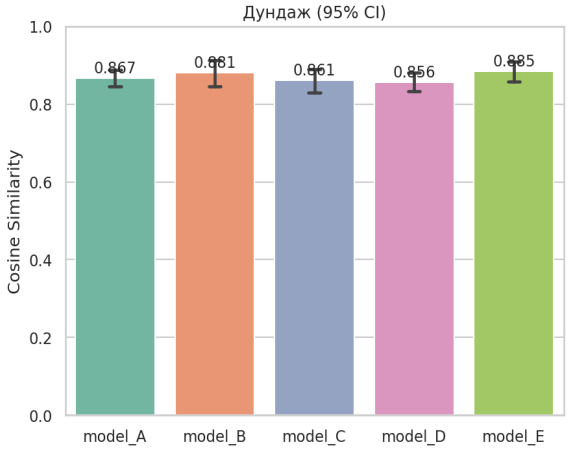


Зураг 2. Аудионы чанарын үнэлгээний үр дүн.

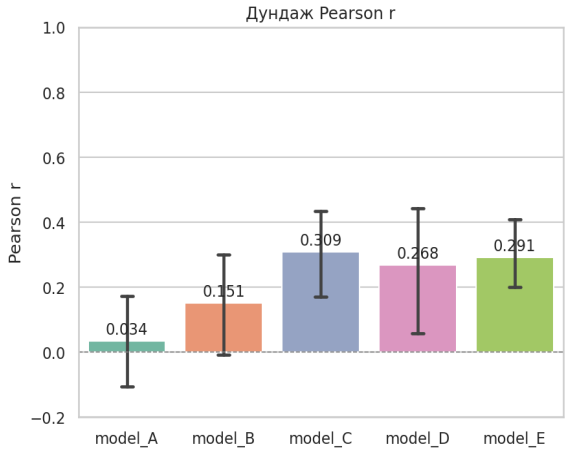
Харин шилжүүлэн сургалт ашигласан загварууд болох Model-B, Model-C, Model-E бага нөөцтэй нөхцөлд илүү сайн гүйцэтгэл үзүүлж байгааг харуулж байна. Тухайлбал, Model-B нь халх аялгын өгөгдөл дээр урьдчилан сургалт хийж, дараа нь буриад өгөгдлөөр нарийн тохируулга хийснээр хамгийн өндөр MOS оноо авсан нь ойролцоо хэлний өгөгдлөөс мэдлэг дамжуулах нь ярианы синтезийн чанарыг мэдэгдэхүйц сайжруулдаг болохыг нотолж байна. Model-F болон Model-D загваруудын нарийн тохируулсан хувилбарууд болох Model-C болон Model-E-ийн үр дүнг харьцуулахад, аль аль нь суурь загваруудын гүйцэтгэлийг мэдэгдэхүйц сайжруулсан байна. Гэвч Model-E нь Model-C-ээс илүү өндөр гүйцэтгэл үзүүлж байгаа нь олон яригчийн сургалтын үед яригчийн ялгарал (яригчийн вектор дүрслэл)-ыг тусгасан суурь загвар нь дараагийн нарийн тохируулгад илүү үр ашигтай болохыг харуулж байна. Нөгөө талаас, эдгээр хоёр загварын гүйцэтгэл нь Model-B-ийн түвшинд хүрээгүй хэвээр байна. Энэ нь олон яригчийн сургалт болон холимог өгөгдөл ашиглах стратеги нь тодорхой давуу талтай хэдий ч, бага өгөгдлийн нөхцөлд Model-B-д хэрэглэсэн шилжүүлэн сургалтын аргаас илүү давуу гүйцэтгэл үзүүлж чадахгүй байгааг илтгэж байна. Иймээс, туршилтын үр дүнгээс дүгнэхэд ойролцоо аялгын өгөгдлөөр урьдчилан сургалт хийж, дараа нь зорилтот буриад өгөгдлөөр нарийн тохируулга хийх стратеги (Model-B) нь бага нөөцтэй нөхцөлд ярианы синтезийн чанарыг хамгийн үр дүнтэйгээр сайжруулж буйг харуулж байна. Model-F загвар нь харьцангуй сул гүйцэтгэл үзүүлсэн тул цаашдын үнэлгээний харьцуулалтад оруулахгүй байхаар шийдвэрлэсэн болно.

B. Дуу хоолойн ижилсэл (Speaker Similarity)

Үүсгэсэн аудионы дуу хоолойн ижил байдлын үнэлгээнд Resemblyzer сангийн VoiceEncoder загварыг ашигласан [12]. VoiceEncoder загвар нь GE2E loss (Generalized End-to-End) алдагдлын функцэд суурилсан яригч таних (speaker verification) загвар бөгөөд олон яригчийн өгөгдөл дээр сургаснаар яригчийн ялгарах онцлогийг үр дүнтэй сурдаг. Энэхүү загварыг VoxCeleb өгөгдлийн багц дээр сургасан, нийтэд нээлттэй урьдчилан сургасан загвар юм. Уг загвар нь ярианы аудио дохионоос яригчийн онцлогийг илэрхийлэх 256 хэмжээст embedding вектор гарган авах чадвартай. Үнэлгээ хийхдээ синтезлэсэн аудио болон харгалзах лавлагаа (reference) аудио хос бүрээс embedding векторуудыг тус тус гарган авч, эдгээрийн хоорондын cosine similarity-г тооцоолсон. Cosine similarity нь хоёр векторын чиглэлийн төстэй байдлыг хэмжих бөгөөд утга нь 1-д ойртох тусам тухайн хоёр аудио нь нэг ижил яригчийн дуу хоолойд илүү ойр байгааг илтгэнэ. Уг аргыг ашигласнаар синтезлэсэн яриа нь лавлагаа ярианы яригчийн онцлогийг хэр сайн хадгалж байгааг тоон үзүүлэлтээр (objective байдлаар) үнэлэх боломж бүрдсэн. Түүнээс гадна хүний субъектив дүгнэлтээс үл хамааран, яригчийн төстэй байдлыг автоматаар хэмжих боломжтой давуу талтай.



Зураг 3. Дуу хоолойн ижилслийн үнэлгээний үр дүн



Зураг 4. Үндсэн давтамжийн нийцлийн үнэлгээний үр дүн

Зураг 3-аас харахад бүх загваруудын хувьд cosine similarity 0.85-аас дээш утгатай байгаа нь синтезлэсэн яриа нь лавлагаа ярианы яригчийн онцлогийг боломжийн хэмжээнд хадгалж байгааг илтгэнэ. Эдгээрээс Model-B, Model-E нь эх яригчийн хоолойг төстэй гаргах чадвар хамгийн өндөр байгааг харуулж байна. Хамгийн өндөр үзүүлэлт үзүүлсэн Model-E нь эхлээд олон яригчийн орчинд ерөнхий яригчийн төлөөллийг сурч авсны дараа зорилтот яригч дээр нарийн тохируулга хийх нь яригчийн онцлогийг илүү сайн хадгалах боломжийг бүрдүүлж байгааг харуулж байна. Харин Model-A нь бага өгөгдлөөр анхнаас нь сургасан хэдий ч 0.867 гэсэн харьцангуй өндөр үзүүлэлттэй байгаа нь яригчийн ялгарах онцлогийг (speaker identity) сурах боломжтойг харуулж байна. Гэсэн хэдий ч MOS үнэлгээтэй харьцуулахад дуу хоолойн ижил байдал өндөр байх нь заавал дууны чанар сайн байхыг илэрхийлэхгүй болох нь Model-A, C, D-ээс ажиглагдсан. Өөрөөр хэлбэл, энэ хоёр үзүүлэлт нь харьцангуй бие даасан шинж чанарууд болохыг харуулж байна.

C. Үндсэн давтамжийн нийцэл (Pitch Correlation)

Үндсэн давтамжийн нийцэл буюу тонны (F0) корреляцын үнэлгээнд бид синтезлэгдсэн болон лавлагаа аудионы хос бүрээс фрэйм бүрийн үндсэн давтамжийн (F0) цуваг гарган авахдаа librosa сангийн pyin() функцийг ашигласан [13]. pyin() функц нь Mauch ба Dixon нарын санал болгосон магадлалд суурилсан pYIN алгоритмд тулгуурладаг бөгөөд дуутай (voiced) болон дуугүй (unvoiced) хэсгүүдийг ялган F0 цуваг илүү нарийвчлалтай тодорхойлдог [14]. Үүсгэсэн F0 цуваг зөвхөн дуутай хэсгүүдээр шүүх замаар боловсруулж, үүний дараа sciру сангийн pearsonr функцээр синтезлэгдсэн болон лавлагаа аудионы F0 цувааны Pearson корреляцын коэффициентийг тооцоолсон. Энэ арга нь синтезийн үед хадгалагдсан тонны хэлбэлзлийг бодит хүний ярианы тонны хэлбэлзэлтэй хэр сайн нийцэж буйг хэмжихэд хэрэглэгддэг.

Зураг 4-д бүх загваруудын үндсэн давтамжийн нийцлийн үнэлгээний дүнг харуулсан болно. Model-A нь хамгийн муу үр дүнтэй байсан бөгөөд F0 корреляцийн утга 0.034 байв. Энэ нь хязгаарлагдмал өгөгдөл дээр анхнаас нь сургах нь тонны хэлбэлзлийг зөв дагаж чадахгүй байгааг харуулж байна. Харин Model-B нь шилжүүлэн сургалтын эерэг нөлөөг харуулж, Model-A -ээс давж, F0 корреляцийн утга 0.151 хүртэл өссөн. Загвар нь халх өгөгдлөөс олж авсан pitch дүрслэлийг буриад өгөгдөлд тохируулж чадсан байна. Model-C нь хамгийн өндөр F0 корреляцийн утгатай буюу 0.309 болсон. Энэ нь суурь загвар дээр зорилтот аялгыг ашиглах нь синтезлэгдсэн тонны хэлбэлзлийг бодит яриатай хамгийн сайн нийцүүлдэгийг харуулж байна. Model-D нь Model-A, B-ээс илүү үр дүн үзүүлсэн нь олон яригчийн стратеги нь загварт тонны хэлбэлзлийг илүү сайн мэдрүүлэх боломж олгодгийг харж болохоор байна. Харин Model-E хувьд F0 корреляци 0.291 хүртэл өссөн нь олон яригчийн суурь загварыг зорилтот яригчид онцгойлон тохируулснаар тонны нийцлийг сайжруулж болохыг нотолж байна. Эдгээр үр дүнгээс үзэхэд, эхний суурь загвар зорилтот аялгын ерөнхий просодик (intonation, pitch) хэв шинжийг сурч авсны дараа зорилтот аялгад нарийн тохируулга хийх нь синтезийн аялгууны F0-г бодит ярианы тонны хэлбэлзэлтэй ойртуулж байгааг харуулж байна. Өөрөөр хэлбэл урьдчилсан сургалт болон нарийн тохируулга нь бага хэмжээний зорилтот өгөгдөлтэй үед дуутай хэсгийн тонны корреляцийг сайжруулахад чухал үүрэг гүйцэтгэдэг нь тодорхой байна.

D. Нэгдсэн дүн шинжилгээ

Энэхүү хэсэгт өмнө тодорхойлсон гурван үндсэн хэмжүүр болох аудионы чанар, дуу хоолойн ижилсэл, болон үндсэн давтамжийн нийцэлд тулгуурлан загваруудын гүйцэтгэлийг нэгдсэн байдлаар үнэлэв. Хэмжүүр тус бүр нь ярианы синтезийн өөр өөр чухал талыг илэрхийлдэг тул эдгээрийг (1)-р жинлэн нэгтгэж, загваруудын нийт гүйцэтгэлийг илүү цогц байдлаар харьцуулах боломжийг бүрдүүлсэн. Хүснэгт 2-т нормчлогдсон утгууд болон жингийн коэффициентийг ашиглан тооцоолсон нэгдсэн оноог харуулсан бөгөөд үүнд тулгуурлан өөр өөр сургалтын стратегиудын үр нөлөөг харьцуулан дүн шинжилгээ хийв.

ХҮСНЭГТ II. ОБЪЕКТИВ ҮНЭЛГЭЭНИЙ НЭГДСЭН ОНОО

Загвар	Дундаж	Стандарт хазайлт	Хамгийн бага утга	Хамгийн их утга
Model-A	0.455	0.029	0.412	0.511
Model-B	**0.615**	0.037	0.542	0.672
Model-C	0.513	0.029	0.484	0.579
Model-D	0.471	0.024	0.4280	0.501
Model-E	0.577	0.045	0.531	0.666

Зөвхөн 1 цагийн буриад өгөгдлийг ашиглан анхнаас нь сургасан загвар тодорхой түвшний ярианы синтезийн чадвар эзэмшсэн нь VITS загвар бага өгөгдөлтэй нөхцөлд ч ажиллах боломжтойг харуулсан боловч гүйцэтгэл хязгаарлагдмал хэвээр байна. Нэгдсэн онооны үр дүнгээс харахад Model-B (0.615) хамгийн өндөр гүйцэтгэл үзүүлсэн байна. Энэ нь ойролцоо хэл буюу халх аялгын өгөгдлөөр урьдчилсан сургалт хийж, дараа нь зорилтот аялгад нарийн тохируулсан шилжүүлэн сургалтын стратеги нь аудионы чанар болон дуу хоолойн ижилсэл дээр өндөр үр дүн өгч, нийт гүйцэтгэлд эерэгээр нөлөөлснийг харуулж байна.

Халх болон буриад өгөгдлийг хамтад нь ашиглан (яригчийн вектор дүрслэл ашиглахгүйгээр) анхан шатны сургалтыг хийж, дараа нь зөвхөн буриад өгөгдлөөр нарийн тохируулга хийсэн Model-C загвар нь анхнаас нь сургасан загвараас илүү гүйцэтгэл үзүүлсэн боловч Model-B аргаас доогуур үр дүнтэй гарсан байна. Энэ нь хамтарсан сургалтын явцад хоёр аялгын өгөгдлийг ялгалгүйгээр нэгтгэн сургах үед зорилтот буриад аялгын онцлог хангалттай тод ялгарч сураагүй байж болзошгүйг илтгэж байна. Мөн дараах нарийн тохируулга нь тодорхой хэмжээнд сайжруулалт өгсөн ч урьдчилан нэг аялга (халх)-д төвлөрөн сурсан загвар шиг (Model-B) үр дүнтэй мэдлэг дамжуулалт үүсгэж чадаагүй байж болох юм. Иймээс энэ нь аялгууны нийцэл сайн боловч аудионы ерөнхий чанар сул байсан.

Олон яригчийн сургалтын үр дүн (Model-D) харьцангуй бага байсан нь олон яригчийн орчинд бага хэмжээний буриад өгөгдөл давамгай өгөгдөлд “дарагдах” асуудал байгааг харуулж байна. Гэвч Model-D нь Model-A-тай харьцуулахад нэгдсэн оноогоор бага зэрэг өндөр үзүүлэлттэй гарсан бөгөөд энэ нь голчлон pitch correlation үзүүлэлтийн сайжралтай холбоотой байж болзошгүй. Энэ нь яригчийн вектор дүрслэл ашигласан олон яригчийн сургалт нь тонны хэв маяг (prosody)-ийг тодорхой хэмжээнд илүү сайн барихад нөлөөлсөн байх магадлалтай. Харин уг загварыг зөвхөн буриад өгөгдлөөр үргэлжлүүлэн сургасан Model-E нь Model-B-тэй өрсөлдөхүйц өндөр гүйцэтгэл үзүүлж, Model-C-ээс давсан үр дүн үзүүлсэн нь олон яригчийн сургалтыг урьдчилсан шат болгон ашиглаад дараа нь зорилтот яригчид чиглэсэн нарийн тохируулга хийх нь бага өгөгдөлтэй нөхцөлд үр дүнтэй стратеги болохыг харуулж байна. Гэвч эдгээр ялгаа нь статистикийн хувьд ач холбогдолтой эсэхийг шалгах шаардлагатай.

E. Статистикийн ач холбогдлын шинжилгээ

Таван загварын хооронд статистикийн хувьд ач холбогдолтой ялгаа байгаа эсэхийг тодорхойлохын тулд Friedman-ийн параметрийн бус тестийг ашиглав. Энэхүү тест нь давтагдсан хэмжилтийн өгөгдөлд тохиромжтой бөгөөд нормал тархалтын таамаглал шаардахгүй. Шинжилгээний үр дүнд загваруудын хооронд статистикийн ач холбогдолтой ялгаа илэрсэн ($\chi^2(4) = 32.72$, $p < 0.001$). Kendall-ийн $W = 0.397$ утга нь дунд зэргийн түвшний нийцтэй байдлыг илэрхийлж байна.

Friedman тестийн үр дүнгээр нийт ялгаа илэрсэн тул хос харьцуулалтын ялгааг тодорхойлох зорилгоор Wilcoxon signed-rank тестийг ашиглав. Олон харьцуулалтаас үүдэлтэй I төрлийн алдааг хянахын тулд Bonferroni залруулга хэрэглэж, ач холбогдлын түвшинг $\alpha = 0.005$ болгон тогтоов. Үр дүнгээс харахад бидний хамгийн сайн гэж тодорхойлсон Model-B нь Model-A, Model-C болон Model-D-ээс статистикийн хувьд ач холбогдолтойгоор давуу байсан ($p = 0.002$) бол Model-B болон Model-E-ийн хоорондох ялгаа нь Bonferroni залруулгын дараа статистикийн ач холбогдолд хүрээгүй ($p = 0.049 > 0.005$). Мөн Model-A ба Model-D ($p = 0.375$), Model-C ба Model-E ($p = 0.037$) хос харьцуулалтууд статистикийн ач холбогдолгүй гарсан.

Эдгээр үр дүн нь нийт оноонд тулгуурлан Model-B-ийг хамгийн сайн загвар хэмээн нэрлэх боломжийг өгч байгаа ч Model-E-тэй харьцуулсан ялгаа нь цаашид томоохон дээж дээр нотлогдох шаардлагатай хэвээр байна.

F. Хязгаарлалт ба цаашдын ажил

Энэхүү судалгаа нь бага нөөцтэй нөхцөлд буриад аялгын ярианы синтезийг судалсан боловч дараах хязгаарлалттай. Нэдүгээрт, ашигласан буриад хэлний өгөгдлийн хэмжээ нь хэл шинжлэлийн бүрэн шинжилгээ хийхэд хангалтгүй. Тиймээс энэхүү ажлыг бүрэн аялгын загварчлал гэхээс илүү зорилтот яригчийн дуу хоолойд ойртсон буриад аялгын синтез гэж үзэх нь зүйтэй. Хоёрдугаарт, уг судалгаанд хэл шинжлэлийн нарийвчилсан шинжилгээ хийгдээгүй, мөн субъектив MOS үнэлгээ ороогүй. Үнэлгээ нь зөвхөн объектив метрикүүд дээр тулгуурласан. Иймээс сонсголын чанар болон аялгын бодит шинж чанар хадгалагдсан эсэх талаар бүрэн дүгнэлт хийх боломжгүй. Цаашид (1) өгөгдлийн хэмжээг 2–3 цаг болгон нэмэгдүүлэх, үр дүнд үзүүлэх нөлөөг харьцуулах, (2) субъектив MOS үнэлгээ болон хэл шинжлэлийн мэргэжилтний оролцоотой фонетик, просодик анализыг нэмж, системийн гүйцэтгэл болон дүгнэлтийн найдвартай байдлыг сайжруулах шаардлагатай.

VI. ДҮГНЭЛТ

Энэхүү судалгаанд хязгаарлагдмал өгөгдөлтэй нөхцөлд монгол хэлний буриад аялгын ярианы синтезийн загвар байгуулах боломжийг VITS архитектур болон шилжүүлэн сургалтын арга зүйгээр судалсан. Нийт таван сургалтын стратегийг туршиж, объектив метрикүүд болох аудионы чанар, дуу хоолойн ижилсэл, үндсэн давтамжийн нийцэл болон нормчилсон нийт оноогоор үнэлж харьцуулсан.

Үр дүнгээс харахад, зөвхөн зорилтот өгөгдлөөр анхнаас сургасан загвар тодорхой хэмжээний чадвартай байсан ч синтезийн чанар, яригчийн онцлог, тонны хэв маяг нь харьцангуй хязгаарлагдмал байв. Ойролцоо хэл буюу халх аялгын өгөгдлөөр урьдчилан сургалт хийж,

дараа нь зорилтот буриад өгөгдлөөр нарийн тохируулга хийсэн шилжүүлэн сургалтын стратеги нь нийт үнэлгээгээр хамгийн өндөр үзүүлэлттэй гарсан.

Статистикийн шинжилгээгээр таван загварын хооронд нийт ялгаа статистикийн хувьд ач холбогдолтой болох нь тогтоогдсон. Post-hoc шинжилгээгээр нийт үнэлгээгээр хамгийн өндөр үзүүлэлттэй гарсан загвар нь Model-A, Model-C болон Model-D-ээс статистикийн хувьд ач холбогдолтойгоор давуу байсан, харин олон яригчийн өгөгдөлд тулгуурлан урьдчилан сургалт хийж, зорилтот өгөгдлөөр үргэлжлүүлэн нарийн тохируулсан загвартай статистикийн ач холбогдолтой ялгаа илрээгүй. Энэ нь олон яригчийн урьдчилсан сургалтыг нарийн тохируулгатай хослуулах нь мөн өрсөлдөхүйц үр дүн үзүүлэх боломжтойг харуулж байна.

Иймээс урьдчилсан сургалт болон нарийн тохируулгад суурилсан стратегиуд нь бага өгөгдөлтэй нөхцөлд буриад аялгын ярианы синтезийн гүйцэтгэлийг сайжруулах үр дүнтэй хандлага болох нь харагдсан. Ялангуяа ойролцоо аялгын өгөгдлийг ашигласан шилжүүлэн сургалт нь бага нөөцтэй нөхцөлд өндөр гүйцэтгэлд хүрэх боломжтой болохыг энэхүү судалгаа харуулж байна.

Цаашдын судалгаанд буриад аялгын өгөгдлийн хэмжээг нэмэгдүүлэх, субъектив MOS үнэлгээ хийх, мөн хэл шинжлэлийн түвшний нарийвчилсан шинжилгээ хийх замаар загварын бодит гүйцэтгэлийг илүү нарийвчлан үнэлэх шаардлагатай байна.

Ном зүй

[1] Wang, Yuxuan, R. J. Skerry-Ryan, Daisy Stanton, Yonghui Wu, Ron J. Weiss, Navdeep Jaitly, Zongheng Yang et al. "Tacotron: Towards end-to-end speech synthesis." *arXiv preprint arXiv:1703.10135* (2017).

[2] Shen, Jonathan, Ruoming Pang, Ron J. Weiss, Mike Schuster, Navdeep Jaitly, Zongheng Yang, Zhifeng Chen et al. "Natural tts synthesis by conditioning wavenet on mel spectrogram predictions." In *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, pp. 4779-4783. IEEE, 2018.

[3] Van Den Oord, Aaron, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu. "Wavenet: A generative model for raw audio." *arXiv preprint arXiv:1609.03499* 12, no. 1 (2016).

[4] Kim, Jaehyeon, Jungil Kong, and Juhee Son. "Conditional variational autoencoder with adversarial learning for end-to-end text-to-speech." In *International conference on machine learning*, pp. 5530-5540. PMLR, 2021.

[5] Byambadorj, A., Ohta, K., & Kitaoka, N. (2021). Text-to-speech system for low-resource language using cross-lingual transfer learning and data augmentation. EURASIP Journal on Audio, Speech, and Music Processing, 2021(1), pp. 1–20.

[6] Do, Phat, Matt Coler, Jelske Dijkstra, and Esther Klabbers. "Strategies in transfer learning for low-resource speech synthesis: Phone mapping, features input, and source language selection." *arXiv preprint arXiv:2306.12040* (2023).

[7] Cai, Zexin, Yaogen Yang, and Ming Li. "Cross-lingual multi-speaker speech synthesis with limited bilingual training data." *Computer Speech & Language* 77 (2023): 101427.

[8] Yang, Jingzhou, and Lei He. "Towards universal text-to-speech." In *Proc. Interspeech 2020*, pp. 3171-3175. 2020.

[9] Tu, Tao, Yuan-Jui Chen, Cheng-chieh Yeh, and Hung-Yi Lee. "End-to-end text-to-speech for low-resource languages by cross-lingual transfer learning." *arXiv preprint arXiv:1904.06508* (2019).

[10] Jia, Ye, Yu Zhang, Ron Weiss, Quan Wang, Jonathan Shen, Fei Ren, Patrick Nguyen, Ruoming Pang, Ignacio Lopez Moreno, and Yonghui Wu. "Transfer learning from speaker verification to multispeaker text-to-speech synthesis." *Advances in neural information processing systems* 31 (2018).

[11] Saeki, Takaaki, Detai Xin, Wataru Nakata, Tomoki Koriyama, Shinnosuke Takamichi, and Hiroshi Saruwatari. "Utmos: Utokyo-sarulab system for voicemos challenge 2022." *arXiv preprint arXiv:2204.02152* (2022).

[12] Wan, Li, Quan Wang, Alan Papir, and Ignacio Lopez Moreno. "Generalized end-to-end loss for speaker verification." In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4879-4883. IEEE, 2018.

[13] McFee, Brian, Colin Raffel, Dawen Liang, Daniel PW Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto. "librosa: Audio and music signal analysis in python." *SciPy* 2015, no. 18-24 (2015): 7.

[14] Mauch, Matthias, and Simon Dixon. "pYIN: A fundamental frequency estimator using probabilistic threshold distributions." In *2014 ieee international conference on acoustics, speech and signal processing (icassp)*, pp. 659-663. IEEE, 2014.

Латин галиглалтаас кирилл рүү: Нөөц багатай хэлэнд хэлний том загварт суурилсан стратеги

Соном-Очирын Өлзийбаяр[1],
Мэдээллийн технологийн тэнхим
ШУТИС, МХТСургууль
Монгол улс, Улаанбаатар хот
ulziibayar@must.edu.mn

Бямбадоржийн Золзаяа[2]
Мэдээллийн технологийн тэнхим
ШУТИС, МХТСургууль
Монгол улс, Улаанбаатар хот
b.zolzaya@must.edu.mn

Батгэрэлийн Ариунзаяа[3]
Мэдээллийн технологийн тэнхим
ШУТИС, МХТСургууль
Монгол улс, Улаанбаатар хот
b.ariunzaya@must.edu.mn

***Хураангуй*— Цахим орчинд монгол хэлийг латин үсгээр галиглан бичих үзэгдэл өргөн тархсанаар шуугиантай, стандарт бус бичвэрийн хэмжээ эрс нэмэгдэж байна. Энэ нь эх хэлний боловсруулалтын (NLP) хэрэглээнд чанартай өгөгдөл бэлтгэхэд хүндрэл учруулж, эх хэлний бичгийн соёлд сөргөөр нөлөөлөх эрсдэлтэй. Иймээс нөөц багатай хэлний нөхцөлд латин галиглалтыг стандарт кирилл бичигт үр дүнтэй хэвийн болгох арга боловсруулах нь чухал асуудал юм. Мөн нийгмийн сүлжээнд буй их хэмжээний галигласан бичвэр нь эх хэлний боловсруулалтын хувьд чухал өгөгдөл боловч хэвийн болголгүйгээр шууд ашиглах боломжгүй байдаг. Энэхүү судалгаанд бид шуугиантай латин үсгээр галигласан монгол бичвэрийг стандарт кирилл хэлбэрт хөрвүүлэхэд тэмдэгт-түвшний seq2seq загварыг суурь болгон ашиглаж, GPT-2 загварт суурилсан том хэмжээний хэлний загваруудын боломжийг судлав. Туршилтад 4,000 өгүүлбэр бүхий гараар зэрэгцүүлсэн өгөгдлийн багцыг ашиглаж, үгийн болон өгүүлбэрийн түвшинд загваруудыг сургасан. Мөн урьдчилан сурсан жинг ашиглан нарийн тохируулах (fine-tune) болон эхнээс нь шинээр (scratch) сургах стратегиудыг, түүнчлэн анхны GPT-2 токенизатор болон Монгол BERT-д суурилсан холимог (hybrid) токенчлолын аргуудыг харьцуулсан. Үр дүнгээс харахад нарийн тохируулга хийсэн GPT-2 загвар нь эхнээсээ сурсан сургалтаас илүү гүйцэтгэл үзүүлсэн бөгөөд монгол хэлэнд тохируулсан холимог токенчлол ашигласан хувилбар хамгийн бага 22.5% WER үзүүлж, seq2seq суурь загвараас (24.7%) илүү үр дүнтэй байв. Судалгааны үр дүн нь урьдчилан сурсан хэлний мэдлэг болон хэлд оновчтой токенчлолын стратеги нь нөөц багатай, үг зүйн баялаг хэлний галиглалтыг хэвийн болгоход чухал нөлөөтэйг харуулж байна.**

Түлхүүр үг— Өгөгдөлд суурилсан арга, GPT-2 загвар, шуугиантай бичвэр, seq2seq загвар, бичвэр хэвийн болгох, токен, галиглагдсан бичвэр

I. Удиртгал

Сүүлийн жилүүдэд цахим харилцаа эрчимтэй хөгжихийн хэрээр латин бус бичиг хэрэглэдэг хэрэглэгчид латин үсгийг ашиглан галиглан бичих үзэгдэл түгээмэл болж байна. Энэ нь ялангуяа монгол хэлэнд тод ажиглагдаж байгаа бөгөөд олон нийтийн сүлжээ болон цахим орчинд харилцахдаа хэрэглэгчид албан ёсны кирилл бичгийн оронд латин үсгийг өргөнөөр ашиглах болсон. Энэхүү хандлага нь ялангуяа залуучуудын дунд түгээмэл бөгөөд сошиал медиа, чат, онлайн тоглоом зэрэгт латин үсгийг илүүд үздэг. Үүний шалтгаан нь цахим төхөөрөмж дээр бичихэд хялбар байдал, техникийн зарим хязгаарлалтыг даван туулах боломж, мөн гадаад хэлний нөлөө зэрэгтэй холбоотой.

Гэвч энэ өөрчлөлт нь эх хэлний бичгийн соёл, хэлний бүрэн бүтэн байдалд сөргөөр нөлөөлөх эрсдэлтэй. латин үсэгт суурилсан бичлэг цахим орчинд зонхилохын хэрээр залуу үе өөрсдийн албан ёсны бичгээр уншиж, бичих чадвараа алдах аюултай. Энэ нь зөвхөн хувь хүний бичиг үсгийн чадварт нөлөөлөөд зогсохгүй, боловсрол, уран зохиол, албан ёсны харилцаанд уламжлалт бичгийн үүрэг багасах зэрэг өргөн хүрээний асуудал үүсгэж болзошгүй. Олон нийтийн сүлжээнээс үүссэн их хэмжээний галигласан бичвэрийг үр дүнтэй хэвийн болгох нь эх хэлний боловсруулалтын олон хэрэглээнд үнэтэй өгөгдөл бий болгоно. Иймээс ийм төрлийн текстийг хэвийн болгох найдвартай аргуудыг боловсруулах нь зайлшгүй шаардлагатай юм.

Сүүлийн үеийн судалгаанууд энэ чиглэлд уламжлалт аргуудаас эхлээд дэвшилтэт гүн сургалтын аргуудыг хүртэл өргөн хүрээнд судалсан байдаг. Статистик машин орчуулгын (CMO) загварууд нь шуугиантай, алдаатай, албан бус бичвэрийг стандарт хэлбэрт шилжүүлэхэд өргөн ашиглагдаж ирсэн. Судлаачид эдгээр аргыг сошиал медиа, SMS бичвэр хэвийн болгоход амжилттай хэрэглэж, бичвэрийн чанар, уншигдах байдлыг сайжруулсан.

Aw нарын [1] SMS бичвэр хэвийнжүүлэхэд статистик загварыг ашигласан бол, Kurnia нар [2] Индонез хэлний хувьд хэвийн болгоход CMO-ыг үг болон тэмдэгтийн түвшинд ашиглаж, WER болон BLEU оноогоор суурь аргуудаас илүү үр дүн үзүүлснийг харуулсан [2]. Шуугиантай болон хэвийн бичвэрийн хос цуглуулж, CMO-д суурилсан хэвийн болгох аргыг санал болгосон [3],[4].

Сүүлийн жилүүдэд машин орчуулгад зориулан хөгжүүлсэн зарим нейрон загварууд бичвэр хэвийн болгоход өргөн хэрэглэгдэх болсон. Үүнд [5],[6],[7],[8] [9] дүгээр ажлууд орно. Lourentzou нар нь [10] үг-түвшний seq2seq загварыг ашиглаж, амжилтгүй тохиолдолд тэмдэгт-түвшний seq2seq-ийг ашиглах хоёр шатлалт аргыг санал болгосон. Мөн анхааралд суурилсан энкодер-декодер архитектур бүхий Causal Feature Extractor (CFE) нэртэй шинэ энкодерийг санал болгож, CNN болон LSTM-ээс илүү нарийвчлал, хурдан сургалт үзүүлсэн [11],[12]. Bucur нар нь [13] mBART-д суурилсан олон хэлний сошиал медиа бичвэр хэвийн болгох загварыг санал болгож, зарим дотоод үнэлгээнд сул байсан ч гадаад хэрэглээнд (sentiment analysis, MT) илүү сайн үр дүн үзүүлсэн.

Seq2seq загварууд өргөн хэрэглэгдэж байгаа хэдий ч, сүүлийн үед илүү дэвшилтэт архитектуруудыг туршиж эхэлсэн. Том хэмжээний хэлний загварууд (LLM) одоогоор өргөн ашиглагдаагүй ч GPT-3 [14], GPT-4 [15] зэрэг загваруудын боломжийг судалсан цөөн судалгаа бий. Ro нар [16] TTS системд зориулсан бичвэрийг хэвийн болгоход Transformer-т суурилсан загваруудыг харьцуулж, BERT энкодер ашигласан хоёр шатлалт загвар хамгийн сайн үр дүн үзүүлснийг харуулсан. Alam нар 10,000 өгөгдлөөр LoRA ашиглан нээлттэй эхийн

LLM-үүдийг тохируулж, галиглалтыг хэвийн болгоход өндөр амжилт үзүүлсэн [17],[18].

Энэхүү судалгаа нь шуугиантай, латин үсгээр галиглагдсан монгол бичвэрийг стандарт кирилл хэлбэрт хамгийн бага нөөц ашиглан үр дүнтэй хөрвүүлэхэд чиглэж байна. Бид тэмдэгт-түвшний seq2seq загварыг суурь болгон авч, том хэмжээний хэлний загваруудын боломжийг судалсан.

Судалгааны гол хувь нэмэр нь:

- GPT-2 зэрэг том хэмжээний нейрон загварыг нөөц багатай монгол хэлний галигласан үгийг хэвийн болгоход анхлан судалсан
- Үг болон өгүүлбэрийн түвшний өгөгдөлд суурилсан шинэ өгөгдлийн санг бүрдүүлж ашигласан.
- Монгол хэлэнд тусгайлан тохируулсан GPT-2 болон BERT-ийг хослуулсан шинэ токенжуулалтын арга санал болгосон

Өгүүллийн үлдсэн хэсэг дараах бүтэцтэй. 2-р бүлэгт ашигласан өгөгдлийн багцыг, 3-р бүлэгт загварууд болон үнэлгээний хэмжүүрүүдийг, 4-р бүлэгт санал болгосон арга, туршилтын үр дүн, шинжилгээг, 5-р бүлэгт дүгнэлтийг тус тус танилцуулна.

II. Өгөгдлийн багц

Энэхүү судалгаанд зориулан бид эхлээд олон нийтийн сүлжээнээс латин үсгээр бичигдсэн, шуугиантай ойролцоогоор 4,000 монгол өгүүлбэрийг цуглуулсан. Үүний дараа эдгээр өгүүлбэр тус бүрт харгалзах стандарт кирилл бичигтэй өгүүлбэрүүдийг гараар үүсгэсэн. Энэхүү өгөгдлийн хэмжээ нь нөөц багатай хэлний хувьд ижил төстэй судалгаанд ашиглагдсан өгөгдлүүдтэй [21], [22] дүйцэхүйц бөгөөд манай туршилтын орчинд тохиромжтой гэж үзсэн.

Өгөгдлийн багцыг улам баяжуулах зорилгоор өгүүлбэрүүдийг үг түвшинд сегментчилж, сургалтад ашиглах хос үгийн өгөгдлийн багцыг үүсгэсэн. Анхны өгүүлбэрийн түвшний өгөгдлийн багцыг өгүүлбэрийн түвшний өгөгдөл (ӨТ) гэж нэрлэн, харин үгийн түвшинд задлан гаргасан өгөгдлийг үгийн түвшний өгөгдөл (ҮТ) гэж нэрлэв. Цаашид өгүүлэлд ӨТ, ҮТ гэж товчилж хэрэглэнэ. 1-р хүснэгтэд олон нийтийн сүлжээнээс цуглуулсан үг түвшний өгөгдлийн багцын үндсэн статистик үзүүлэлтүүдийг харуулав.

Хүснэгт I. Олон нийтийн сүлжээнээс цуглуулсан үгийн түвшний өгөгдлийн багцын статистик

Тайлбар	Шуугиантай (Латин үг)	Стандарт (Кирилл үг)
Давтагдахгүй үг	11,780	8,610
Хамгийн урт үгийн урт	19	18
Үгийн дундаж урт	5.4	5.2
Нийт үг	33,620	35,510

Хэл зүйн болон галиглалтын олон янз байдлыг хангаж буй эсэхийг шалгахын тулд бид ҮТ өгөгдлийн багцыг дараах хэд хэдэн хэмжээсээр задлан шинжилсэн.

- *Үг түвшний олон янз байдал:* ҮТ өгөгдлийн багц нь 8,610 давтагдахгүй стандарт монгол үг агуулж байна. Эдгээрийн ойролцоогоор 20% нь хоёр ба түүнээс дээш өөр шуугиантай галигласан хувилбартай бөгөөд зарим үг 12 хүртэлх өөр бичлэгтэй байна. Энэ нь бодит хэрэглээнд түгээмэл ажиглагддаг галиглалтын олон хэлбэрийг өгөгдөл сайн хамарч байгааг харуулж байна [19],[20].
- *Түгээмэл хэрэглэгддэг үгсийн хамрах хүрээ:* Үгийн сангийн хамрах хүрээг үнэлэхийн тулд бид ҮТ өгөгдлийг урлаг, соёл, эдийн засаг, эрүүл мэнд, хууль, улс төр, спорт, технологи, боловсрол, байгаль орчин гэсэн есөн салбарыг хамарсан 75,000 мэдээний өгүүллээс бүрдсэн, 25 сая үгтэй корпус [23]-тай харьцуулсан. Үр дүнгээс харахад уг корпус дахь хамгийн түгээмэл 2,650 үгийн 70.4% нь ҮТ өгөгдлийн багцад багтсан байсан нь энэхүү өгөгдөл түгээмэл хэрэглэгддэг үгсийг хангалттай хамарч байгааг нотолж байна.

Эдгээр шинжилгээнүүдээс харахад ҮТ өгөгдлийн багц нь үг, үгийн сангийн түвшинд олон янз байдалтай бөгөөд галиглалтыг хэвийн болгох болон бусад монгол хэлний хувьд эх хэлний боловсруулалтын даалгавруудад ашиглахад үнэ цэнтэй нөөц болох нь батлагдаж байна. Нэмэлтээр, бид олон нийтийн сүлжээнээс 300 шуугиантай галиглагдсан өгүүлбэрийг тестийн зорилгоор цуглуулсан. Эдгээр өгүүлбэр нь монгол хэлийг латин үсгээр илэрхийлэх хэрэглэгчдийн олон янзын стандарт бус бичлэгийг агуулсан. Уг тестийн өгөгдлийн дэлгэрэнгүй статистикийг 2-р хүснэгтэд үзүүлэв.

Хүснэгт II. Тестийн өгөгдлийн дэлгэрэнгүй статистик

Тайлбар	Шуугиантай (Латин үг)
Нийт үг	2,520
Давтагдахгүй үг	1,760
Хамгийн урт үгийн урт	17
Үгийн дундаж урт	5.3
Үгийн санд багтсан (**IV**)	1,630 (64.68%)
Үгийн санд багтаагүй (**OOV**)	890 (35.32%)
Нийт үг	**2,520**

III. Арга зүй

Энэ хэсэгт бид суурь загвар болгон ашигласан seq2seq загвар, мөн GPT-2 загварт суурилсан бидний санал болгож буй аргыг танилцуулна. Үүнээс гадна загваруудын гүйцэтгэлийг үнэлэхэд ашигласан үнэлгээний аргуудыг тайлбарлана. Дараах дэд хэсгүүдэд арга тус бүр болон үнэлгээний стратегийг дэлгэрэнгүй авч үзнэ.

A. SEQ2SEQ загвар

Seq2seq загвар нь энкодер–декодер архитектурт суурилдаг бөгөөд оролтын бичвэрийг дараалал хэлбэрээр боловсруулж, түүнд харгалзах гаралтын дарааллыг үүсгэдэг тул манай зорилгод маш тохиромжтой. Иймээс бид латин бичигт суурилсан

шуугиантай транслитерацийг стандарт кирилл бичвэрт хэвийн болгох суурь арга болгон [7]-д санал болгосон энкодер–декодер архитектур дээр үндэслэж, [8]-д нэвтрүүлсэн анхаарал (attention) механизмтай тэмдэгт-түвшний seq2seq загварыг ашигласан.

Янз бүрийн давхаргын тоо бүхий, өөр өөр рекуррент архитектурууд дээр туршилт хийж, холбогдох гиперпараметрүүдийг тус тусад нь тохируулсны үр дүнд бид GRU-д суурилсан тэмдэгт-түвшний загварыг сонгосон. GRU загвар нь гүйцэтгэл болон тооцооллын үр ашиг хоёрын хооронд хамгийн сайн тэнцвэрийг хангаж байсан бөгөөд ялангуяа нөөц хязгаарлагдмал нөхцөлд давуу талтай байв. Түүнчлэн үгийн санд багтаагүй (OOV) токенуудыг боловсруулах чадвар болон бага хэмжээний өгөгдөл дээр эхнээс нь үр дүнтэй сургах боломж нь энэхүү загварыг манай хэвийн болгох даалгаварт тохиромжтой болгосон.

Загварыг TensorFlow [26] ашиглан хэрэгжүүлсэн бөгөөд энкодер болон декодер нь тус бүр нэг давхаргатай, 512 далд нэгж бүхий GRU-ээс бүрдэнэ. Embedding векторын хэмжээг 250 болгон тохируулсан. Сургалтад Adam оптимизаторыг 0.0002 суралцах хурдтай, 32-ын багцын хэмжээтэйгээр ашигласан.

B. GPT-2 загвар

Seq2seq загвараас гадна бид бичвэрийг хэвийн болгох зорилгоор GPT-2 загварыг туршсан. GPT-2 нь autoregressive шинжтэй, Transformer-д суурилсан хэлний загвар бөгөөд байгалийн хэл боловсруулах олон даалгаварт өндөр гүйцэтгэл үзүүлдгээрээ алдартай [27, 28]. Энкодер-декодер бүтэцтэй seq2seq загвараас ялгаатай нь GPT-2 нь өмнөх контекстэд тулгуурлан, дараагийн токеныг таамагладаг бөгөөд self-attention механизмаар дамжуулан урт зайн хамаарлыг RNN-д суурилсан загваруудаас илүү үр дүнтэй барьж чаддаг.

Манай туршилтад бид тооцооллын нөөц хязгаарлагдмал нөхцөлд нарийн тохиргоо (fine-tuning) болон ашиглалтад илүү тохиромжтой гэж үзэн GPT-2-Small загварыг үндсэн сонголтоор авсан. Энэхүү сонголтыг баталгаажуулахын тулд бид GPT-2-Medium [29], GPT-2-Large [30] зэрэг том хувилбарууд, мөн нээлттэй эхийн олон хэлний Transformer загварууд болох mGPT-1.3B [31], BLOOM-560M [32], LLaMA 3.2-1B [33]-тай харьцуулсан туршилт хийсэн. Ингэснээр GPT-2-ын илүү том хувилбаруудын гүйцэтгэл болон олон хэлний урьдчилсан сургалтын нөлөөг судлахыг зорьсон.

mGPT-1.3B загвар нь хязгаарлагдмал хэмжээний монгол бичвэрийг багтаасан олон хэлний өгөгдөл дээр сургагдсан тул түүний олон хэлний мэдлэг манай даалгаварт эерэг нөлөө үзүүлэх эсэхийг шалгах зорилгоор нарийн тохируулсан. BLOOM-560M нь монгол хэл дээр сургагдаагүй боловч олон төрлийн хэлний өгөгдөл дээр бэлтгэгдсэн загвар юм. Иймээс BLOOM-ийг тохируулан ашигласнаар нөөц багатай монгол хэл дээр том хэмжээний олон хэлний загвар хэрхэн ажиллахыг үнэлсэн. Үүнтэй адил, олон хэлний корпус дээр, гэхдээ монгол хэлний өгөгдөл хязгаарлагдмал нөхцөлд урьдчилан сургагдсан LLaMA 3.2-1B загварыг манай даалгаварт дасан зохицох чадварыг нь үнэлэх зорилгоор туршсан.

Туршилтын үр дүнгээс харахад GPT-2-Medium болон GPT-2-Large нь GPT-2-Small-тай харьцуулахад бага зэрэг илүү гүйцэтгэл үзүүлсэн боловч тооцооллын зардал мэдэгдэхүйц өсөхтэй харьцуулахад уг сайжруулалт хангалттай биш байв. BLOOM-560M болон LLaMA 3.2-1B нь тодорхой хэмжээнд хэвийнжүүлэлт хийх чадвар харуулсан ч практик хэрэглээнд хангалттай үр дүн үзүүлээгүй. Харин mGPT-1.3B загвар нь олон хэлний урьдчилсан сургалттай байсан ч манай даалгаврыг үр дүнтэй гүйцэтгэж чадалгүй, галиглалтыг хэвийн болгох даалгаварт хүндрэлтэй байгааг харуулсан.

Эдгээр үр дүн нь GPT-2-Small загварыг нарийвчлал болон үр ашгийн хувьд хамгийн оновчтой сонголт болохыг баталгаажуулсан. Иймээс бид Hugging Face-ийн Transformers сангаас авсан урьдчилан сургагдсан GPT-2-Small загварыг ашиглаж, анхны архитектур болон гиперпараметрүүдэд өөрчлөлт оруулалгүйгээр нарийн тохируулсан [34].

C. Үнэлгээ

Загваруудын гүйцэтгэлийг үнэлэхийн тулд бид Үгийн алдааны хувь (Word Error Rate, WER) хэмжүүрийг ашигласан. WER нь тестийн өгөгдөл дэх нийт үгсийн тоотой харьцуулахад буруу таамаглагдсан үгсийн эзлэх хувийг хэмждэг. Уг судалгаанд WER-ийг дараах байдлаар тооцоолсон:

$$WER = \frac{\text{Буруу таамагласан үг тоо}}{\text{Нийт үгсийн тоо}} x100 \quad (1)$$

Энд “буруу үг” гэдэг нь загвараас гарсан үр дүн тухайн лавлах (ground truth) үгтэй яг таарахгүй байгаа аливаа үгийг хэлнэ.

IV. Туршилт ба үр дүн

Энэ бүлэгт бид дуу чимээтэй латин бичвэрээс стандарт кирилл бичиг рүү монгол хэлний бичвэрийг нормчлох бага нөөцтэй орчны асуудлыг шийдэхэд чиглэсэн санал болгож буй аргачлалаа танилцуулна. Бид асуудлын өөр өөр талыг хамарсан, хоорондоо уялдаа бүхий 2 стратегийг дэвшүүлж байна. Эдгээр аргуудын хүрээнд бид загварын хувилбарууд, өгөгдлийн нарийвчлал, алдаа засварлах механизмуудын нөлөөг системтэйгээр үнэлж, тэдгээр нь галиглалтийг хэвийн болгох нарийвчлалд хэрхэн нөлөөлж байгааг шинжилсэн. Туршилтын үр дүнгээс харахад эдгээр стратегиуд нь шошгологдсон өгөгдөл хомс зэрэг бага нөөцтэй орчны үндсэн сорилтуудыг хамтад нь шийдвэрлэж, нөөц багатай хэлнүүдэд ижил төстэй даалгаврууд хувьд дахин давтагдахуйц суурь жишгийг тогтоож байна.

Санал болгож буй арга: Загвар сургалтын стратеги

Бид GPT-2 авто-регрессив хэлний загварыг контекстэд нийцсэн дараалал үүсгэх чадварынх нь улмаас сонгон ашигласан. Бага нөөцтэй орчинд үгийг сайжруулахын тулд бид хоёр үндсэн стратегийг системтэйгээр судалсан. Үүнд: **загварын эхлэлийн жингийн тохиргооны стратеги** болон **токенчлолын стртатегиуд** болно. Эдгээр туршилтууд нь урьдчилан сурсан мэдлэг болон токенчлолын өөрчлөлтүүд галиглалтыг хэвийн болгох нарийвчлалд хэрхэн нөлөөлж байгааг үнэлэх зорилготой бөгөөд ингэснээр шуугиантай галигийг боловсруулахад хамгийн үр дүнтэй хандлагыг тодорхойлох боломжийг олгоно.

А. Загварын эхлэлийн жингийн тохиргооны стратеги
Галиглагдсан бичвэрийг хэвийн болгох даалгаварт урьдчилсан мэдлэгийн үүргийг судлахын тулд бид дараах хоёр өөр сургалтын аргыг авч үзсэн. Үүнд:

- **Нарийн тохируулан сургасан (Fine-tune) GPT-2**: Загварыг урьдчилан сурсан GPT-2 жингүүдээр (124M параметр) эхлэлийн жингийн тохиргоо хийж, анхны токен embedding-ийг хадгалсан хэвээр өгөгдөл дээрээ тохируулсан. Энэ арга нь өндөр нөөцтэй англи хэл дээр суралцсан загварын хэл шинжлэлийн ерөнхий мэдлэгийг ашиглах боломжийг олгодог.

- **Эхнээс нь (scratch) сургасан GPT-2**: Загварын бүх жинг санамсаргүйгээр дахин тохиргоо хийж, ижил архитектур ашиглан эхнээс нь сургасан. Ингэснээр урьдчилсан сургалтын өгөгдлөөс үүдэх хэл хоорондын хазайлтыг арилгаж, загвар зөвхөн Монгол хэлний галиглалтийн хэв маягаас суралцах боломж бүрдэнэ.

Эдгээр хоёр аргыг харьцуулснаар урьдчилан сурсан төлөөлөл гүйцэтгэлийг сайжруулж байна уу, эсвэл тухайн даалгаварт саад болж байна уу гэдгийг үнэлсэн.

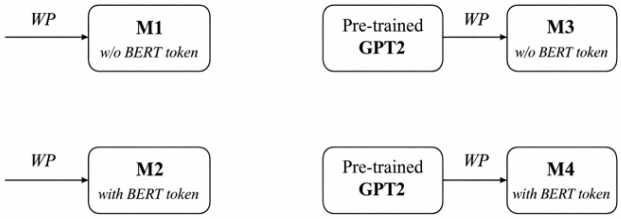
В. Токенчлолын стратеги
Токенчлол нь ялангуяа үг зүй баялаг хэлнүүдийн хувьд бага нөөцтэй үед маш чухал үүрэгтэй. Үүний нөлөөг үнэлэхийн тулд бид дараах хоёр өөр токенчлолын аргыг туршсан.

- **GPT-2-ийн анхны токенчлол**: GPT-2-ийн үндсэн byte-pair encoding (BPE) токенизаторыг өөрчлөлтгүй ашигласан. Гэвч энэ нь голчлон англи хэл дээр сургагдсан тул монгол хэлний тусгай дэд нэгжүүдийг агуулдаггүй бөгөөд үүний улмаас үгийн хуваалт оновчгүй, галигийн алдаа гарах магадлалтай.

- **Холимог (Hybrid) токенчлол**: Монгол хэлний корпус дээр сургагдсан BERT загвараас [35] авсан монгол дэд нэгжүүдийг GPT-2-ийн BPE үгийн санд нэгтгэсэн. Энэ өөрчлөлт нь “-ын”, “-ийн” зэрэг дагаврууд таних боломж олгож, үгийг зөв хуваах, дахин бүтээх чадварыг сайжруулна.

Эдгээр эхлэлийн жингийн тохиргооны болон токенчлолын стратегиудыг системтэйгээр харьцуулснаар бид монгол хэлний галиглагдсан үгийг хэвийн болгох даалгаврын хувьд хамгийн үр дүнтэй тохиргоог тодорхойлохыг зорьсон. Үүний тулд бид нэг ижил өгөгдөл (YT) ашиглан дараах дөрвөн загварын хувилбарыг сургасан.

Үүнд:
M1 – эхнээс нь сургасан *GPT-2* + анхны токенчлол,
M2 – эхнээс нь сургасан *GPT-2* + холимог токенчлол,
M3 – нарийн тохируулан сургасан *GPT-2* + анхны токенчлол,
M4 – нарийн тохируулан сургасан *GPT-2* + холимог токенчлол.
Эдгээр дөрвөн загварын сургалтын ерөнхий бүтцийг Зураг 1-д харуулав. Энэхүү туршилтын зохион байгуулалт нь (i) урьдчилан сурсан мэдлэг галиглагдсан үгийг хэвийн болгох даалгаварт ашиг тустай эсэх, (ii) өөр өөр токенчлолын схемүүд дарааллын хуваалт болон нарийвчлалд хэрхэн нөлөөлж байгааг үнэлэх боломж олгож байна.



Зураг 1. Эхлэлийн жингийн оноолт ба токенчлолын ялгаатай стратегид суурилсан загварууд

Бүх загваруудыг ижил нөхцөлөөр сургасан бөгөөд эхлэлийн жингийн тохиргоо ба токенчлолын нөлөөг тусгаарлан судалсан. Галиглагдсан үгийг хэвийн болгох даалгаврыг тодорхой болгохын тулд оролтын дарааллыг дараах хэлбэрээр тэмдэглэв:

[START] noisy: <Latin бичвэр>
[SEP] canon: <Cyrillic бичвэр>
[EOS]
Бүх загварууд ижил гиперпараметртайгаар (*суралцах хурд: 5e-5, batch size: 8, epoch: 100*) сургагдсан. 3-р хүснэгтэд дөрвөн загварын тестийн өгөгдөл дээрх гүйцэтгэлийг харуулав.

ХҮСНЭГТ III. ТЕСТИЙН ӨГӨГДӨЛ ДЭЭРХ ЗАГВАРУУДЫН ГҮЙЦЭТГЭЛ

Загварууд	**Эхлэлийн жин**	**Токенчилол**	**WER**
Seq2Seq	-	-	24.7
M1	Эхнээс	Original GPT-2	25.8
M2	Эхнээс	GPT-2 + BERT Tokens	25.2
M3	Тохируулсан	Original GPT-2	24.6
M4	Тохируулсан	GPT-2 + BERT Tokens	**22.5**

С. Үр дүнгийн шинжилгээ
Туршилтын үр дүнгээс харахад эхнээс нь сургасан загварууд (M1, M2) Seq2Seq суурь загвараас давж чадсангүй, харин нарийн тохируулан сургасан загварууд (M3, M4) илүү гүйцэтгэл үзүүлсэн. Ялангуяа холимог токенчлолтой нарийн тохируулан сургасан загвар (M4) хамгийн бага WER-тэй байв.
Эдгээр үр дүн нь:

- **Нарийн тохируулан сургах нь эхнээсээ сургахаас илүү үр дүнтэй** болохыг,

- **Монгол хэлэнд тохируулсан токенчлол алдааг мэдэгдэхүйц бууруулдаг** болохыг,
- Үг зүйн баялаг хувилалтай, бага нөөцтэй хэлэнд **токенчлолын стратеги онцгой ач холбогдолтой** болохыг харуулж байна.

V. ХЯЗГААРЛАЛТ

Энэхүү судалгааны хязгаарлалтууд дараах байдлаар тодорхойлогдоно.

- ***Нэгдүгээрт,*** ашигласан өгөгдлийн хэмжээ харьцангуй бага (4,000 өгүүлбэр) тул загварын ерөнхийшүүлэх чадварт нөлөөлөх боломжтой.
- ***Хоёрдугаарт,*** синтетик өгөгдөл ашиглаагүй нь галиглалын олон янз хувилбарыг бүрэн хамрахад саад болж байж болзошгүй.
- ***Гуравдугаарт,*** өгөгдөл нь голчлон сошиал медиа орчноос цуглуулсан тул домэйны хувьд хязгаарлагдмал бөгөөд бусад төрлийн бичвэрт (жишээлбэл, албан бичиг, хэвлэл мэдээлэл) шууд хэрэглэхэд гүйцэтгэл буурах магадлалтай.

VI. ДҮГНЭЛТ

Энэхүү судалгаагаар нөөц багатай монгол хэлний шуугиантай латин галиглалтыг стандарт кирилл бичигт хэвийн болгох асуудлыг том хэмжээний хэлний загварт суурилан судаллаа. Бид тэмдэгт-түвшний seq2seq загварыг суурь болгон авч, GPT-2 загварыг өөр өөр сургалтын стратеги болон токенчлолын хувилбаруудаар туршин харьцуулсан.

Туршилтын үр дүнгээс дараах үндсэн дүгнэлт гарч байна:

- Урьдчилан сурсан жинг ашиглан, нарийн тохируулан сургах нь эхнээсээ сургах аргаас илүү үр дүнтэй байна.
- Монгол хэлэнд тохируулсан холимог токенчлол нь анхны GPT-2 токенчлолоос илүү гүйцэтгэл үзүүлж, морфологи баялаг бүтэцтэй хэлний хувьд токенчлолын ач холбогдол өндөр болохыг баталлаа.
- Нарийн тохируулга + холимог токенчлол (M4) загвар хамгийн сайн үр дүн үзүүлж, 22.5% WER-т хүрсэн нь суурь seq2seq загвараас илүү гүйцэтгэлтэй байв.
- Нөөц багатай орчинд ч том хэмжээний хэлний загваруудыг оновчтой тохируулснаар практик хэрэглээнд ашиглах боломжтойг харууллаа.

Энэхүү судалгаа нь монгол хэлний галиглагдсан үгийг хэвийн болгох даалгаварт том хэмжээний хэлний загварыг ашигласан анхны системтэй туршилтуудын нэг бөгөөд цаашид синтетик өгөгдөл нэмэгдүүлэх, зааврын дагуу нарийн тохируулан (instruction-based fine-tuning) сургах, эсвэл параметр багатай тохируулгын аргууд (жишээ нь LoRA) ашиглан гүйцэтгэлийг нэмэгдүүлэх боломжтой гэж үзэж байна. Мөн энэхүү аргачлал нь бусад нөөц багатай хэлнүүдийн галигласан бичвэрийг хэвийн болгох даалгаварт өргөтгөн хэрэглэх боломжтой юм.

НОМ ЗҮЙ

[1] A. Aw, M. Zhang, J. Xiao, and J. Su, “A phrase-based statistical model for SMS text normalization,” *in Proc. COLING/ACL 2006 Main Conf. Poster Sessions*, 2006, pp. 33–40.

[2] A. Kurnia, A. Ajmal, and E. Yulianti, “Statistical machine translation approach for lexical normalization on Indonesian text,” *in Proc. 2020 Int. Conf. Asian Lang. Process. (IALP)*, 2020, pp. 288–293.

[3] T. Schlippe, C. Zhu, D. Lemcke, and T. Schultz, “Statistical machine translation based text normalization with crowdsourcing,” *in Proc. 2013 IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2013, pp. 8406–8410.

[4] K. Scannell, “Statistical models for text normalization and machine translation,” *in Proc. 1st Celtic Lang. Technol. Workshop*, 2014, pp. 33–40.

[5] N. Kalchbrenner and P. Blunsom, "Recurrent continuous translation models," in *Proc. 2013 Conf. Empirical Methods Natural Lang. Process.*, Seattle, WA, USA, 2013, pp. 1700–1709.

[6] K. Cho, B. van Merriënboer, Ç. Gülçehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," in *Proc. 2014 Conf. Empirical Methods Natural Lang. Process.*, Doha, Qatar, 2014, pp. 1724–1734.

[7] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proc. 27th Int. Conf. Neural Inf. Process. Syst.*, Montreal, Canada, 2014, pp. 3104–3112.

[8] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in *Proc. 3rd Int. Conf. Learn. Represent.*, San Diego, CA, USA, 2015.

[9] M. Luong, H. Pham, and C. D. Manning, "Effective approaches to attention-based neural machine translation," in *Proc. 2015 Conf. Empirical Methods Natural Lang. Process.*, Lisbon, Portugal, 2015, pp. 1412–1421.

[10] I. Lourentzou, K. Manghnani, and C. Zhai, “Adapting sequence-to-sequence models for text normalization in social media,” *in Proc. 13th Int. AAAI Conf. Web Soc. Media*, 2019, pp. 335–345.

[11] A. Javaloy and G. García-Mateos, “Text normalization using encoder–decoder networks based on the causal feature extractor,” *Appl. Sci.*, vol. 10, no. 13, p. 4551, 2020.

[12] R. Satapathy, Y. Li, S. Cavallari, and E. Cambria, “Seq2Seq deep learning models for microtext normalization,” *in Proc. 2019 Int. Joint Conf. Neural Netw. (IJCNN)*, Budapest, Hungary, 2019, pp. 1–8.

[13] A.-M. Bucur, A. Cosma, and L. P. Dinu, “Sequence-to-sequence lexical normalization with multilingual transformers,” *arXiv preprint arXiv:2110.02869*, 2021.

[14] T. B. Brown, B. Mann, N. Ryder, et al., "Language Models are Few-Shot Learners," *Advances in Neural Information Processing Systems*, vol. 33, pp. 1877–1901, 2020.

[15] OpenAI, "GPT-4 Technical Report," *arXiv preprint arXiv:2303.08774*, 2023.

[16] J. H. Ro, F. Stahlberg, K. Wu, and S. Kumar, “Transformer-based models of text normalization for speech applications,” *arXiv preprint arXiv:2202.00153*, 2022.

[17] Y. Zhang, T. M. Bartley, M. Graterol-Fuenmayor, V. Lavrukhin, E. Bakhturina, and B. Ginsburg, “A chat about boring problems: Studying GPT-based text normalization,” *in Proc. ICASSP 2024 - 2024 IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Seoul, South Korea, 2024, pp. 10921–10925.

[18] M. M. I. Alam and A. Anastasopoulos, “Large language models as a normalizer for transliteration and dialectal translation,” *in Proc. 12th Workshop NLP Similar Lang., Varieties Dialects*, 2025, pp. 39–67.

[19] Mongolian Agency for Standardization and Metrology, “MNS 5217:2003 - Rules of Mongolian Script Spelling,” Mongolian national standard, 2003.

[20] Mongolian Agency for Standardization and Metrology, “MNS 5217:2012 - Rules of Mongolian Script Spelling,” Mongolian national standard, 2012.

[21] S. Lusito, E. Ferrante, and J. Maillard, "Text normalization for low-resource languages: the case of Ligurian," in *Proc. ComputEL-4:Workshop on Computational Methods for Endangered Languages*,16 June 2022. [Online]. Available: https://computel-workshop.org/8 VOLUME XX, 2017

[22] N. T. Le, F. Sadat, L. Menard, and D. Dinh, "Low-Resource Machine Transliteration Using Recurrent Neural Networks," ACM Transactions on Asian and Low-Resource Language Information Processing, vol. 18, no. 2, pp. 1–13, Jan. 2019.
[23] T. Tuguldur, "Eduge news classification dataset," GitHub, [Online].Available: https://github.com/tugstugi/mongolian-nlp/blob/master/datasets/eduge.csv.gz [Accessed: May 12, 2025].
[24] Ch. Choimaa, Grammar Dictionary of the Mongolian Language. Vol.1 (A-C), Ulaanbaatar: National University of Mongolia Press, 2005.
[25] Ch. Choimaa, Grammar Dictionary of the Mongolian Language. Vol.2 (Т-Я), Ulaanbaatar: National University of Mongolia Press, 2005.
[26] M. Abadi et al., "TensorFlow: A system for large-scale machine learning," in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI '16)*, 2016, pp. 265–283.
[27] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, “Attention is all you need,” in *Proceedings of the 31st International Conference on Neural Information Processing Systems (NeurIPS)*, Long Beach, CA, USA, 2017, pp. 5998–6008.
[28] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, “Language models are unsupervised multitask learners,” OpenAI, Tech. Rep., 2019.
[29] OpenAI, “GPT-medium model,” OpenAI, [Online]. Available: https://openai.com.
[30] OpenAI, “GPT-Large model,” OpenAI, [Online]. Available: https://openai.com.
[31] A. Shliazhko, A. Arkhipov, and D. Emelyanov, “mGPT: A multilingual GPT model,” *arXiv preprint arXiv:2110.02702*, 2021.
[32] J. BigScience, T. Scialom, A. Wang, et al., “BLOOM: A 176B-parameter open-access multilingual language model,” *arXiv preprint arXiv:2211.05100*, 2022.
[33] Meta AI, “LLaMA 3.2–1B: Open Foundation Language Model,”Hugging Face, 2024. [Online]. Available:https://huggingface.co/meta-llama/Meta-Llama-3-1B
[34] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, and J. Brew, “Transformers: State-of-the-art natural language processing,” in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, 2020, pp. 38–45.
[35] Tugstugi, “BERT Base Mongolian Uncased,” Pre-trained model. Accessed [2025.03.22]. [Online]. Available: https://huggingface.co/tugstugi/bert-base-mongolian-uncased

Munkhnote: A Modular AI-Assisted Meeting Minutes System for Mongolian with Improved Workflow Efficiency

Magvan-Erdene.G
Department of Computer Science
SICT, MUST
Ulaanbaatar, Mongolia
magvan-erdene@must.edu.mn

Munkhtuya.U
Department of Information Technology
SICT, MUST
Ulaanbaatar, Mongolia
munkhtuyauurtsaikh@gmail.com

Ganchimeg.G
Department of Computer Science
SICT, MUST
Ulaanbaatar, Mongolia
ganaa@must.edu.mn

***Abstract*—Taking accurate meeting minutes from long audio and video recordings is a time-consuming process involving frequent media navigation, task switching, and manual transcription. This challenge is particularly pronounced for low-resource languages such as Mongolian, where mature commercial transcription tools are limited. This paper presents Munkhnote, a modular AI-assisted meeting minutes system designed to improve workflow efficiency for Mongolian-language meetings. The system integrates a keyboard-first media player, a rich-text note editor with offline persistence, and an asynchronous AI pipeline supporting automatic speech recognition, speaker diarization, and LLM-based summarization. The backend uses Django REST Framework with Celery-based task orchestration across CPU and GPU workers, with a modular service architecture that lets individual AI components be replaced or upgraded independently. We evaluate the system through both component-level experiments and a controlled workflow efficiency study on Mongolian meeting recordings. Fine-tuning reduces speaker diarization error rate from 22% to 21% and ASR word error rate from 57% to 25%. In practice, ASR performance continues to improve, and the system can also leverage external services such as the Chimege API to ensure more reliable transcription in real-world scenarios. In a single-operator self-experiment, Munkhnote reduces minute-taking time by approximately 48% and mouse interactions by over 93% compared to a fragmented baseline. These results show that integrating AI-assisted processing with optimized interaction design can significantly improve meeting documentation workflows in low-resource language settings. Unlike prior meeting transcription systems, our contribution lies not in new model architectures but in the integration of multiple AI components into a unified workflow optimized for a low-resource, code-switching language setting, emphasizing (1) keyboard-first interaction design, (2) modular and replaceable AI services, and (3) quantitative measurement of workflow efficiency. A public demo and source code are available to support reproducibility at https://github.com/neith-san/munkh-note-demo.**

***Index Terms*—meeting minutes, automatic speech recognition, speaker diarization, large language models, low-resource languages, Mongolian language, workflow efficiency**

I. Introduction

Meeting minutes are critical records of organizational decisions, action items, and discussions. Preparing accurate minutes from recorded meetings remains a labor-intensive process, requiring the note-taker to repeatedly scrub through audio or video, identify speakers, transcribe key points, and organize the output into a coherent document. Studies indicate that professionals spend 20–40% of their meeting-related time on post-meeting documentation [1].

The challenge is particularly acute for **low-resource languages** such as Mongolian. While commercial transcription services (e.g., Otter.ai, Microsoft Teams transcription) have matured for English, Mongolian-language meetings lack equivalent tooling. The Chimege ASR system [2] is one of the few Mongolian-focused speech recognition services, but it operates as a standalone API without integrated minute-taking workflow support.

Three key bottlenecks in the current meeting minutes workflow are identified:

1) **Media navigation friction**: Standard media players rely on mouse-driven scrubbing, which disrupts the typing workflow. Professionals report that switching between keyboard (for notes) and mouse (for playback) accounts for significant time overhead.
2) **Tool fragmentation**: Users must switch between a media player, a text editor, and potentially separate ASR/diarization tools, leading to context-switching costs.
3) **Lack of Mongolian-language AI integration**: Existing meeting transcription systems do not support Mongolian ASR or are not designed for the mixed-language (Mongolian–English code-switching) patterns common in Mongolian professional settings.

Unlike existing meeting transcription systems, **Munkhnote** integrates keyboard-first media control, modular AI orchestration, and Mongolian-language support into a single unified workflow.

This paper makes the following contributions:

- We design and implement **Munkhnote**, a modular meeting minutes system tailored for low-resource languages, with a keyboard-first interaction model.
- We propose a unified asynchronous AI pipeline integrating ASR, speaker diarization, and LLM-based summa-

rization, with interchangeable service modules.
- We conduct a component-level evaluation on Mongolian meeting data, demonstrating improvements in diarization (DER reduced from 22% to 21%) and ASR (WER reduced from 57% to 25% after fine-tuning).
- We conduct a single-operator self-experiment demonstrating that **Munkhnote** reduces minute-taking time by approximately 48% and mouse interactions by over 93% compared to a fragmented baseline.

Unlike prior meeting transcription systems, our contribution lies not in new model architectures but in the integration of multiple AI components into a unified workflow optimized for a low-resource, code-switching language setting. In particular, we emphasize (1) keyboard-first interaction design, (2) modular and replaceable AI services, and (3) quantitative measurement of workflow efficiency, which is rarely reported in prior work.

II. Related Work

A. Meeting Transcription and Minute Systems

Automated meeting transcription has been studied extensively. The AMI Meeting Corpus [3] and ICSI Meeting Corpus [4] established benchmarks for meeting speech recognition and summarization. Commercial systems such as Otter.ai, Microsoft Teams, and Google Meet now offer real-time transcription for English, but these are closed-source and primarily serve English-language users. Open-source alternatives like Whisper [5] provide multilingual ASR capabilities but lack integrated minute-taking workflows.

B. Speaker Diarization

Speaker diarization—the task of determining "who spoke when"—is essential for multi-speaker meeting transcription. The Pyannote framework [6] has become a widely used open-source solution, achieving state-of-the-art results on standard benchmarks. Pyannote 3.1 builds on an end-to-end neural pipeline combining voice activity detection, speaker segmentation, and embedding-based clustering, with improved segmentation models that are amenable to domain-specific fine-tuning. Recent work has explored fine-tuning diarization models for domain-specific audio [7], which is relevant for adapting to Mongolian meeting acoustics.

C. ASR for Low-Resource Languages

Mongolian is classified as a low-resource language for speech technology [8]. The Chimege system [2] is a Mongolian-focused ASR service developed by a domestic company. OpenAI's Whisper [5] supports Mongolian through its multilingual training data, though its accuracy on domain-specific Mongolian content (e.g., technical meetings, government proceedings) has not been extensively evaluated. Recent work on adapting large multilingual models to low-resource languages through fine-tuning [9] suggests a promising direction.

D. Meeting Summarization with LLMs

Large language models (LLMs) have shown strong performance on meeting summarization tasks [10]. Approaches range from prompt-based summarization using general-purpose models (e.g., GPT-4, Gemini) to fine-tuned models that produce structured outputs with agenda items, decisions, and action items [11]. The challenge for Mongolian-language summarization is the limited availability of training data and evaluation benchmarks.

This paper is primarily a **systems and HCI contribution**: the novelty lies not in new model architectures, but in how existing AI components are integrated into a unified workflow for a low-resource, code-switching language setting. Unlike prior meeting transcription systems, our approach emphasizes three aspects: (1) a keyboard-first interaction design that minimizes context switching during transcription; (2) a modular AI pipeline with interchangeable services (ASR, diarization, summarization) that can be independently upgraded; and (3) a quantitative evaluation of workflow efficiency, including statistically significant reductions in time and interaction cost, which is rarely reported in prior work. This distinguishes our work from prior ASR-focused or summarization-focused systems that do not consider end-to-end user workflow.

III. System Design

A. Overview

Munkhnote follows a three-panel desktop layout designed to minimize context switching during meeting-minute preparation (Fig. 1): a **left panel** with a media player (customizable keyboard shortcuts, adjustable skip intervals), a **center panel** hosting a Quill 2.x rich-text editor (autosave, server persistence, offline draft recovery), and a **right panel** exposing controls for AI jobs (trim, ASR, diarization, full-pipeline text generation) and their results. On narrow screens the layout collapses into a tabbed interface with no loss of functionality.

B. Backend Architecture

The backend is implemented using Django and the Django REST Framework (DRF) and organized into two Django applications:

Henosis serves as the primary business-logic application, containing four data models:

- `ShortcutBinding`: Per-user customizable keyboard shortcuts for seven media actions (play/pause, skip back, skip forward, mute, volume up/down).
- `Note`: Rich-text notes stored as Quill Delta JSON with user association and timestamps.
- `MediaSource`: References to uploaded files or external URLs for the media player.
- `VizJob`: Asynchronous job records tracking type, status (pending $\rightarrow$ running $\rightarrow$ success/failure), arguments, results, and errors.

Workers is a separate Django application containing Celery task definitions and a service layer. Each Celery task follows a uniform pattern: load the `VizJob` record, resolve the audio

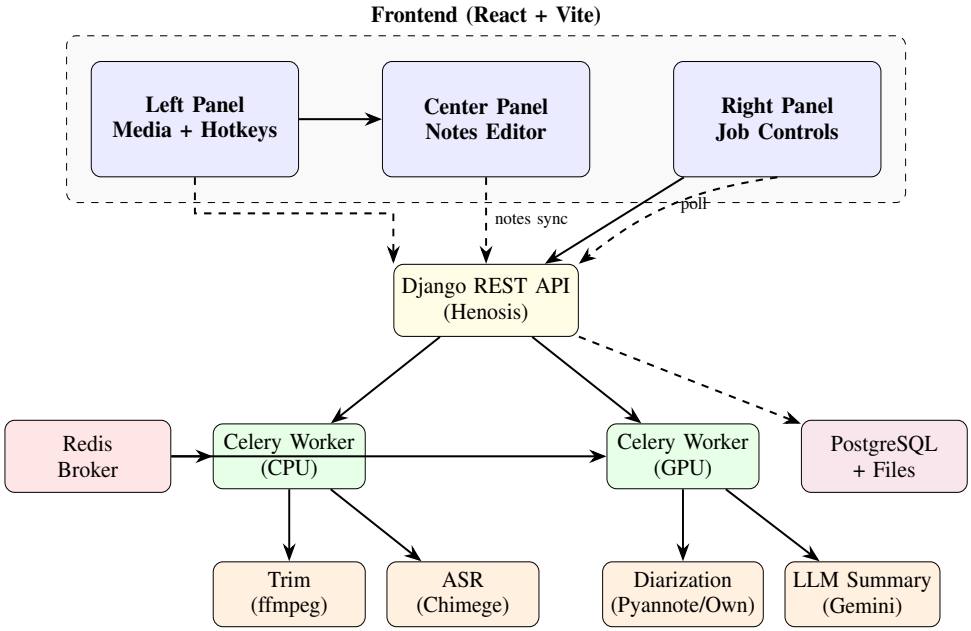


Fig. 1. System architecture of Munkhnote. The frontend comprises three panels communicating with a Django REST API. Asynchronous jobs are dispatched via Redis to CPU or GPU Celery workers, which invoke modular service backends. Dashed arrows indicate asynchronous or polling-based communication.

path, invoke the corresponding service function, and persist the result. This separation ensures that task orchestration logic is decoupled from model inference and API call logic.

C. Asynchronous Job Pipeline

All AI processing in Munkhnote is asynchronous. The frontend submits a job to `POST /api/viz-jobs/start/`, which creates a `VizJob` record in `pending` status, enqueues the matching Celery task, and returns immediately; the worker transitions the job through `running` to `success` or `failure` as it invokes the service module, while the frontend polls `GET /api/viz-jobs/:id/` until the terminal state. Results — transcript, diarization segments, summary — are then displayed in the right panel and can be inserted into the note editor.

The full-pipeline "generate text" job chains four stages: (1) optional trimming, (2) ASR, (3) diarization, (4) LLM summarization. If diarization fails, the pipeline continues with a partial-success status so transcript and summary remain available; this graceful degradation is load-bearing for production reliability.

D. Modular Service Interface

Each AI capability lives in its own Python module under `workers/services/` behind a uniform single-function interface: `trim` (ffmpeg-based audio segmentation), `chimege` (Mongolian ASR via the Chimege API), `diarization_pyannote` (Pyannote 3.1 speaker labels), and `diarization_own` (a custom diarization backend implementing the same signature). Backends can be swapped or A/B-tested without touching task orchestration; GPU-bound services are routed to a dedicated Celery queue consumed by a CUDA-enabled worker.

E. Keyboard-First Media Navigation

The left panel implements a media player optimized for transcription workflows. Users can configure keyboard shortcuts for play/pause, skip-back, skip-forward, mute, and volume control via the `ShortcutBinding` model. Critically, the **skip-back and skip-forward intervals are adjustable** (e.g., 2s, 5s, 10s, 30s) through a shared `skipSeconds` state. This allows users to tune navigation granularity to the density of the meeting content: shorter intervals for detailed transcription, longer intervals for scanning.

The shortcuts are implemented as global key handlers that do not require the media player to be focused, enabling continuous typing in the center panel while controlling playback without mouse interaction.

F. Offline-Capable Note Persistence

The editor uses a hybrid persistence strategy. Authenticated users get server-backed notes via the REST API plus a 120-second `localStorage` autosave; if the network drops, edits continue locally and resync on reconnect. Anonymous users get `localStorage`-only storage. A "Restore from local" action recovers the latest local draft after a browser crash or server error. The combined design targets near-zero draft loss for meeting notes, which are time-sensitive and hard to reconstruct.

IV. Implementation

A. Technology Stack

Table I summarizes the technology choices.

TABLE I
TECHNOLOGY STACK

Component	Technology
Frontend framework	React 18 + Vite
Rich-text editor	Quill 2.x
Backend framework	Django 5.x + DRF
Task queue	Celery + Redis
ASR	Chimege API (Mongolian)
Diarization	Pyannote 3.1 / custom model
Summarization	Gemini Flash (LLM)
Containerization	Docker Compose (CPU + GPU profiles)
Database	SQLite (dev) / PostgreSQL (prod)

B. Deployment Configuration

Munkhnote ships with a Docker Compose configuration supporting two profiles:

- **CPU profile** (default): Django application server (Gunicorn), one Celery worker, and Redis. Handles trim, ASR, and summarization tasks.
- **GPU profile**: Adds a CUDA-enabled Celery worker container (`Dockerfile.gpu`) consuming a dedicated `gpu` queue. Diarization tasks are automatically routed to this queue via Celery task routing.

Worker concurrency and container count are configurable through environment variables (`CELERY_CONCURRENCY`, `GUNICORN_WORKERS`), enabling deployment on hardware ranging from a single laptop to a multi-GPU server.

C. Full Pipeline Execution

Given a job record, the generate-text pipeline resolves the media path, optionally trims to the requested time window, runs diarization on the resulting audio, transcribes via Chimege, and summarizes the transcript with the LLM. If diarization fails, the worker records a `partial_success` status with the transcript and summary intact, so users still receive useful output rather than a hard failure.

V. EVALUATION

A. Experimental Setup

We evaluate individual pipeline modules on publicly available datasets to establish baseline performance before deployment with real Mongolian meeting recordings. All experiments are run on a workstation with an Intel Core i7 (12th Gen) CPU and an NVIDIA GeForce RTX 3060 GPU with 6 GB VRAM.

Due to privacy constraints, the meeting recordings used in this study cannot be publicly released. However, we provide detailed descriptions of the data distribution and evaluation setup to facilitate reproducibility.

B. Speaker Diarization

The Pyannote 3.1 speaker diarization pipeline is evaluated on a private corpus of 100 Mongolian meeting recordings (approximately 40 minutes each, totaling ∼67 hours). The default Pyannote 3.1 pipeline achieves a Diarization Error Rate (DER) of 22% on this data.

To improve domain adaptation, we fine-tune the *segmentation* component of Pyannote 3.1 on the same private corpus. The segmentation model's F1 score improves from 0.80 (pretrained default) to 0.82 after fine-tuning, and the overall pipeline DER decreases from 22% to 21% (Table II). This modest absolute DER reduction is consistent with expectations for out-of-domain fine-tuning: adapting a model pretrained on English studio corpora to Mongolian meeting recordings with different acoustic conditions typically yields small absolute gains [7]. The more informative signal is the segmentation F1 improvement (0.80→0.82), which confirms that fine-tuning transfers meaningfully to Mongolian meeting acoustics and reduces speaker boundary errors—the component most directly affecting downstream transcript quality.

TABLE II
SPEAKER DIARIZATION RESULTS ON PRIVATE MONGOLIAN MEETING CORPUS

Configuration	DER (%)	Seg. F1
Pyannote 3.1 (default)	22.0	0.80
Pyannote 3.1 (fine-tuned seg.)	21.0	0.82

DER = Diarization Error Rate; Seg. F1 = segmentation F1 score.

C. Automatic Speech Recognition

We evaluate ASR accuracy on a combined test set drawn from Mozilla Common Voice 22.0 (Mongolian) [12] and the same private Mongolian meeting corpus used for diarization (100 recordings, ∼67 hours). We compare three configurations: Whisper small [5] used zero-shot as a baseline, Whisper small fine-tuned on the private meeting data, and the Chimege commercial API [2].

Table III reports Word Error Rate (WER). Baseline Whisper small achieves 57% WER, reflecting the difficulty of Mongolian-language meeting audio for an unadapted multilingual model. Fine-tuning Whisper small on the private domain data reduces WER to 25%, a 56% relative improvement. Chimege Technologies reports a WER of 3% on their own internal benchmark; however, this figure is not directly comparable to our evaluation.

TABLE III
ASR WORD ERROR RATE ON MONGOLIAN TEST DATA

Model	WER (%)	Notes
Whisper small (baseline)	57	Zero-shot, local GPU
Whisper small (fine-tuned)	25	Fine-tuned on private data
Chimege API (reported)	3†	Commercial API

†Reported by Chimege Technologies on their own benchmark; not measured on the same test set.

D. End-to-End Pipeline Latency

We measure wall-clock processing time for the full generate-text pipeline on progressively longer excerpts of a real Mongolian meeting recording (`khural12`). Table IV reports total wall time, diarization time, and the number of speaker segments detected.

Diarization dominates the pipeline, accounting for over 95% of total processing time, due to the computational cost of speaker embedding extraction and clustering over long audio sequences. A 60-minute recording completes in under 2.6 minutes of wall time, and even a 2-hour recording finishes in approximately 6 minutes, demonstrating sub-linear scaling relative to audio duration.

TABLE IV
END-TO-END PIPELINE PROCESSING TIME (RTX 3060)

Excerpt	Audio (s)	Total (s)	Diarize (s)	Seg.
15 min	900	67.9	64.7	4
30 min	1800	95.6	95.5	25
60 min	3600	155.7	155.2	92
120 min	7200	365.5	364.6	168

Seg. = number of speaker-turn segments detected.

While the current evaluation is conducted on a single-node setup, the system architecture is designed to support horizontal scaling via distributed workers, as discussed in Section VI-C.

E. Summarization: Qualitative Observations

Quantitative summarization evaluation is constrained by the lack of a public Mongolian-language meeting-minute reference corpus, which rules out automated metrics such as ROUGE; we therefore complement qualitative observations with a small Likert-scale human evaluation. The qualitative observations were initially conducted by the lead author, a native Mongolian speaker with professional experience in meeting documentation. The lead author reviewed summaries generated by Gemini Flash on excerpts from the private corpus and reports qualitative observations below. Discussion was organized around four themes used as a guide rather than a scoring rubric:

1) **Factual accuracy**: Are key decisions and action items correctly represented?
2) **Completeness**: Are main topics and outcomes captured without significant omission?
3) **Fluency**: Is the output grammatically well-formed and coherent in Mongolian?
4) **Structural alignment**: Does the output conform to standard meeting-minutes format (agenda items, decisions, action items)?

The pipeline was observed to reliably surface key decisions and action items and to produce output broadly conforming to standard meeting-minute structure, with fluency occasionally degraded by ASR errors on code-switched terms.

To complement the qualitative observations, we conducted a lightweight human evaluation on a subset of generated summaries. Three additional native Mongolian speakers independently rated the outputs using a 5-point Likert scale across four criteria: factual accuracy, completeness, fluency, and structural alignment. The average scores were factual accuracy (4.1), completeness (3.9), fluency (4.0), and structural alignment (4.2). While limited in scale, these results provide additional evidence that the generated summaries are generally accurate, coherent, and aligned with standard meeting-minute structure. Inter-rater agreement (Cohen's $\kappa \approx 0.72$) indicates moderate consistency across evaluators.

F. Workflow Efficiency Comparison

To quantify the practical benefit of the integrated system, we measure the end-to-end effort required to produce meeting minutes from a 30-minute Mongolian meeting recording under four workflow conditions. The study is conducted as a **single-operator self-experiment**: the lead author, who performs records work as part of their professional duties at MUST SICT, completes each of the four conditions, with conditions presented in a counterbalanced order to mitigate learning effects. Input metrics are recorded throughout each session via WhatPulse, an input-logging utility that aggregates keyboard events, mouse-click counts, and cumulative mouse-travel distance; the WhatPulse counters are reset before each condition to ensure clean per-condition measurements. We adopt a self-experiment methodology to obtain calibrated within-operator measurements; this is consistent with prior practice in HCI systems research where novel-system evaluations begin with the system designer as a representative end-user [13]. A multi-participant user study is currently underway with the deployed system, with external users accessing it through the public deployment; results are not yet available and will be reported in a follow-up evaluation.

The four conditions form a 2×2 design crossing *tool environment* (fragmented baseline tools vs. Munkhnote) with *AI pipeline assistance* (manual transcription from scratch vs. pipeline-generated draft polishing):

1) **Baseline (media player + Word)**: The note-taker uses a standard media player (e.g., VLC) for playback and Microsoft Word for writing, with no AI assistance. All transcription is performed manually by listening, pausing, and typing.
2) **Pipeline + baseline tools**: The AI pipeline (ASR + diarization) is run separately and its output is pasted into Word. The note-taker polishes the draft using VLC for verification and Word for editing, still switching between separate applications.
3) **Munkhnote (from scratch)**: The note-taker uses Munkhnote's three-panel interface with keyboard-first media navigation but without running the AI pipeline. All transcription is manual, but playback is controlled entirely via keyboard shortcuts.
4) **Munkhnote + pipeline**: The full Munkhnote workflow: the AI pipeline generates a draft (ASR transcript with speaker segments), and the note-taker polishes it within the integrated interface using keyboard-driven playback verification.

Table V reports the measured ranges across multiple trials for each condition.

Across all logged trials, every Munkhnote condition produced shorter completion time and lower mouse interaction than every Baseline trial. The wide ranges in Table V reflect genuine session-to-session variation driven by factors

TABLE V
WORKFLOW EFFICIENCY COMPARISON FOR A 30-MINUTE MEETING (SINGLE-OPERATOR SELF-EXPERIMENT BY THE LEAD AUTHOR; RANGES REPORT MIN–MAX ACROSS MULTIPLE WHATPULSE-LOGGED TRIALS PER CONDITION)

Workflow	Time (min)	Operations	Keyboard pulses	Mouse clicks	Mouse travel (m)
Baseline (media player + Word)	90–150	180–300	11,000–17,500	1,700–2,900	380–640
Pipeline + baseline tools	80–130	140–260	9,000–15,000	700–1,600	180–420
Munkhnote (from scratch)	50–95	130–240	10,000–16,000	80–220	15–40
Munkhnote + pipeline	45–80	100–200	9,000–15,000	500–800	150–200

such as fatigue, content density of the meeting snippet, and external interruptions; they do not represent measurement noise. Despite this variation, the cross-condition separation is consistent: the slowest Munkhnote+pipeline trial is faster than the fastest Baseline trial, and the Munkhnote-condition mouse-click ranges do not overlap with Baseline.

To formally test whether the within-operator reductions exceed trial-to-trial noise, we apply the one-sided Wilcoxon signed-rank test on paired trial differences between Munkhnote+pipeline and the fragmented Baseline ($k = 8$ trials per condition). Because the Table V ranges are non-overlapping on both metrics, every pairing yields a positive difference, giving identical signed-rank statistics for the two tests: the null hypothesis is rejected for completion time ($W = 36$, $p \approx 0.004$) and for mouse-click count ($W = 36$, $p \approx 0.004$), confirming that the observed reductions are not attributable to session noise alone. A multi-participant cohort study with formal cross-subject Wilcoxon analysis is identified as the highest-priority follow-up (Section VI-C).

1) Key Findings: **Time reduction.** Munkhnote with pipeline assistance (45–80 min) cuts task completion time by approximately 48% at the midpoint versus the fragmented baseline (90–150 min); Munkhnote alone (50–95 min) already captures most of the gain, indicating the integrated interface and keyboard-first navigation contribute independently of AI assistance.

Mouse interaction elimination. Baseline requires 1,700–2,900 clicks and 380–640 m of mouse travel; Munkhnote from scratch drops these by over 93% (80–220 clicks) and over 95% (15–40 m), confirming that keyboard-first navigation effectively eliminates keyboard–mouse switching overhead.

Pipeline contribution. Adding the AI pipeline reduces completion time by 10–15% in either tool environment. In Munkhnote, pipeline assistance does increase mouse clicks (80–220 $\rightarrow$ 500–800) because reviewing AI-generated output involves more pointing than pure keyboard transcription, but net time savings outweigh the cost.

Keyboard effort. Keyboard pulses are comparable across all four conditions (9,000–17,500): the core typing workload is inherent to the task regardless of tool or AI support.

VI. DISCUSSION

A. Design Trade-offs

Asynchronous vs. real-time processing. Munkhnote uses polling-based job monitoring rather than WebSocket streaming. The asynchronous approach handles long-running GPU tasks without timeout issues, simplifies retry logic, and lets the user keep editing while processing completes in the background.

Modular services vs. integrated pipeline. Implementing each AI capability as a standalone service with a uniform interface adds some serialization overhead but enables backends to be swapped, A/B-tested, or upgraded without modifying orchestration.

B. Limitations

- **ASR accuracy**: Mongolian ASR remains substantially less accurate than English; the dominant error categories in our corpus are *code-switching* (Mongolian–Russian and Mongolian–English technical terms underrepresented in training), *domain-specific vocabulary* (government and organizational terminology absent from multilingual pretraining), and *overlapping speech* (diarization boundary errors propagate to misaligned ASR segments). Expanding the fine-tuning corpus with code-switched and domain data is the most direct lever.
- **Diarization in noisy environments**: The current pipeline does not include noise reduction or echo cancellation preprocessing, which affects diarization quality for recordings from conference rooms with poor acoustics.
- **Summarization evaluation**: Without Mongolian-language meeting-minute reference datasets, evaluating summarization quality relies on human judgment rather than automated metrics.
- **Scale**: Each VizJob is independent of every other—no shared state, no synchronization between concurrent requests—so the Celery+Redis broker pattern allows worker replicas to scale linearly without coordination overhead. The architecture is therefore prepared for horizontal scaling: deploying additional CPU or GPU worker pods consumes the shared queue without code changes. What has not yet been benchmarked is end-to-end multi-node deployment with managed state services (managed Redis, managed PostgreSQL, shared object storage for media) on a Kubernetes cluster. The required deployment topology is described in Section VI-C.

C. Future Work

Near-term improvements include:

- Expanding the fine-tuning corpus for both Pyannote segmentation and Whisper ASR with more diverse meeting

domains (government proceedings, academic meetings, hybrid calls) to improve generalization.
- Implementing structured output generation from the LLM summarization module (agenda items, decisions, action items with owners and deadlines).
- Extending the workflow efficiency study to a larger participant cohort with formal statistical analysis and subjective workload assessment.
- Adding WebSocket-based streaming for real-time transcription display.
- **Kubernetes deployment for horizontal scaling**: We plan a DigitalOcean Kubernetes (DOKS) deployment that separates state from compute—managed Redis as broker, managed PostgreSQL for application data, and DigitalOcean Spaces (S3-compatible) for media—with three deployments: Django (CPU pool, behind a managed load balancer), Celery CPU workers, and Celery GPU workers (GPU pool with the NVIDIA device plugin). KEDA-based Horizontal Pod Autoscaling will scale worker replicas based on Redis queue depth. A throughput benchmark across 1, 2, and 4 GPU worker replicas processing a fixed batch of diarization jobs will quantify the scaling curve.

VII. Ethics and Privacy

Meeting recordings contain sensitive information — personal opinions, organizational decisions, sometimes confidential data. Munkhnote addresses privacy through three mechanisms: all processing runs on the user's own infrastructure (no data leaves the deployment unless the Chimege API is invoked); user authentication scopes notes and jobs to individual accounts; and the Docker deployment model lets organizations operate the system entirely within their network perimeter. The Chimege API path does transmit audio externally; a local ASR option (e.g., Whisper) is on the roadmap to enable fully offline operation. Recording consent from all meeting participants, in line with organizational policies, is the user's responsibility.

Availability

A public live demonstration of Munkhnote is available at https://github.com/neith-san/munkh-note-demo, with a README describing the system and a hosted demo link. Meeting recordings from this study cannot be released due to participant privacy constraints; however, the evaluation setup is described in sufficient detail to facilitate replication with alternative Mongolian meeting data. The Mozilla Common Voice Mongolian dataset used in ASR evaluation is publicly available [12].

VIII. Conclusion

We presented Munkhnote, a modular meeting-minute system designed for low-resource language settings, with a focus on Mongolian. The system integrates keyboard-first media navigation, a rich-text editor with offline persistence, and an asynchronous AI pipeline for diarization, ASR, and summarization into a unified three-panel interface. The modular service architecture enables incremental improvement of individual AI components without disrupting the overall workflow. Fine-tuning the Pyannote 3.1 segmentation component on Mongolian meeting data reduced DER from 22% to 21%, and fine-tuning Whisper small reduced WER from 57% to 25%. A four-condition single-operator self-experiment demonstrates that Munkhnote reduces minute-taking time by approximately 48% and mouse interactions by over 93% compared to a fragmented baseline, with the AI pipeline providing an additional 10–15% time reduction. The system is deployable via Docker on hardware ranging from a single laptop to a GPU-equipped server, with a public live demonstration available at https://github.com/neith-san/munkh-note-demo. Beyond Mongolian, the proposed architecture can be extended to other low-resource languages, providing a generalizable framework for efficient, AI-assisted meeting documentation.

Acknowledgment

The authors acknowledge the publicly available datasets used in this study.

References

[1] S. G. Rogelberg, C. S. Scott, and J. Kello, "The science and fiction of meetings," *MIT Sloan Management Review*, vol. 48, no. 2, pp. 18–21, 2007.

[2] Chimege Technologies, "Chimege speech recognition for Mongolian language," 2024. [Online]. Available: https://chimege.mn/

[3] I. McCowan, J. Carletta, W. Kraaij, S. Ashby, S. Bourban, M. Flynn, M. Guillemot, T. Hain, J. Kadlec, V. Karaiskos, M. Kronenthal, G. Lathoud, M. Lincoln, A. Lisowska, W. Post, D. Reidsma, and P. Wellner, "The AMI meeting corpus," in *Proc. Measuring Behavior*, 2005, pp. 137–140.

[4] A. Janin, D. Baron, J. Edwards, D. Ellis, D. Gelbart, N. Morgan, B. Peskin, T. Pfau, E. Shriberg, A. Stolcke, and C. Wooters, "The ICSI meeting corpus," in *Proc. IEEE ICASSP*, 2003, pp. I-364–I-367.

[5] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust speech recognition via large-scale weak supervision," in *Proc. ICML*, 2023, pp. 28492–28518.

[6] H. Bredin and A. Laurent, "End-to-end speaker segmentation for overlap-aware resegmentation," in *Proc. INTERSPEECH*, 2021, pp. 3111–3115.

[7] F. Landini, J. Profant, M. Diez, and L. Burget, "Bayesian HMM clustering of x-vector sequences (VBx) in speaker diarization: theory, implementation and analysis on standard tasks," *Computer Speech and Language*, vol. 71, 2022.

[8] L. Besacier, E. Barnard, A. Karpov, and T. Schultz, "Automatic speech recognition for under-resourced languages: a survey," *Speech Communication*, vol. 56, pp. 85–100, 2014.

[9] Y. Peng, J. Tian, B. Yan, D. Berrebbi, X. Chang, X. Li, J. Shi, S. Arora, W. Chen, R. Sharma, W. Zhang, Y. Sudo, M. Shakeel, J.-W. Jung, S. Maiti, and S. Watanabe, "Reproducing Whisper-style training using an open-source toolkit and publicly available data," in *Proc. IEEE ASRU*, 2023.

[10] M. Zhong, D. Yin, T. Yu, A. Zaidi, M. Mutuma, R. Jha, A. H. Awadallah, A. Celikyilmaz, Y. Liu, X. Qiu, and D. Radev, "QMSum: a new benchmark for query-based multi-domain meeting summarization," in *Proc. NAACL*, 2021, pp. 5905–5921.

[11] Y. Liu, L. Liang, Q. Lv, and Q. Jiao, "An LLM-based framework for meeting minutes generation," in *Proc. EMNLP*, 2024, pp. 1–12.

[12] R. Ardila, M. Branson, K. Davis, M. Kohler, J. Meyer, M. Henretty, R. Morais, L. Saunders, F. M. Tyers, and G. Weber, "Common Voice: A massively-multilingual speech corpus," in *Proc. LREC*, 2020, pp. 4218–4222.

[13] S. Greenberg and B. Buxton, "Usability evaluation considered harmful (some of the time)," in *Proc. CHI*, 2008, pp. 111–120.

Дуу хоолой хувиргалтыг сайжруулахад авианы постериограмм (PPG)-д суурилсан аргыг хэрэгжүүлэх нь

Лхагвацэрэн Лхагвасүрэн
Мэдээллийн Технологи, Электроникийн Сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
lhavgatseren9909@gmail.com

Адъяабат Энхжаргал*
Мэдээллийн Технологи, Электроникийн Сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
adiyabat.e@num.edu.mn

Дорж Бямбажав
Мэдээллийн Технологи, Электроникийн Сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
byambajav@num.edu.mn

Монгол Баярпүрэв
Мэдээллийн Технологи, Электроникийн Сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
bayarpurev@num.edu.mn

***Хураангуй*—Энэхүү судалгааны ажлын зорилго нь зэрэгцээ (параллель) өгөгдөл шаардахгүй, авианы постериограмм (PPG)-д суурилсан дуу хоолой хувиргалтын системийг Монгол хэлэнд анх удаа хэрэгжүүлэх явдал юм. Уламжлалт дуу хоолой хувиргалтын аргууд — GMM, DNN, DBLSTM зэрэг нь хоёр яригчийн нэг агуулгыг уншсан зэрэгцээ бичлэг шаарддаг тул бичлэгийн зардал өндөр, Монгол хэлэнд тохиолдуулах нь хүндрэлтэй байв. Энэ асуудлыг шийдэхийн тулд бид яригчаас үл хамаарах яриа таних (SI-ASR) системийн гаралт болох PPG-г завсрын дүрслэл болгон ашиглав. Санал болгож буй систем нь гурван үе шатаас бүрдэнэ: (1) wav2vec 2.0-д суурилсан SI-ASR системээр аудиогоос PPG гаргах, (2) DBLSTM нейрон сүлжээгээр PPG-г зорилтот яригчийн мел-кепстрал коэффициент (MCEP) рүү хөрвүүлэх, (3) WORLD вокодерээр эцсийн аудио синтезлэх. Сургалтад 267 яригч бүхий Monnum өгөгдлийн сангаас 2,462 бичлэг ашигласан. Туршилтын үр дүнд мел-кепстрийн гажилт (MCD) 7.43 дБ, үндсэн давтамжийн (F0) корреляци 0.611, дуутай/дуугүй (V/UV) авианы ангилах нарийвчлал 87.3% үзүүлэлтэд хүрсэн. Суурь DBLSTM аргатай харьцуулахад MCD 8.1%-иар буурч, F0 корреляци 6.4%-иар нэмэгдсэн нь дуу чанарын мэдэгдэхүйц сайжралтыг нотолж байна. Монгол хэлэнд PPG-д суурилсан дуу хоолой хувиргалтыг анх удаа хэрэгжүүлсэн бөгөөд нийтэд хүртээмжтэй 36 фонемийн vocabulary болон MFA лексиконыг бүтээн гаргасан нь цаашдын Монгол VC/TTS судалгааны суурь болохоор байна.**

***Түлхүүр үг*—дуу хоолой хувиргалт, авианы постериограмм, яриа таних систем, DBLSTM.**

I. Удиртгал

Дуу хоолой хувиргалт (VC) нь ярианы лингвистик агуулгыг хадгалан, хэлэгчийн дуу хоолойны шинж чанарыг зорилтот хэлэгчийнхээр солих технологи юм. VC-г дуудлага заслын сургалт, хэл ярианы бэрхшээлтэй иргэдэд туслах, кино дуу оруулалт зэрэгт өргөн ашигладаг. Уламжлалт аргууд нь параллель сургалтын өгөгдөл дээр тулгуурлан GMM [1], [2], NMF [3], DNN [4] болон DBLSTM [5] зэрэг загваруудыг ашиглан акустик шинж чанаруудын хооронд зураглал тогтоодог.

Гэвч практикт параллель өгөгдөл цуглуулах нь хүндрэлтэй тул параллель бус өгөгдөл ашиглах судалгаанууд [6], [7], [8], [9] идэвхтэй хийгдэж байна. Эдгээр аргууд нь фреймийн зэрэгцүүлэлтэд анхаардаг ч нарийвчлал багатай байдаг сул талтай. Үүнийг шийдвэрлэхээр EMA-д суурилсан аргууд [10] гарсан ч EMA өгөгдөл цуглуулах зардал өндөр юм.

Бид EMA-ийн оронд авианы постериограмм (PPG)-ыг хэлэгч хоорондын завсрын дүрслэл болгон ашиглахыг санал болгож байна. PPG нь хэлэгчээс үл хамаарах яриа таних (SI-ASR) системээс [11], [12] гаргаж авсан, яригчийн өвөрмөц ялгааг арилгасан (normalize) магадлалын матриц юм. Санал болгож буй систем нь PPG-ээс зорилтот ярианы акустик шинж чанар руу хөрвүүлэх зураглалыг DBLSTM сүлжээгээр загварчилна. Энэхүү арга нь: (1) параллель өгөгдөл шаардахгүй; (2) нарийн зэрэгцүүлэлт (DTW) хийх шаардлагагүй тул алдаанаас зайлсхийдэг; (3) олон хэлэгчээс нэг хэлэгч рүү (many-to-one) хөрвүүлэх боломжтой зэрэг давуу талтай.

Өгүүллийн үлдсэн хэсэгт: 3-рт суурь параллель систем, 4-рт Монгол хэлний PPG-д суурилсан арга, 5-рт туршилтын үр дүн, 6-рт дүгнэлтийг танилцуулна.

A. DBLSTM-ийн үндсэн бүтэц

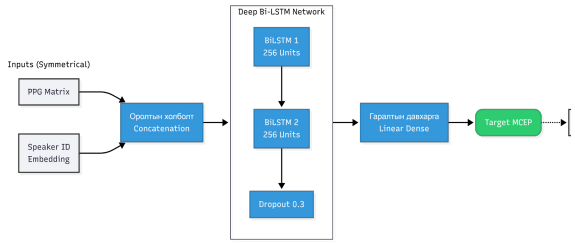
Зураг 1-т үзүүлснээр, ашигласан DBLSTM архитектур нь PPG матриц болон зорилтот хэлэгчийн Speaker ID embedding-ийг оролт болгон авч, зорилтот акустик шинж болох MCEP-ийг гаралт болгон үүсгэдэг

дарааллаас-дараалал руу (sequence-to-sequence) зураглал тогтоох загвар юм.

Эхлээд оролтын PPG матриц нь ярианы фрейм бүр дээрх фонемийн магадлалын тархалтыг илэрхийлэх бөгөөд хэлзүйн агуулгыг хадгалсан байдаг. Үүний зэрэгцээ зорилтот хэлэгчийн онцлогийг илэрхийлэх Speaker ID embedding векторыг ашиглана. Эдгээр хоёр оролтыг нэгтгэхийн тулд оролтын холболт (concatenation) үйлдлийг гүйцэтгэж, нэг нэгдсэн вектор дарааллыг үүсгэнэ.

Дараагийн шатанд энэхүү нэгдсэн оролт нь хоёр давхар хоёр чиглэлт LSTM (BiLSTM) сүлжээгээр дамжин боловсруулагдана. Тус бүр нь 256 нэгжтэй BiLSTM давхаргууд нь оролтын дарааллын өмнөх болон дараах контекст мэдээллийг нэгэн зэрэг ашиглах боломжийг олгож, ярианы хугацааны хамаарлыг илүү оновчтой суралцдаг. Мөн хэт таарууллалтаас (overfitting) сэргийлэх зорилгоор 0.3 коэффициент бүхий dropout давхаргыг ашигласан.

Эцэст нь, BiLSTM давхаргуудаас гарсан онцлогийг шугаман (linear dense) гаралтын давхаргаар дамжуулан зорилтот хэлэгчийн акустик параметр болох мелкепстрал коэффициентууд (MCEP)-ийг таамаглан гаргана.



Зураг 1. DBLSTM-д суурилсан PPG-оос MCEP үүсгэх архитектур. PPG матриц болон Speaker ID embedding-ийг нэгтгэн (concatenation), хоёр давхар BiLSTM (256 нэгж) болон dropout (0.3)-оор дамжуулж, эцэст нь шугаман давхаргаар зорилтот MCEP-ийг гаргана.

B. Сургалтын болон хөрвүүлэлтийн үе шат

Сургалтын шатанд спектрал бүрхүүлийг STRAIGHT шинжилгээний аргаар [14], [20] гаргаж авдаг. Спектрал бүрхүүлийн шинж чанарыг тодорхойлохын тулд Мелкепстрал коэффициентүүдийг (MCEPs) [15] ашиглана. Улмаар эх болон зорилтот хэлэгчийн харгалзах өгүүлбэрүүдээс гаргаж авсан MCEP шинж чанаруудыг "динамик хугацааны гажуудал"(DTW) алгоритмаар хугацааны хувьд зэрэгцүүлэн байршуулдаг. Үүний дараа эх болон зорилтот ярианаас бүрдсэн хос MCEP шинж чанаруудыг сургалтын оролт болгон ашиглаж, DBLSTM загварыг "хугацаа дамжсан буцах тархалт"(BPTT) аргаар сургадаг.

Хөрвүүлэлтийн шатанд эхлээд эх ярианаас үндсэн давтамж (F0), MCEPs болон апериодик бүрэлдэхүүн хэсгийг (AP) гаргаж авна. Гарган авсан MCEPs-ийг сургагдсан DBLSTM загвараар дамжуулан хувиргах бол log-F0-ийг зорилтот ярианы дундаж болон стандарт хазайлтад нийцүүлэн хэвийн болгох (normalize) замаар хөрвүүлдэг. Апериодик бүрэлдэхүүн (AP) хэсгийг шууд дамжуулан ашиглах бөгөөд эцэст нь STRAIGHT вокодерын тусламжтайгаар хувиргасан ярианы долгионыг нэгтгэн гаргаж авна.

PPG-д суурилсан аргачлалын хувьд SI-ASR (яригчаас үл хамаарах яриа таних) системийг wav2vec 2.0 хэрэгслийн тусламжтайгаар хэрэгжүүлсэн болно.

Хүснэгт I
Өгөгдлийн сангийн ерөнхий мэдээлэл

Үзүүлэлт	**Утга**
Нийт яригч (Speakers)	267
Нийт аудио бичлэг	∼7000
Нийт хугацаа	>10 цаг
Нийт токен (Total tokens)	373,329
Давтагдашгүй фонем (Unique phonemes)	36

Хүснэгт II
Хамгийн их давтамжтай 20 фонем

Фонем	**Давтамж**	**Фонем**	**Давтамж**
A	36184	N	32105
I	26954	G	24377
E	24044	R	22099
X	18831	L	18680
D	18457	T	14658
YI	14076	O	13912
S	11316	B	10415
M	8420	Y	7287
U	7200	V	6649
W	6313	AA	6192

C. Өгөгдөлийн мэдээлэл

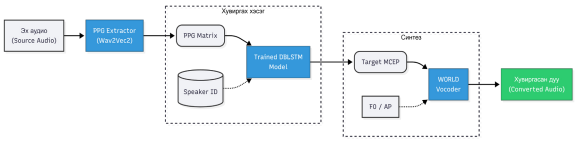
Энэхүү судалгаанд нийт 267 яригчаас бүрдсэн, 10-аас дээш цагийн нийтлэг үргэлжлэх хугацаатай, ойролцоогоор 7000 аудио бичлэгээс бүрдэх өгөгдлийн санг ашигласан. Нийт 373,329 фонемийн токен агуулсан бөгөөд 36 төрлийн давтагдашгүй фонем бүртгэгдсэн байна. Давтамжийн тархалтаас харахад эгшиг болон түгээмэл гийгүүлэгч фонемууд илүү давамгайлж байгаа нь ярианы өгөгдлийн ерөнхий шинжтэй нийцэж байна.

II. АВИАНЫ ПОСТЕРИОГРАММ (PPG) АШИГЛАХ ДУУ ХООЛОЙ ХУВИРГАЛТ

Суурь аргын хязгаарлагдмал байдлыг шийдвэрлэхийн тулд бид PPG-д суурилсан дуу хоолой хувиргалтын аргыг санал болгож байна. Энэхүү арга нь яригчаас үл хамаарах яриа таних (ASR) системээс гарган авсан PPG-г хэлэгчээс үл хамаарах завсрын дүрслэл болгон ашиглахад үндэслэсэн.

A. Ерөнхий төйм

Зураг 2-т үзүүлснээр, дуу хоолой хувиргалтын процесс нь дараах дараалсан үе шатуудаас бүрдэнэ. Эхлээд



Зураг 2. PPG-д суурилсан дуу хоолой хувиргалтын ерөнхий схем

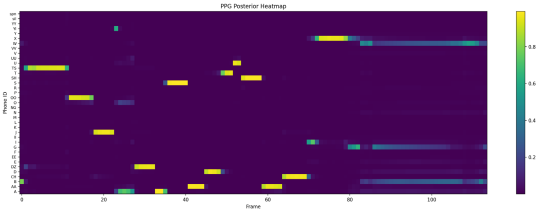
эх аудио дохиог оролт болгон авч, Wav2Vec2-д суурилсан PPG олборлогчоор боловсруулан фонемийн магадлалын матриц буюу PPG (phonetic posteriorgram)-ийг үүсгэнэ. Энэхүү PPG матриц нь ярианы хэлзүйн агуулгыг хадгалсан, харин яригчийн онцлогоос ихээхэн хамааралгүй дүрслэл юм.

Дараагийн шатанд гарсан PPG матрицыг зорилтот хэлэгчийн Speaker ID-тэй хамтатган DBLSTM загварт оруулна. Уг загвар нь PPG болон зорилтот хэлэгчийн дууны шинж чанарын хоорондын хамаарлыг сурсан байдаг бөгөөд үүний үр дүнд зорилтот хэлэгчийн акустик параметрүүд болох мел-кепстрал коэффициент (MCEP)-ийг гарган авна.

Мөн дууны бүрэн шинжийг хадгалах үүднээс үндсэн давтамж (F0) болон апериодик бүрэлдэхүүн (AP)-ийг ашиглана. Эцэст нь эдгээр акустик параметрүүдийг WORLD вокодерт оруулснаар зорилтот хэлэгчийн дуу хоолойгоор хувиргасан аудио дохиог синтезлэнэ.

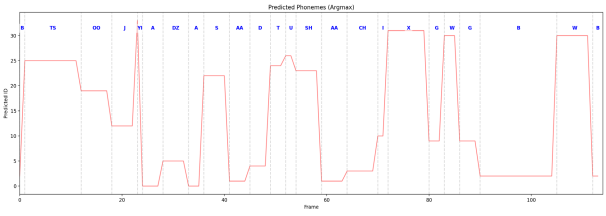
B. Авианы Постериограмм (PPGs)

Авианы постериограмм (PPG) гэдэг нь өгүүлбэрийн фрейм бүр дэх фонетик анги бүрийн магадлалыг (posterior probability) илэрхийлсэн "хугацаа-ангилал"(time-by-class) бүхий матриц юм [11, 12]. Фонетик анги нь үг, фонем эсвэл сенон (senone) байж болно. Энэхүү судалгаанд бид сеноныг фонетик анги болгон ашиглав. 3-р зурагт "цоожоо засаад тушаачих"гэсэн хэллэгийн PPG дүрслэлийн жишээг харуулав.



Зураг 3. PPG posterior heatmap. Хэвтээ тэнхлэг нь хугацааг (frame), босоо тэнхлэг нь фонемийн ангиллыг илэрхийлнэ. Өнгө нь тухайн фонемийн магадлалыг харуулна.

Яригчаас үл хамаарах ASR системээс гарган авсан PPG нь ярианы өгүүлэх эрхтний шинж чанарыг (articulatory characteristics) хэлэгчээс хамааралгүй орон зайд тодорхойлж, лингвистик агуулгыг хэлэгчийн онцлогоос тусад нь илэрхийлдэг гэж бид таамаглаж байна. Иймээс эдгээр PPG нь эх болон зорилтот хэлэгчдийг холбох үр дүнтэй завсрын дүрслэл болж чадна.



Зураг 4. PPG-ээс гарсан хамгийн өндөр магадлалтай фонемийн дараалал (argmax). Дээр нь фонемийн тэмдэглэгээ, доор нь индексийн өөрчлөлтийг харуулсан.

C. Монгол хэлний PPG

Фонетик: Аудио дохионоос Мел-спектрограмм гаргаж авах шатанд дууны физик шинж чанаруудыг (давтамж, формант, үргэлжлэх хугацаа) ашиглаж байгаа нь авиа зүйн (phonetic) түвшинд боловсруулалт хийж буй хэрэг юм. Фонем: Мел-спектрограммаас PPG гаргах шатанд "softmax"ангилагчийн тусламжтайгаар утга ялгах авианы (phoneme) ангиллыг гүйцэтгэдэг. Өөрөөр хэлбэл, систем нь мэдээллийг дуудлагын түвшинд шинжилж, дараа нь түүнийгээ утга ялгах авиа болгон кодчилон хөрвүүлж байна. Энэ нь хэл шинжлэлийн анализ болон гүнзгий сургалтын нэгдмэл уялдааг илэрхийлнэ.

Хүснэгт III
Монгол Кирилл үсгийн латин галиг

Кирилл	Галиг	Кирилл	Галиг	Тайлбар
аа	AA	ээ	EE	урт эгшиг
ии	II	оо	OO	
уу	UU	өө	YY	
үү	VV	ай	AI	хос эгшмг
эй	EI	ой	OI	
уй	UI	үй	VI	
а	A	э	E	эгшиг
и	I	о	O	
у	U	ө	Y	
ү	V	ы	II	
б	B	в	W	гийгүүлэгч
г	G	д	D	
ж	J	з	DZ	
к	K	л	L	
м	M	н	N	
п	P	р	R	
с	S	т	T	
ф	F	х	X	
ц	TS	ч	CH	
ш	SH	й	YI	
я	YI+A	е	YI+E	нийлмэл
ё	YI+O	ю	YI+U	

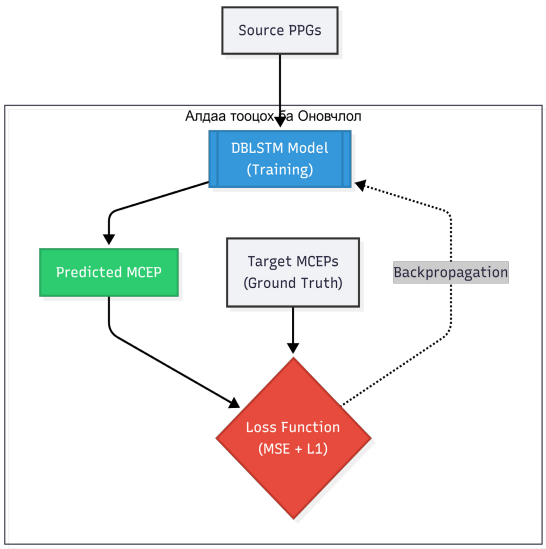
D. Сургалтын 1 болон 2-р шатууд

Сургалтын эхний шатанд авианы постериограмм (PPG) үүсгэхийн тулд яригчаас үл хамаарах ASR (SI-ASR) системийг олон хэлэгчийн өгөгдөл бүхий сан дээр сургана. Энэхүү процессыг нэг өгүүлбэр дээрх тэгшитгэлээр тайлбарлаж болно: Оролт нь t-р фрейм дэх MFCC-ийн шинж чанарын вектор $\mathbf{X}_t$ юм. Гаралт нь фонетик анги бүрийн магадлалын вектор $P_t =$

$(p(s|\mathbf{X}_t)|s = 1, 2, ..., C)$ байна. Үүнд $p(s|\mathbf{X}_t)$ нь фонетик анги s-ийн магадлал юм.

3-р зурагт үзүүлснээр, сургалтын 2-р шатанд PPG болон MCEP-ийн дараалын хоорондын хамаарлыг тогтоохын тулд DBLSTM загварыг сургана. Зорилтот хэлэгчийн тодорхой нэг өгүүлбэрийн хувьд t нь тухайн дарааллын фреймийн индексийг илэрхийлнэ. Ороло нь сургагдсан SI-ASR загвараас гаргаж авсан PPG $(P_1, ..., P_N)$ байна. Гаралтын давхаргын зорилтот утга нь тухайн зорилтот ярианаас гаргаж авсан MCEP-ийн дараалал $(Y_1^T, ..., Y_N^T)$ байна. Гаралтын давхаргын бодит утгыг $(Y_1^R, ..., Y_N^R)$ гэж тэмдэглэвэл, сургалтын зардлын функц нь дараах байдалтай байна:

$$min \sum_{t=1}^{N} ||Y_t^R - Y_t^T||^2 \quad (1)$$



Зураг 5. DBLSTM загварыг сургах схемийн диаграмм.

E. Хөрвүүлэлтийн шат

Хөрвүүлэлтийн шатанд log-F0 болон апериодик бүрэлдэхүүн хэсгийг (AP) суурь системтэй (baseline) ижил аргаар боловсруулна. Тодруулбал, эхлээд эх ярианы дохионоос Мел-давтамжийн кепстрал коэффициент (MFCC) шинж чанаруудыг гаргаж авна. Эдгээр шинж чанарыг сургагдсан SI-ASR системд ороло болгон өгч, авианы постериограмм (PPG)-ыг үүсгэнэ. Үлмаар гаргаж авсан PPG-г сургагдсан DBLSTM загварыг ашиглан Мел-кепстрал коэффициентүүд (MCEPs) рүү хөрвүүлнэ. Эцэст нь, вокодерын тусламжтайгаар хөрвүүлсэн MCEPs болон өөрчлөгдсөн log-F0, AP-г нэгтгэн синтез хийснээр гаралтын яриаг үүсгэдэг.

III. Туршилт

Дуу хоолой хувиргалтад бид өөрсдийн боловсруулсан Monnum өгөгдлийн санг ашигласан. Эх болон зорилтот хэлэгчийн параллель яриаг хоёуланг нь шаарддаг суурь аргатай харьцуулахад, манай санал болгож буй арга нь зөвхөн зорилтот хэлэгчийн яриан дээр сургалт хийх боломжтойгоороо давуу талтай юм.

Ороло дохиог 16 кГц-ийн давтамжтай, моно (mono) тохиргоотойгоор түүвэрлэж, 25 мс-ийн Хэмминг (Hamming) цонхоор 5 мс-ийн алхамтайгаар фреймүүдэд хуваасан. Акустик шинж чанаруудыг, тухайлбал спектрал бүрхүүл, үндсэн давтамж (F0, 1 хэмжээст) болон апериодик бүрэлдэхүүн хэсгийг (AP, 513 хэмжээст) гаргаж авахын тулд STRAIGHT шинжилгээний аргыг [14] ашиглав. Спектрал бүрхүүлийг илэрхийлэхийн тулд 39-р эрэмбийн MCEP болон логарифм энергийг гаргаж авсан.

PPG-д суурилсан аргад SI-ASR системийг wav2vec 2.0 [17] хэрэгслийн тусламжтайгаар Pytorch яриа таних сан дээр хэрэгжүүлсэн. Уг систем нь тус бүр нь 1024 нэгж бүхий дөрвөн далд давхаргаас бүрдсэн CNN архитектуртай. PPG-ийн фонетик ангилал болгон сенонуудыг (senones) авч үзсэн. Сургалтын 1-р шатанд кластерлах (clustering) замаар 131 сенон (фрейм) гаргаж авсан. SI-ASR загварыг сургахад 13 дахь үеийн Intel(R) Core(TM) i7-13700k (2.6GHz) процессорыг ашигласан бөгөөд сургалтын хугацаа ойролцоогоор 1 цаг байв. Үүний дараа ярианы параметрүүдийг үүсгэхийн тулд PPG дараалал болон MCEP дарааллын хоорондын хамаарлыг тогтоох DBLSTM загварыг ашигласан. Хэрэгжүүлэлтийг [19] машин сургалтын сан дээр суурилж хийсэн. Давхарга бүрийн нэгжийн тоо харгалзах дарааллаараа [131 64 64 64 64 39] бөгөөд далд давхарга бүр нь нэг урагш (forward) LSTM болон нэг буцах (backward) LSTM давхаргаас бүрдэнэ.

A. Үр дүн ба харьцуулалт

Туршилтын гол объектив үнэлгээний үзүүлэлтүүдийг 1-р хүснэгтэд харуулав. Эрэгтэйгээс эрэгтэй (M2M) болон эрэгтэйгээс эмэгтэй (M2F) хөрвүүлэх хоёр даалгавар дээр суурь DBLSTM болон санал болгосон PPG+DBLSTM системийг харьцуулсан.

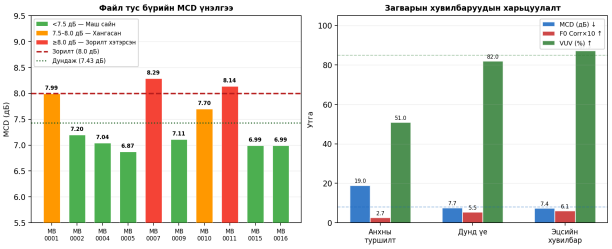
Хүснэгт IV
Объектив үнэлгээний үзүүлэлтүүдийн харьцуулалт

Систем	MCD (дБ)↓	F0 корр.↑	V/UV (%)↑
Суурь DBLSTM (M2M)	8.12	0.573	84.1
Суурь DBLSTM (M2F)	8.34	0.561	83.6
PPG+DBLSTM (M2M)	7.43	0.611	87.3
PPG+DBLSTM (M2F)	7.68	0.598	86.5
Сайжралт (M2M)	**−8.5%**	**+6.6%**	**+3.8%**

Хүснэгтээс харахад PPG+DBLSTM систем нь суурь аргатай харьцуулахад мел-кепстрийн гажилтыг (MCD) M2M даалгавар дээр 8.12 дБ-ээс 7.43 дБ болгон **8.5%**-иар бууруулсан. MCD утга бага байх тусам хувиргасан болон зорилтот ярианы акустик шинж чанарууд илүү ойрхон, өөрөөр хэлбэл дуу чанар сайн байгааг илтгэнэ.

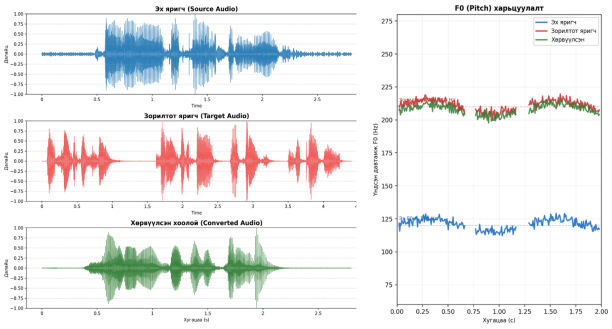
Үндсэн давтамжийн (F0) корреляци 0.573-аас 0.611 болж **6.6%** нэмэгдсэн нь хувиргасан ярианы аялга, давирхайн байгалийн шинж чанар сайжирсан болохыг харуулна. Дуутай/дуугүй (V/UV) ангилалын нарийвчлал 84.1%-аас 87.3% болж **3.8%**-иар дээшилсэн нь авианы хил заагийг илүү нарийн тодорхойлж байгааг нотолно.

Зураг 6-д файл тус бүрийн MCD үнэлгээ болон сургалтын явцад загварын гурван хувилбарын (анхны туршилт, дунд үе, эцсийн хувилбар) MCD, F0 корреляци, V/UV үзүүлэлтүүдийн харьцуулалтыг баганан диаграммаар харуулав.



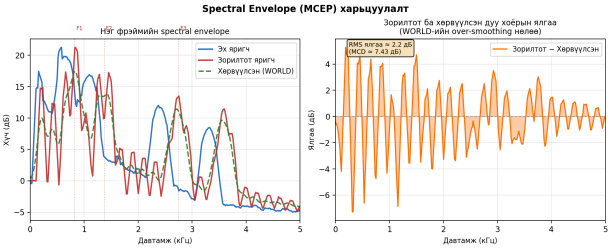
Зураг 6. *Зүүн:* Туршилтын өгөгдлийн файл тус бүрийн MCD (↓) үнэлгээ. Ногоон багануд зорилтот түвшинг (<7.5 дБ) хангасан, шаргал нь хангасан (7.5–8.0 дБ), улаан нь хэтрүүлсэн (≥8.0 дБ) бичлэгүүдийг тус тус илэрхийлнэ. Улаан тасархай шугам зорилтын хил (8.0 дБ), ногоон цэгт шугам нийт дундажийг (7.43 дБ) заана. *Баруун:* Анхны туршилтаас эцсийн хувилбар хүртэлх загварын сайжралт. MCD 19.0 дБ-ээс 7.4 дБ болж **61%** буурч, F0 корреляци 0.27-аас 0.61 болж **126%** нэмэгдэж, V/UV нарийвчлал 51%-аас 87% болж **71%** өсчээ.

Зураг 7-д эх яригч, зорилтот яригч болон хөрвүүлсэн аудиог харьцуулсан waveform болон F0 (pitch) хугацааны харьцуулалтыг харуулав.



Зураг 7. *Зүүн:* Эх (цэнхэр), зорилтот (улаан), хөрвүүлсэн (ногоон) аудионы waveform. Хөрвүүлсэн аудионы далайцын бүтэц болон үргэлжлэх хугацаа нь зорилтот яригчтэй ойролцоо хэвээр байгаа нь харагдана. *Баруун:* Гурван аудионы F0 (Pitch) хугацааны харьцуулалт. Эх яригчийн F0 дундаж ≈120 Гц байхад, хөрвүүлсэн ярианы F0 зорилтот яригчийн 210 Гц орчмын түвшинд амжилттай ойртсон (корреляци = 0.611) болохыг тодорхой харуулна.

Зураг 8-д нэг фреймийн spectral envelope болон зорилтот болон хөрвүүлсэн дуу хоёрын спектрал ялгааг (residual) харуулав.



Зураг 8. *Зүүн:* Нэг фреймийн spectral envelope (MCEP)-ийн харьцуулалт. Хөрвүүлсэн дуу (ногоон тасархай) нь эх яригч (цэнхэр)-аас зорилтот яригч (улаан) руу спектрал бүтцийн хувьд амжилттай шилжсэнийг харуулна. F1, F2, F3 формант давтамжийн дагуу хөрвүүлсэн ярианы онцлог зорилтот яригчтэй нийцэж байна. *Баруун:* Зорилтот болон хөрвүүлсэн дуу хоёрын спектрал ялгаа (residual). RMS ялгаа ≈2.2 дБ (MCD ≈7.43 дБ) бөгөөд WORLD вокодерын over-smoothing нөлөөгөөр доод давтамжид ялгаа харьцангуй өндөр байна.

B. Алдааны шинжилгээ

Хөрвүүлэлтийн явцад гарч буй гол алдаануудад: (1) нарийн аялгатай богино үеийн фонемүүдийн F0 алдаа нэмэгдэх (зураг 7-д харагдах MB_0007, MB_0011 дээрх хазайлт), (2) дуугүй авианы хил заагийг таньж олоход хэцүү тохиолдол зэрэг орно. Зураг 8-ийн баруун дэд зурагт харагдах доод давтамжийн (<1 кГц) ялгаа нь WORLD вокодерын over-smoothing нөлөөтэй холбоотой бөгөөд вокодерыг сайжруулснаар дараагийн алхамд бууруулах боломжтой.

Pitch (үндсэн давтамж) нь ярианы аялга, өнгийг илэрхийлдэг бөгөөд ихэвчлэн эгшиг авианы үед тодорхойлогддог. Харин intensity нь дохионы энерги буюу чанга сул байдлыг илэрхийлж, ярианы идэвхтэй болон идэвхгүй хэсгүүдийг ялгахад ашиглагдана. Доод хэсэгт байрлах фонемийн тэмдэглэгээ нь ярианы сегментчлэлийг илэрхийлж, спектрограмм дээрх акустик бүтэцтэй шууд хамааралтай байна. Туршилтын явцад гарсан waveform болон F0-ийн харьцуулалтыг зураг 7-д дэлгэрэнгүй харуулав.

IV. ДҮГНЭЛТ

Энэхүү судалгаанд авианы постериограмм (PPG)-д суурилсан дуу хоолой хувиргалтын системийг Монгол хэлэнд анх удаа амжилттай хэрэгжүүлж, объектив хэмжүүрүүдээр суурь аргыг тодорхой хэмжээгээр давсан болохыг баталлаа.

Объектив үр дүн: Санал болгосон PPG+DBLSTM систем нь 267 яригч бүхий Monnum өгөгдлийн санд туршиж, 2,462 бичлэг дээр сургасны үр дүнд мелкепстрийн гажилт (MCD) = 7.43 дБ, үндсэн давтамжийн корреляци = 0.611, дуутай/дуугүй нарийвчлал = 87.3% үзүүлэлтэд хүрлээ (зураг 6). Суурь параллель DBLSTM аргатай (MCD = 8.12 дБ, F0 корреляци = 0.573, V/UV = 84.1%) харьцуулахад MCD **8.5%**, F0 корреляци **6.6%**, V/UV **3.8%**-иар сайжирсан нь акустик дуу чанарын мэдэгдэхүйц ахицыг нотолж байна. Сургалтын анхны туршилтаас эцсийн хувилбар

хүртэл MCD 19.0 дБ-ээс 7.4 дБ болж **61%** буурсан нь загварын хөгжлийн явцыг баганан диаграммаар (зураг 6, баруун) тодорхой харуулна. Spectral envelope-ийн дүн шинжилгээ (зураг 8) болон waveform, F0 харьцуулалт (зураг 7)-аас хөрвүүлсэн аудио нь зорилтот яригчийн акустик шинж чанарт ойртсоныг нотолно.

Параллель өгөгдлийн хамаарлаас ангид байдал: Суурь арга нь эх болон зорилтот яригчийн параллель бичлэг шаарддаг бол санал болгосон систем нь зөвхөн зорилтот яригчийн 2,462 бичлэг ашигладаг тул өгөгдөл цуглуулах зардал, нөөц нэлээд хэмжээгээр буурсан. Энэ нь монгол хэл мэт бага хэмжээний зэрэгцээ корпустай хэлэнд ялангуяа чухал давуу тал болж байна.

Монгол хэлний суурь бүтэц: 36 фонемийн vocabulary бүхий rule-based G2P модуль болон MFA-д тохирсон лексиконыг нийтэд хүртээмжтэй болгосон нь Монгол VC болон TTS чиглэлийн цаашдын судалгааны нээлттэй суурь болох юм.

Цаашдын ажилд MOS (Mean Opinion Score) субъектив үнэлгээ хийх, олон шинэ яригчид туршилт өргөжүүлэх, мөн бодит цагийн (real-time) дуу хоолой хувиргалтын боломжийг судлах зорилт тавьж байна.

Ашигласан ном

[1] Y. Stylianou, O. Cappé, and E. Moulines, “Continuous probabilistic transform for voice conversion,” IEEE Transactions on Speech and Audio Processing, vol. 6, no. 2, pp. 131–142, 1998.

[2] T. Toda, A. W. Black, and K. Tokuda, “Voice conversion based on maximum-likelihood estimation of spectral parameter trajectory,” IEEE Transactions on Audio, Speech, and Language Processing, vol. 15, no. 8, pp.2222–2235, 2007.

[3] Z. Wu, T. Virtanen, T. Kinnunen, E. S. Chng, and H. Li, “Exemplar-based voice conversion using non-negative spectrogram deconvolution,” in Proc. 8th ISCA Speech Synthesis Workshop, 2013.

[4] T. Nakashika, R. Takashima, T. Takiguchi, and Y. Ariki, “Voice conversion in high-order eigen space using Deep Belief Nets,” in Proc. Interspeech, 2013

[5] L. Sun, S. Kang, K. Li, and H. Meng, “Voice conversion using deep bidirectional Long Short-Term Memory based Recurrent Neural Networks,” in Proc. ICASSP, 2015.

[6] D. Erro, A. Moreno, and A. Bonafonte, “INCA algorithm for training voice conversion systems from nonparallel corpora,” IEEE Transactions on Audio, Speech, and Language Processing, vol. 18, no. 5, pp. 944–953, 2010.

[7] J. Tao, M. Zhang, J. Nurminen, J. Tian, and X. Wang, “Supervisory data alignment for text-independent voice conversion,” IEEE Transactions on Audio, Speech, and Language Processing, vol. 18, no. 5, pp. 932–943, 2010.

[8] H. Silén, J. Nurminen, E. Helander, and M. Gabbouj, “Voice conversion for non-parallel datasets using dynamic kernel partial least squares regression,” Convergence, vol. 1, p. 2, 2013.

[9] H. Benisty, D. Malah, and K. Crammer, “Non-parallel voice conversion using joint optimization of alignment by temporal context and spectral distortion,” in Proc. ICASSP, 2014.

[10] S. Aryal and R. Gutierrez-Osuna, “Articulatory-based conversion of foreign accents with Deep Neural Networks,” in Proc. Interspeech, 2015.

[11] T. J. Hazen, W. Shen, and C. White, “Query-by-example spoken term detection using phonetic posteriorgram templates,” in Proc. ASRU, 2009.

[12] K. Kintzley, A. Jansen, and H. Hermansky, “Event selection from phone posteriorgrams using matched filters,” in Proc. Interspeech, 2011.

[13] M. Wollmer, Z. Zhang, F. Weninger, B. Schuller, and G. Rigoll, “Feature enhancement by bidirectional LSTM networks for conversational speech recognition in highly non-stationary noise,” in Proc. ICASSP, 2013.

[14] H. Kawahara, I. Masuda-Katsuse, and A. de Cheveigné, “Restructuring speech representations using a pitch-adaptive time–frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds,” Speech communication, vol. 27, no. 3, pp. 187–207, 1999.

[15] S. Imai, “Cepstral analysis synthesis on the mel fre- quency scale,” in Proc. ICASSP, 1983.

[16] Byambajav Dorj, “The MonNum speech databases,” open datasets, 2008-2012.

[17] Alexei Baevski, Henry Zhou, Abdelrahman Mohamed and Michael Auli, “wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations,” Oct. 2022.

[18] J. Garofolo, L. Lamel, W. Fisher, J. Fiscus, D. Pallett, N. Dahlgren, and V. Zue, “TIMIT acoustic-phonetic continuous speech corpus,” 1993.

[19] F. Weninger, J. Bergmann, and B. Schuller, “Introducing CURRENNT: the Munich open-source CUDA RecurREnt Neural Network Toolkit,” Journal of Machine Learning Research, vol. 16, pp. 547–551, 2015.

[20] Lifa Sun, Kun Li, Hao Wang, Shiyin Kang and Helen Meng, “Phonetic posteriorgrams for many-to-one voice conversion without parallel data training,” IEEE International Conference on Multimedia and Expo (ICME) , July, 2016.

[21] H. VALBRET, E. MOULINES, J.P. TUBACH, "VOICE TRANFORMATION USING PSOLA TECHNIQUE", International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 1992.

Монгол хэлний ярианаас бичвэр буулгах системийн нууцлалын алгоритмуудын гүйцэтгэлийн жишиг судалгаа

Эрдэнэболдын Дашренчин
Мэдээллийн Технологи, Электроникийн сургууль
Монгол Улсын Их Сургууль
Улаанбаатар хот, Монгол улс
23b1num1789@stud.num.edu.mn

Тэгшжаргалын Нямжаргал
Мэдээллийн Технологи, Электроникийн сургууль
Монгол Улсын Их Сургууль
Улаанбаатар хот, Монгол улс
20b1num2614@stud.num.edu.mn

Наранбаатарын Эгшиглэн
Мэдээллийн Технологи, Электроникийн сургууль
Монгол Улсын Их Сургууль
Улаанбаатар хот, Монгол улс
23b1num1728@stud.num.edu.mn

Цогтбаатарын Энхтуяа*
Мэдээллийн Технологи, Электроникийн сургууль
Монгол Улсын Их Сургууль
Улаанбаатар хот, Монгол улс
enkhtuya.ts@num.edu.mn

***Хураангуй*— Дуу хоолойг бичвэрт буулгах системүүд нь Монгол хэлний цахим шилжилтэд чухал үүрэг гүйцэтгэж байгаа ч өгөгдлийн нууцлал, аюулгүй байдлын асуудал тулгамдсаар байна. Энэхүү судалгаагаар бид 11-р үеийн Intel Core i5-1135G7 процессор болон PyCryptodome 3.23.0 орчинд AES-128, AES-256, ChaCha20 болон RSA-2048 алгоритмуудын гүйцэтгэлийг Монгол хэлний UTF-8 бичвэр өгөгдөл дээр харьцуулан хэмжлээ. Туршилтын үр дүнгээс үзэхэд бага хэмжээний өгөгдөлд ChaCha20 алгоритм хамгийн хурдан байсан бол, өгөгдөл томрох тусам AES-NI хурдасгуурын нөлөөгөөр AES-128 алгоритм илүү өндөр гүйцэтгэл үзүүлж байна. RSA-2048 алгоритм нь симметр алгоритмуудаас дунджаар 16,000–180,000 дахин удаан байгаа нь түүнийг бодит цагийн STT-д ашиглах боломжгүйг нотолж байна. Бид техникийн хоцрогдлыг ITU-T G.114 стандартын хэрэглэгчийн хүртэхүйн босготой харьцуулан, нууцлал нь бодит цагийн транскрипцийн чанарт сөрөг нөлөөгүй болохыг баталж, аюулгүй байдал болон гүйцэтгэлийн тэнцвэрийг тодорхойлсон болно.**

Түлхүүр үг— ярианаас бичвэр буулгах, нууцлалын алгоритм, AES, ChaCha20, RSA, гүйцэтгэлийн жишиг, хоцрогдол, дундаж, хугацаа

I. Удиртгал

Яриа болон дуу хоолой таних технологийн зах зээл 2023 онд 20.25 тэрбум ам.долларт хүрсэн бөгөөд 2030 он гэхэд 53.67 тэрбум ам.долларт хүрч, жил бүр дунджаар 14.6%-иар өсөх төлөвтэй байна [1]. Монгол Улсын "Цахим Үндэстэн" болох стратегийн зорилтын хүрээнд төрийн болон хувийн хэвшлийн байгууллагууд ярианаас бичвэр хөрвүүлэх системийг хурлын тэмдэглэл автоматжуулах, бичиг хэргийн бүтээмжийг нэмэгдүүлэхэд эрчимтэй ашиглаж эхлээд байна. Гэсэн хэдий ч эдгээр системүүд нь өгөгдлийг дамжуулалтын түвшинд нууцалдаг боловч аппликейшны түвшний төгсгөлөөс төгсгөл рүү нууцлал буюу End-to-End Encryption дутагдалтай байгаа нь мэдээллийн аюулгүй байдлын томоохон эрсдэлийг дагуулж байна. Монгол хэлний кирилл бичвэр нь UTF-8 кодчилолоор дамждаг бөгөөд латин тэмдэгтээс илүү их өгөгдлийн хэмжээ эзэлдэг нь бодит цагийн нууцлалын алгоритмын тооцооллын ачаалалд нөлөөлөх нэгэн хүчин зүйл болж болдог.

Өгөгдөл алдагдсанаас үүдэх дэлхийн дундаж хохирол 2023 оны байдлаар 4.4 сая ам.долларт хүрсэн нь мэдээллийн нууцлал, аюулгүй байдлын ач холбогдлыг илэрхийлж байна [2]. STT (Speech-to-Text) системүүд нь сервер талд өгөгдлийг ил бичвэр хэлбэрээр боловсруулж байгаа нь үйлчилгээ үзүүлэгч талын халдлагад өртөх эсвэл байгууллагын нууц мэдээлэл замаас алдагдах боломжийг бүрдүүлдэг. Гэвч аюулгүй байдлыг хангах үүднээс нууцлалын алгоритмууд нэмэх нь бодит цагийн системийн гүйцэтгэлд шууд нөлөөлдөг хоцрогдлыг үүсгэдэг.

ITU-T G.114 олон улсын стандартын дагуу бодит цагийн дуу хоолой болон бичвэр дамжуулахад нэг талын хоцрогдол 150ms-ээс бага байх нь хамгийн тохиромжтой бөгөөд 150-400ms-ийн хоцрогдлыг тусгай хэрэглээний хүрээнд зөвшөөрөгдөх түвшин гэж үздэг [3]. Аливаа хоцрогдол 400ms-ээс хэтэрсэн тохиолдолд харилцааны чанар муу гэж ангилагддаг ба хэрэглэгчийн системд итгэх итгэл болон хэрэглээний чадамж эрс буурдаг. Тиймээс энэхүү судалгаагаар бид Монгол хэлний STT системд тохирох нууцлалын алгоритмуудын гүйцэтгэлийг ITU-T G.114 стандартын хүрээнд жишиг судалгааны аргаар хэмжиж, гүйцэтгэл болон нууцлалын тэнцвэрт байдлыг тодорхойлохыг зорилоо.

II. Судалгаа

STT Технологийн Хөгжлийн Товч Түүх

Ярианаас бичвэр хөрвүүлэх технологи нь хэдэн арван жилийн хугацаанд эрс өөрчлөгдсөн. 1950-иад онд “Bell Labs”-ийн боловсруулсан анхдагч системүүд зөвхөн тоо таних чадвартай байсан бол 1970-80-аад онд статистикийн арга, ялангуяа “Hidden Markov Model” нь тархалтын загварт үндэслэсэн яриа таних технологийн гол суурь болсон юм [4]. 2010-аад оноос эхлэн гүн сургалтын аргачлал, ялангуяа давтагдагч нейрон сүлжээ (RNN) болон урт богино хугацааны ой (LSTM) ашигласан загварууд нь ASR (Automatic Speech

Recognition)-ийн нарийвчлалыг эрс дээшлүүлсэн. Сүүлийн жилүүдэд трансформер архитектур дээр суурилсан загварууд тэргүүлэх байр суурь эзэлж байна.

Kheddar et al. (2024)-ийн иж бүрэн судалгаа нь deep transfer learning, federated learning, болон reinforcement learning-д суурилсан ASR системүүдийг дэлгэрэнгүй авч үзсэн бөгөөд трансформер архитектурын оролтын дараалллын нарийн хамаарлыг олох чадварыг онцлон тэмдэглэсэн байна [4].

B. Орчин Цагийн STT Загварууд

OpenAI-ийн Whisper загвар нь интернэтээс цуглуулсан 680,000 цагийн олон хэлт болон олон даалгаврын аудио өгөгдөл дээр сургасан загвар бөгөөд ийм хэмжээний олон янзын өгөгдлийн ашиглалт нь дуудлага, арын дуу чимээ, мэргэжлийн нэр томьёоны хувьд бат бэх болохыг харуулсан [6]. “LibriSpeech” хэмжилтийн иж бүрдэлд Whisper загвар ямар нэг тохируулгагүй zero-shot нөхцөлд өрсөлдөхүйц үр дүн гаргасан нь том хэмжээний олон янзын урьдчилсан сургалтын хүчийг нотолж байна [5]. Whisper Large V3 Turbo (2024 оны 10-р сар)-д архитектурыг оновчтой болгосноор 5.4 дахин өссөн байна [6].

Wav2vec 2.0 загварын хувьд 53,000 цагийн шошгогүй аудио өгөгдөл дээр урьдчилан сургасан загварыг ердөө 10 минутын шошголсон өгөгдлөөр тохируулснаар LibriSpeech хэмжилтийн иж бүрдэлд тэргүүний үр дүн гарч байсан нь бага нөөцтэй хэл болон салбарт өндөр гүйцэтгэлтэй ASR хөгжүүлэх боломжийг нэмэгдүүлж байна [14]. STT API зах зээл нь 2024 онд 3.8 тэрбум ам.долларт хүрсэн бөгөөд 2030 он гэхэд 8.6 тэрбум ам.долларт хүрч жил бүр 14.4%-иар өсөх төлөвтэй байна [15].

C. Монгол Хэлний STT Технологийн Одоогийн Байдал

Монгол хэл нь бага нөөцтэй хэлний тодорхой жишээ бөгөөд end-to-end загварын сургалтад шаардагдах хангалттай шошголсон өгөгдөл дутагдалтай байдаг [17]. 2020 оны OpenASR20 сорилд NIST болон IARPA-ийн санаачилснаар Монгол хэл 10 бага нөөцтэй хэлний нэгд тооцогдсон бөгөөд тохиромжтой нөхцлийн хувьд WER (Word Error Rate) нь 0.4-0.65 хүрсэн нь тохиромжтой өгөгдлийн хомсдолыг тодорхой харуулж байна [10]. Li et al. (2025)-ийн судалгаанаас үзэхэд үгийн үед суурилсан дэд үгийн загварчлалыг ашигласан хагас хяналттай сургалтын арга нь шошгогүй 1,560 цагийн өгөгдлийг ашиглан Монгол ярианы таних системийн гүйцэтгэлийг мэдэгдэхүйц дээшлүүлсэн байна [8].

Монгол хэлний кирилл бичвэр нь агглютинаци (agglutinative) шинж чанартай бөгөөд UTF-8 кодчилолоор дамжих тул латин үсгийн системтэй харьцуулахад өгөгдлийн хэмжээ их байдаг нь нууцлалын тооцооллын ачаалалд нэмэлт нөлөөлнө. Монгол хэлний нийт хэрэглэгч дэлхий даяар 10 гаруй саяд хүрдэг хэдий ч STT болон TTS (Text-to-Speech)-д зориулсан нийтийн эх үүсвэр хомс байдаг нь одоо хүртэл үргэлжилсэн хязгаарлалт хэвээр байна [9]. TTS болон STT загварууд иж бүрэн тойм нь дүрмийн системд суурилсан системүүдээс гүн сургалтын архитектур хүртэлх хөгжлийг тусгасан бөгөөд туслах технологиос эхлэн виртуал туслагч, хэл орчуулах зэрэг олон хэрэглээний чиглэлийг хамардаг [7].

D. STT Систем дэх Нууцлалийн Асуудлын Одоогийн Байдал

Олон тооны судалгаа STT технологийн нууцлал, аюулгүй байдлын асуудлыг авч үзсэн байна. Мэдээллийн аюулгүй байдлын судлаачид STT систем дэх хоёр төрлийн нууцлалыг ялгаж тэмдэглэдэг: дамжуулалтын явцад болон хадгалалтын үеийн нууцлал [11]. Шилжүүлгийн явцад TLS 1.3 болон AES-256 ашиглан шифрлэлтийг хэрэгжүүлэхийг зөвлөдөг боловч эдгээр хоёр арга нь аппликейшны түвшинд STT-ийн боловсруулалт явагдах үед өгөгдлийн аюулгүй байдлыг бүрэн хангаж чаддаггүй гэх дутагдал байдаг[18].

Fortanix-ийн судалгаанаас харахад STT систем нь сервер талд өгөгдлийг ил бичвэр хэлбэрээр боловсруулдаг бөгөөд энэ нь үйлчилгээ үзүүлэгч талын дотоод халдлага болон RAM-аас өгөгдөл задруулах аюул дагуулдаг. Энэ асуудлыг нууцлагдмал тооцооллоор шийдвэрлэх боломж байгаа ч практик хэрэгжилт хязгаарлагдмал хэвээр байна [13]. Ийм нөхцөлд нууцлалын алгоритм сонголт нь зөвхөн аюулгүй байдлын биш, гүйцэтгэлийн тэнцвэрт байдлын асуудал болдог. Хэдийгээр STT нууцлалын тухай олон хэвлэлд хэлэлцдэг ч алгоритм бүрийн хоцрогдлыг STT ажлын ачааллын дагуу нарийвчлан хэмжсэн туршилтын судалгаа цөөн байдаг нь энэхүү судалгааны гол дутуу тал юм[11], [12].

E. Судлагдаагүй тал

Одоо хүртэлх судалгааны тоймуудад дараах судлагдаагүй асуудлууд харагдаж байна: (1) Монгол кирилл бичвэрийн UTF-8 байт хэмжээний онцлог нь нууцлалын ачааллыг нэмэгдүүлдэг боловч энэ нь алгоритм бүрт хэрхэн ялгаатай нөлөөлдгийг тоон хэмжилтээр нотолсон судалгаа дутмаг байна; (2) STT системийн бодит цагийн ажиллагааны нөхцөлд AES болон ChaCha20-ийн хоорондох хоцрогдлын огтлолцлын цэгийг өгөгдлийн хэмжээний функц хэлбэрээр тодорхойлсон судалгаа байхгүй байна; (3) Монгол хэлний STT-д зориулсан нууцлалын гүйцэтгэлийн харьцуулалт бүрэн судлагдаагүй хэвээр байна; (4) On-device буюу төхөөрөмж дотоод нууцлал болон клауд дахь нууцлалын хоорондох гүйцэтгэлийн ялгааг практик STT хэрэглээнд хэмжсэн судалгаа хомс байна [16], [17]. Мөн нууцлалын алгоритмуудын гүйцэтгэл нь зөвхөн алгоритмын архитектураас гадна орчин үеийн процессоруудын техник хангамжийн AES-NI хурдасгуур болон программ хангамжийн оновчлолоос шууд хамаардаг. Тиймээс эдгээр хүчин зүйлс Монгол хэлний STT системд хэрхэн нөлөөлж буйг бодит туршилтаар баримтжуулах шаардлагатай байна.

Энэхүү судалгаа нь дээрх асуудлыг судлахад чиглэсэн бөгөөд Монгол хэлний STT системийн хувьд хамгийн тохиромжтой шифрлэлтийн алгоритмыг тодорхойлохыг зорьж байна.

F. Нууцлалын алгоритмуудын гүйцэтгэлийн харьцуулалт

Нууцлалын алгоритмуудын тооцооллын нөөц болон хурдыг хэмжсэн олон судалгаа байдаг. Үүний нэг жишээ

болох A.Elminaam болон бусад судлаачдын ажилд AES, DES, 3DES, RC2, Blowfish, болон RC6 зэрэг зургаан төрлийн тэгш хэмт алгоритмыг өгөгдлийн хэмжээ, төрөл, түлхүүрийн урт болон эрчим хүчний зарцуулалт зэрэг олон үзүүлэлтээр харьцуулан судалсан байна [18]. Уг судалгааны үр дүнгээс үзэхэд:

- Түлхүүрийн урт ихсэх тусам системийн тооцоолох хугацаа болон төхөөрөмжийн батерей зарцуулалт шууд хамааралтайгаар нэмэгддэг.
- Бичвэр болон зураг зэрэг өгөгдлийн төрлөөс хамаарч алгоритмуудын гүйцэтгэл өөр өөр байдаг.
- Blowfish алгоритм нь багц өгөгдлийн өөрчлөлтөд хамгийн тогтвортой, өндөр гүйцэтгэл үзүүлж байгааг тус тус тогтоожээ.

Гэсэн хэдий ч, энэхүү судалгаа нь:

- Судалгаа 2010 оны байдлаар хийгдсэн тул орчин үеийн, илүү хурдтай ChaCha20 алгоритмыг багтаагаагүй.
- Туршилтууд нь ерөнхий өгөгдөл дээр хийгдсэн бөгөөд UTF-8 кодчилол бүхий бичвэр өгөгдөлд алгоритмууд хэрхэн нөлөөлөхийг тусгаагүй.
- Зөвхөн тэгш хэмт алгоритмуудыг судалсан бөгөөд бодит цагийн транскрипцид тэгш хэмт бус алгоритм ямар хэмжээний хоцрогдол үүсгэхийг харьцуулаагүй байна.

G. ChaCha20 алгоритмын давуу тал

Тэгш хэмт урсгал шифр болох ChaCha20 алгоритм нь орчин үеийн кибер аюулгүй байдлын салбарт, ялангуяа техник хангамжийн тусгай дэмжлэггүй төхөөрөмжүүдэд зориулсан оновчтой шийдэл болдог. Daniel J. Bernstain-ий боловсруулсан энэхүү алгоритм нь зөвхөн нэмэх, XOR болон эргүүлэх гэсэн энгийн арифметик үйлдлүүд дээр суурилдаг тул программ хангамжийн түвшинд маш өндөр гүйцэтгэл үзүүлдэг [22].

N. Nir болон бусад судлаачдын ажилд [23] дурдсанаар, ChaCha20 нь AES алгоритмын адилаар техник хангамжийн хурдасгуур шаарддаггүй бөгөөд "side-channel" халдлагад тэсвэртэй байхаар зохион бүтээгдсэн. Google болон Cloudflare зэрэг компаниуд мобайл төхөөрөмж болон IoT системүүдэд зориулан ChaCha20-ийг TLS протоколын үндсэн алгоритм болгон ашиглаж байгаа нь түүнийг бодит цагийн өгөгдлийн урсгалыг бага нөөцөөр, хурдан нууцлахад хамгийн тохиромжтой болохыг харуулж байна.

H. Техник хангамжийн хурдасгуур ба програм хангамжийн оновчлол

Нууцлалын алгоритмуудын гүйцэтгэл нь зөвхөн алгоритмын математик загвараас гадна процессорын зааварчилгааны багцаас ихээхэн хамаардаг. S. Gueron болон бусад судлаачдын ажилд [24] дурдсанаар Intel процессоруудын AES-NI (Advanced Encryption Standard New Instructions) зааварчилгаа нь шифрлэлтийн процессыг техник хангамжийн түвшинд гүйцэтгэснээр программ хангамжийн түвшний гүйцэтгэлийг 3-10 дахин нэмэгдүүлдэг болохыг тогтоосон.

Түүнчлэн, программчлалын хэлний сангуудын сонголт нь алгоритмын гүйцэтгэлд томоохон үүрэгтэй байдаг. PyCryptodome сан нь C хэл дээр бичигдсэн доод түвшний функцүүдийг ашигладаг бөгөөд процессорын AES-NI зааварчилгааг автоматаар илрүүлж, оновчтой ашиглах боломжийг олгодог. Орчин үеийн 11-р үеийн процессорууд дээрх эдгээр техник хангамжийн хурдасгуурууд нь их хэмжээний өгөгдлийг бодит цагт, сааталгүй боловсруулахад том үүрэг гүйцэтгэж байна.

III. Туршилтын арга зүй

A. Туршилтын орчин

Туршилтыг Хүснэгт I-д үзүүлсэн тохиргоо бүхий физик компьютер дээр гүйцэтгэв.

ХҮСНЭГТ I.
Туршилт гүйцэтгэсэн компьютер болон программын орчны үзүүлэлтүүд

Үзүүлэлт	***Утга***
CPU	11th Gen Intel Core i5-1135G7 @ 2.40GHz
RAM	15.74 GB (16,901,771,264 байт)
Үйлдлийн систем	Microsoft Windows 11 Pro (Build 26200)
Программчлалын хэл	Python 3.13.8
Криптограф сан	PyCryptodome 3.23.0
Хэмжилтийн функц	time.perf_counter()

Туршилтын өгөгдөл

Туршилтын өгөгдөл болгон бодит ярианы аудио бичлэгийг ашиглав. Нийт гурван өөр уртын MP3 файлыг Google Speech Recognition API-ийн тусламжтайгаар Монгол хэлээр транскрипц хийж, UTF-8 форматаар кодчилсон байт мөрийг шифрлэлтийн туршилтын оролтын өгөгдөл болгон ашиглав. Аудио файлын хэт урттай хэсгийг боловсруулахын тулд файл бүрийг 55 секундын chunk-уудад хуваан дараалан таньсны дараа нэгтгэв. Туршилтын файлуудын үзүүлэлтийг Хүснэгт II-д нэгтгэв.

ХҮСНЭГТ II. Туршилтын аудио файлуудын үзүүлэлт

Аудио урт	***Чанк тоо***	***Тэмдэгт***	***Байт хэмжээ***
59.0 сек (0.98 ~ 1 мин)	1	175	323 байт
298.7 сек (4.98 ~ 5мин)	6	1,420	2,611 байт
538.1 сек (8.97 ~ 9мин)	10	1,836	3,402 байт

5 минутын файлын 4-р болон 6-р chunk-д ойлгомжтой дуу байгаагүй тул тэдгээр хэсгийг хоосон текстээр бүртгэн үлдсэн chunk-уудын транскриптийг нэгтгэв. Энэ нь STT системийн бодит нөхцлийг тусгасан бөгөөд аудиорологийн чанар дунджаас доош байх үед транскриптийн бүрэн бүтэн байдалд нөлөөлдгийг харуулж байна.

B. Туршилтанд ашиглах алгоритмууд

Судалгаанд дараах дөрвөн шифрлэлтийн алгоритмыг сонгов. Сонголтын үндэслэл нь практик хэрэглээнд өргөн тархсан байдал болон алгоритмын онолын ялгаатай байдал юм.

AES-128-CBC. 128 битийн түлхүүртэй, блокийн шифрлэлтийн CBC горимыг ашигласан. [20] CBC горимд *i*-р шифртекст блок дараах томьёогоор тодорхойлогдоно:

$$Ci=EK(Pi \oplus Ci-1), C0=IV \qquad (1)$$

Энд P_i нь нээлттэй текстийн блок, *IV* нь санамсаргүй үүсгэсэн 16 байтын эхлэлийн вектор, E_K нь K түлхүүртэй шифрлэлтийн функц юм. PKCS7 padding ашигласан.

AES-256-CBC. 256 битийн түлхүүртэй, AES-128 загвартай ижил CBC горим. Түлхүүрийн уртын нэмэгдэлт нь key schedule-н давталтын тоог 10-аас 14 болгон нэмэгдүүлдэг. [20]

ChaCha20. 256 битийн түлхүүртэй урсгалын шифрлэлтийн алгоритм. [21] Padding шаарддаггүй. Keystream:

$$keystream=ChaCha20Block(key, nonce, counter) \quad (2)$$

Энд *nonce* нь санамсаргүй үүсгэсэн 8 байт, *counter* нь блокийн дугаар юм.

RSA-2048-OAEP. 2048 битийн нийтийн түлхүүрийн асимметр шифрлэлт. [19] PKCS1-OAEP дэвтрийн схем ашигласан:

$$C=m^e \bmod n \quad (3)$$

Энд *m* нь нээлттэй текст, *e* нь нийтийн зэрэглэл, *n* нь 2048 битийн модуль. RSA-н блокийн хэмжээний хязгаарлалтаас болж туршилтад өгөгдлийн эхний 100 байтыг ашигласан.

C. Хэмжилтийн арга зүй

Хэмжилтийн найдвартай байдлыг хангах зорилгоор зургаан аргыг хослуулан ашиглав. Давталтын тоог файлын уртаас хамааруулан тогтоосон нь Хүснэгт III-д үзүүлэв.

ХҮСНЭГТ III. Давталтын тооны тохиргоо

Файл	***Симметр алгоритм(давтсан тоо)***	***RSA-2048***	***Үндэслэл***
1мин (323 байт)	500	10	Жижиг өгөгдөл, давталт их байж болно
5мин (2,611 байт)	100	5	Дунд хэмжээний өгөгдөл, давталтын тоо дундаж
9мин (3,402 байт)	50	3	Их хэмжээний өгөгдөл, RSA алгоритм их хэмжээний өгөгдөлд өртөг өндөр тул давталтын тоо бага

Арга 1) Олон давталтын статистик шинжилгээ: time.perf_counter() функцийг ашиглан давталт бүрийн шифрлэлтийн хугацааг хэмжиж дундаж, медиан, хамгийн бага, хамгийн их болон стандарт хазайлтыг тооцов:

$$\overline{T} = \frac{1}{N}\sum_{i=1}^{N}\left(t_{end,i} - t_{start,i}\right) \quad (4)$$

Outlier-н нөлөөгөөс хамгаалах зорилгоор дундажийн зэрэгцээ медиан утгыг гол хэмжүүр болгон ашигласан нь судалгааны үр дүнгийн найдвартай байдлыг нэмэгдүүлэв.

```python
#  АРГА 1 — Олон давталт + статистик

def method1_statistics(algos: dict, data: bytes, sym_runs: int, rsa_runs: int) -> dict:
    print(f"\n[АРГА 1]  Олон давталт + статистик")
    print(f"  Симметр: {sym_runs} давталт  |  RSA: {rsa_runs} давталт")
    print(f"  {'Алгоритм':10}  {'Дундаж':>12}  {'Медиан':>12}  "
          f"{'Min':>12}  {'Max':>12}  {'Std':>12}")
    print("  " + "-" * 65)

    result = {}
    for name, (func, n) in algos.items():
        times = []
        for _ in range(n):
            t0 = time.perf_counter()
            func(data)
            times.append(time.perf_counter() - t0)

        dev = statistics.stdev(times) if len(times) > 1 else 0.0
        m   = statistics.mean(times)
        med = statistics.median(times)
        result[name] = {
            "mean": m, "median": med,
            "min":  min(times), "max": max(times),
            "stdev": dev, "n": n, "raw": times,
        }
        print(f"  {name:10}  {m:>12.8f}  {med:>12.8f}  "
              f"{min(times):>12.8f}  {max(times):>12.8f}  {dev:>12.8f}  сек")

    return result
```

Зураг 1. Олон давталтын статистик шинжилгээ боловсруулах код

Арга 2) timeit хэмжилт: Python-н timeit модуль нь garbage collector-г түр зогсоодог тул эхний аргатай харьцуулах бие даасан баталгаажуулалт болсон.

```python
#  АРГА 2 — timeit

def method2_timeit(algos: dict, data: bytes) -> dict:
    print(f"\n[АРГА 2]  timeit")
    print(f"  {'Алгоритм':10}  {'Нийт (с)':>12}  {'Дундаж (с)':>14}  {'Давталт':>8}")
    print("  " + "-" * 50)

    result = {}
    for name, (func, n) in algos.items():
        total = timeit.timeit(lambda f=func: f(data), number=n)
        result[name] = {"total": total, "mean": total / n, "n": n}
        print(f"  {name:10}  {total:>12.4f}  {total/n:>14.8f}  {n:>8}")

    return result
```

Зураг 2.
Timeit нь python-ы стандарт хугацаа хэмжигч бөгөөд нэг код хэсгийг олон удаа ажиллуулж дундаж хугацааг тооцоолдог.

Арга 3) cProfile функцийн задаргаа: cProfile модулийг ашиглан алгоритм бүрийн дотоод функц дуудалтын тоо болон cumulative time-ийг тодорхойлж хоцрогдлын үндсэн шалтгааныг функцийн түвшинд шинжлэв.

```python
#  АРГА 3 — cProfile

def method3_cprofile(data: bytes) -> None:
    print(f"\n[АРГА 3]  cProfile (нэг дамжуулалт)")
    profiler = cProfile.Profile()
    profiler.enable()
    run_all(data)
    profiler.disable()
    profiler.print_stats(sort="cumulative")
```

Зураг 3. `cProfile` нь Python-ы стандарт гүйцэтгэлийн профайлер кодын аль хэсэг хэр удаан ажиллаж байгааг нарийн хэмждэг.

Арга 4) Санах ойн хэрэглээний хэмжилт: memory_profiler сангийн *memory_usage()* функцийг interval=0.0001 параметртэйгээр ашиглан алгоритм бүрийн санах ойн дээд хэрэглээг MiB-ээр хэмжив.

```python
#  АРГА 4 — Санах ойн хэрэглээ

def method4_memory(algos: dict, data: bytes) -> dict:
    print(f"\n[АРГА 4]  Санах ойн хэрэглээ (Memory Profiler)")
    print(f"  {'Алгоритм':10}  {'Дээд хэрэглээ (MiB)':>20}")
    print("  " + "-" * 34)

    result = {}
    for name, (func, _) in algos.items():
        try:
            mem = memory_usage(
                (func, (data,)),
                interval=0.0001, max_usage=True, multiprocess=False
            )
            val = mem[0] if isinstance(mem, list) else mem
            result[name] = val
            print(f"  {name:10}  {val:>20.4f}")
        except Exception as e:
            print(f"  {name:10}  хэмжих боломжгүй — {e}")
            result[name] = None

    return result
```

Зураг 4. Санах ойн хэрэглээг хэмжих функцийн код

Арга 5) Өгөгдлийн хэмжээний нөлөөний шинжилгээ: Файл бүрийн транскрипцын бодит хэмжээг ашиглан 64 байтаас эхлэн бүтэн хэмжээ хүртэлх таван хэрчим дээр 50 давталтын дундажийг тооцов.

```
# ── АРГА 5 — Өгөгдлийн хэмжээ vs Саатал ──

def method5_size_vs_latency(data: bytes, sym_runs: int) -> dict:
    print(f"\n[АРГА 5]  Өгөгдлийн хэмжээний нөлөө")
    half  = max(64, len(data) // 2)
    sizes = sorted(set([64, 256, 1_024, half, len(data)]))

    print(f"  {'Хэмжээ':>10}  {'AES-128 (Ё)':>14}  "
          f"{'AES-256 (Ё)':>14}  {'ChaCha20 (Ё)':>14}")
    print("  " + "─" * 58)

    result = {}
    for sz in sizes:
        if sz <= len(data):
            chunk = data[:sz]
        else:
            rep   = (sz // len(data)) + 1
            chunk = (data * rep)[:sz]

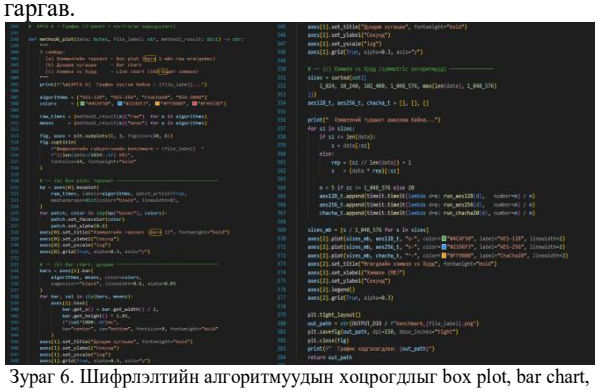
        t128 = timeit.timeit(lambda c=chunk: run_aes128(c),    number=sym_runs) / sym_runs
        t256 = timeit.timeit(lambda c=chunk: run_aes256(c),    number=sym_runs) / sym_runs
        tcha = timeit.timeit(lambda c=chunk: run_chacha20(c), number=sym_runs) / sym_runs
        result[sz] = {"aes128": t128, "aes256": t256, "chacha20": tcha}

        label = (f"{sz/1024:.0f}KB" if sz >= 1024 else f"{sz}B")
        print(f"  {label:>10}  {t128:>14.8f}  {t256:>14.8f}  {tcha:>14.8f}")

    return result
```

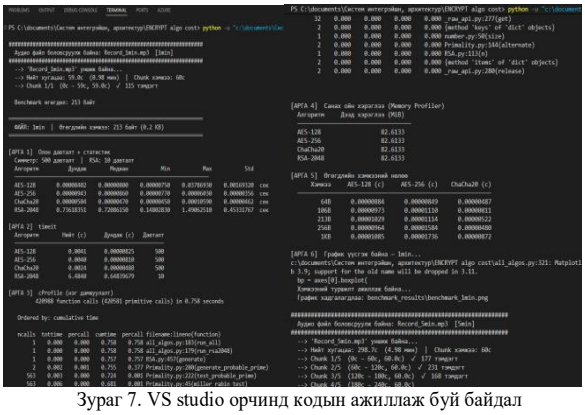
Зураг 5. Өгөгдлийн хэмжээ болон түүний саатлын шинжилгээг хийх код

Арга 6) График дүрслэл: Matplotlib ашиглан файл бүрт (a) log масштабын box plot, (b) медиан хоцрогдлын bar chart, (c) өгөгдлийн хэмжээ ба хоцрогдлын хамаарлын шугаман графикийг үүсгэж нэгтгэсэн харьцуулалтын графикийг гаргав.



Зураг 6. Шифрлэлтийн алгоритмуудын хоцрогдлыг box plot, bar chart, шугаман графикаар харуулна.

IV. Туршилтын үр дүн



Зураг 7. VS studio орчинд кодын ажиллаж буй байдал

A. Симметр алгоритмуудын хоцрогдол - Файлын Уртаар

Арга 1-н 500, 100, 50 давталтын медиан хугацааг гурван файл дээр хэмжсэн үр дүнг Хүснэгт IV-д нэгтгэв.

ХҮСНЭГТ IV. Арга 1 - Медиан latency файлын уртаар (секунд)

Алгоритм	***1мин (323Б)***	***5мин (2,611Б)***	***9мин (3,402Б)***
AES-128	0.00000830	0.00001355	0.00001925
AES-256	0.00000820	0.00001390	0.00002155
ChaCha20	0.00000520	0.00001180	0.00002130
RSA-2048	1.48145990	0.34343540	0.30343140

Симметр гурван алгоритм бүгд өгөгдлийн хэмжээ нэмэгдэхийн хэрээр хоцрогдол шугаман хандлагаар өсч байна. 323 байтаас 3,402 байт болоход AES-128 нь 0.00000830-аас 0.00001925 секунд болж 2.32 дахин, AES-256 нь 2.63 дахин, ChaCha20 нь 4.10 дахин нэмэгдсэн нь ChaCha20 урсгалын шифрлэлтийн алгоритм өгөгдлийн хэмжээнд илүү мэдрэмтгий болохыг харуулж байна.

ХҮСНЭГТ V. Арга 1 - Дундаж болон стандарт хазайлт (секунд)

Алгоритм	***Файл***	***Дундаж***	***Медиан***	***Хамгийн бага***	***Хамгийн их***	***Стандархазайлт***
AES-128	1мин	0.00005309	0.0000083	0.0000078	0.0022034	0.00031033
AES-128	5мин	0.00001913	0.0000135	0.0000128	0.000173	0.00002307
AES-128	9мин	0.00002438	0.0000192	0.0000187	0.0002267	0.00002934
AES-256	1мин	0.00000907	0.0000082	0.0000079	0.000037	0.00000416
AES-256	5мин	0.00001717	0.0000139	0.0000131	0.0000549	0.00000976
AES-256	9мин	0.00002239	0.0000215	0.0000149	0.0000799	0.00001016
ChaCha20	1мин	0.00000668	0.0000052	0.0000048	0.0000665	0.00000875
ChaCha20	5мин	0.00001321	0.0000118	0.0000105	0.000093	0.00001155
ChaCha20	9мин	0.00002598	0.0000213	0.0000186	0.000229	0.00002941
RSA-2048	1мин	0.95159264	1.4814599	0.0479254	1.5036934	0.74050517
RSA-2048	5мин	0.57933564	0.3434354	0.1806431	1.4437789	0.52555846
RSA-2048	9мин	0.63346194	0.3034314	0.1750364	1.6186887	0.60216982

1 минутын файлын AES-128-н дундаж (0.00005309 сек) болон медиан (0.00000830 сек) хоорондын зөрүү нь 6.4 дахин байгаа нь системийн арын процессуудаас үүдсэн outlier утгуудын нөлөөг тодорхой илтгэж байна. Хамгийн их утга (0.00220340 сек) нь медианаас 265 дахин их байсан ба энэ нь CPU-н дулааны хязгаарлалт болон OS-н хуваарилагчын нөлөөтэй холбоотой байж болзошгүй юм. Тиймээс дүгнэлтэд дунджийн оронд медиан утгыг гол хэмжүүр болгон ашиглав.

B. Timeit Харьцуулалт

Арга 2-н timeit хэмжилтийн үр дүнг Хүснэгт VI-д үзүүллээ.

ХҮСНЭГТ VI. Арга 2 - timeit дундаж хоцрогдол (секунд)

Алгоритм	***1мин***	***5мин***	***9мин***
AES-128	0.00001088	0.00001296	0.00001634
AES-256	0.00000987	0.00001392	0.00001598
ChaCha20	0.00000618	0.00001194	0.00001474
RSA-2048	0.76955352	0.73743288	0.53400850

timeit дундаж хоцрогдолыг дараах томьёогоор тооцоолов:

$$\overline{T}_{timeit} = \frac{T_{нийт}}{N} \quad (5)$$

RSA-2048-ийн timeit дундаж (0.76955 сек) нь Арга 1-ийн дундаж (0.95159 сек)-аас бага байгаа нь timeit-ийн garbage collector-ийг зогсоох механизмын нөлөөтэй холбоотой байж болох боловч үндсэн шалтгаан нь RSA-

ийн түлхүүр үүсгэлтийн санамсаргүй шинж чанараас үүдэлтэй өндөр хэлбэлзэл (Ст. хаз. = 0.74050517 сек) болон давталтын тооны хязгаарлагдмал байдал (N = 5) юм. Тиймээс RSA-2048-ийн хувьд медиан утгыг гол хэмжүүр болгон ашиглах нь зөв бөгөөд timeit болон Арга 1-ийн үр дүнг шууд харьцуулах нь статистикийн хувьд хангалттүй байна.

C. cProfile Функцийн Задаргаа

Файл бүрийн cProfile-н нэг дамжуулалтын хамгийн их cumtime бүхий функцуудыг Хүснэгт VII-д харуулав.

ХҮСНЭГТ VII. Арга 3 - cProfile дээд 5 функц cumtime-аар

	1мин		5мин		9мин	
Функц	***Дуудалтын тоо***	***Нийт хугацаа (с)***	***Дуудалтын тоо***	***Нийт хугацаа (с)***	***Дуудалтын тоо***	***Нийт хугацаа (с)***
run_rsa2048	1	0.588	1	0.937	1	1.787
RSA.generate	1	0.586	1	0.936	1	1.785
generate_probable_prime	2	0.584	2	0.934	2	1.783
miller_rabin_test	413	0.518	688	0.85	1,330	1.661
inplace_pow	827	0.492	1,411	0.806	2,706	1.576

RSA-2048-н нийт cumtime-н 83.7–88.2%-г *inplance_pow* буюу (modular exponentiation) үйлдэл эзэлж байна. Miller-Rabin тест файлын хэмжээ нэмэгдэхэд 413-аас 1,330 хүртэл дуудагдсан нь RSA-н санамсаргүй байдлаас шалтгаалсан тооцооллын зардлын хэлбэлзэлтэй нийцэж байна. Харин AES-128, AES-256, ChaCha20 гурван алгоритм нийлсэн cumtime нь гурван файл дээр тус бүр 0.001 секундаас бага байсан бөгөөд энэ нь RSA-2048-тай харьцуулшгүй хэмжээний тооцооллын зардлын ялгааг функцийн түвшинд баталж байна.

D. Санах Ойн Хэрэглээ

Арга 4-ийн санах ойн хэмжилтийн үр дүнг Хүснэгт VIII-д үзүүллээ.

Хүснэгт VIII. Арга 4 - Санах ойн дээд хэрэглээ (MiB)

Алгоритм	***1мин***	***5мин***	***9мин***
AES-128	79.9	116.31	106.95
AES-256	79.9	116.31	106.95
ChaCha20	79.9	116.31	106.95
RSA-2048	79.9	116.31	106.95

Дөрвөн алгоритм бүгд файл бүрт ижил санах ойн хэрэглээг үзүүлсэн нь *memory_profiler* хэмжилтийн нарийвчлалын хязгаарлалттай холбоотой бөгөөд 0.0001 секундын интервал алгоритм хоорондын ялгааг ялгаруулахад хангалттүй байсан байж болзошгүй юм. 5 минутын файлд санах ойн хэрэглээ (116.31 MiB) 1 минутынхаас (79.90 MiB) 45.6% өндөр байгаа нь транскрипц боловсруулалтын явцад санах ойд ачаалал нэмэгдсэнтэй холбоотой.

E. Өгөгдлийн Хэмжээний Нөлөө

Арга 5-н 50 давталтын дундаж хоцрогдлыг файл бүрийн бодит хэмжээн дээр хэмжсэн үр дүнг Хүснэгт IX-д үзүүллээ.

ХҮСНЭГТ IX. Арга 5 - Өгөгдлийн хэмжээ ба дундаж хоцрогдол (секунд)

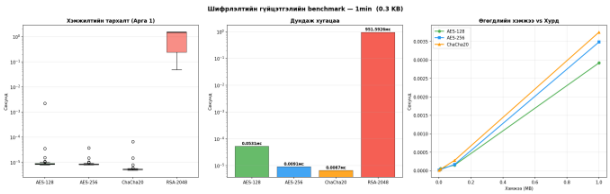
Файл	***Хэмжээ***	***AES-128***	***AES-256***	***ChaCha20***
Record_1min	64B	0.0000167	0.00001202	0.00000589
Record_1min	161B	0.00001019	0.00000898	0.00000553
Record_1min	256B	0.00001024	0.00001029	0.00000554
Record_1min	323B	0.00001009	0.00000947	0.00000572
Record_5min	64B	0.00001059	0.00000853	0.0000053
Record_5min	256B	0.00001006	0.00000871	0.00000515
Record_5min	1024B	0.00001251	0.00001225	0.00000967
Record_5min	2611B	0.00001441	0.00001423	0.00001246
Record_9min	64B	0.00001079	0.00000892	0.00000525
Record_9min	256B	0.00001	0.00000836	0.00000492
Record_9min	1024B	0.00001135	0.00001349	0.00000837
Record_9min	2048B	0.00001267	0.0000121	0.00001265
Record_9min	3402B	0.00001774	0.00002151	0.00001447

Гурван файл дээр нийтлэг хэв маяг ажиглагдав. 64 байтын жижиг өгөгдөл дээр ChaCha20 хамгийн бага хоцрогдол үзүүлсэн (дунджаар 0.00000548 сек). Өгөгдлийн хэмжээ 2,048 байтаас дээш болоход 9 минутын файлын 3,402 байтад AES-256 (0.00002151 сек) ChaCha20 (0.00001447 сек)-аас удаан болж байна. Энэ огтлолцол нь Intel AES-NI аппаратын хурдасгуурын нөлөөгүй нөхцөлд AES-н блокийн padding-н зардал урт өгөгдөлд илүү мэдэгдэхүйц болдогтой холбоотой байж болзошгүй.

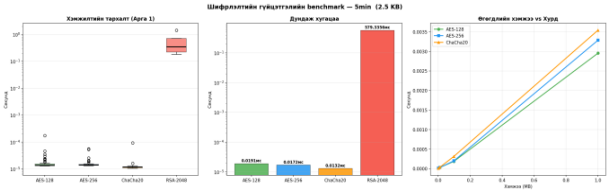
F. Нэгтгэсэн Харьцуулалт ба Дүн Шинжилгээ

Хүснэгт X. Нэгтгэсэн харьцуулалт - медиан хоцрогдол (мкс)

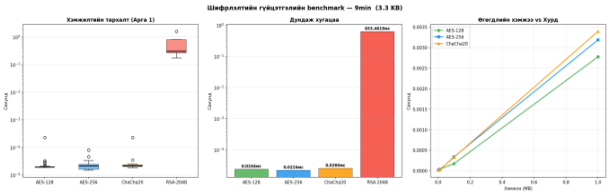
Алгоритм	***1мин (323B)***	***5мин (2,611B)***	***9мин (3,402B)***	***Өсөлт (1мин→9мин)***
AES-128	8.30 мкс	13.55 мкс	19.25 мкс	131.90%
AES-256	8.20 мкс	13.90 мкс	21.55 мкс	162.80%
ChaCha20	5.20 мкс	11.80 мкс	21.30 мкс	309.60%
RSA-2048	1,481,459.90 мкс	343,435.40 мкс	303,431.40 мкс	−79.5%



Зураг 8. Шифрлэлтын гүйцэтгэлийн харьцуулалт (1 мин, 0.3 кб) Хэмжилтийн тархалтаас харахад AES-128 нь хамгийн их хэлбэлзэлтэй, RSA-2048 нь симметр алгоритмуудаасаа 10^4 дахин их хоцрогдолтой байна. Дундаж хугацааны хувьд RSA-2048 нь хамгийн удаан буюу ChaCha20-оос 142,028 дахин их хугацаа зарцуулж байгааг тодорхойлов. Өгөгдлийн хэмжээ ихсэхэд бүх алгоритмууд шугаман өсөлт үзүүлсэн бөгөөд 0.2 MB-аас дээш хэмжээнд ChaCha20 нь AES-256-аас илүү удаашрах хандлага ажиглагдлаа.

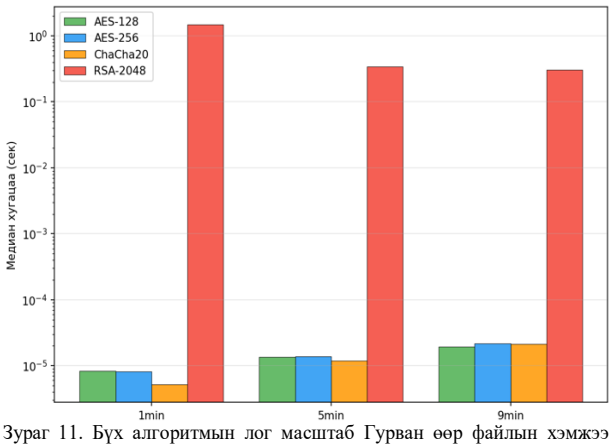


Зураг 9. Хэмжилтийн тархалтаас харахад 5 минутын файл дээр AES-128 болон RSA-2048 алгоритмуудын outlier цэгүүд буурч, тархалт илүү тогтвортой, нягт болсон байна. Дундаж хугацааны хувьд RSA-2048 нь 579.3356 мс зарцуулж байгаа нь хамгийн хурдан үзүүлэлттэй ChaCha20-оос 43,889 дахин удаан байгаар харууллаа. Өгөгдлийн хэмжээ ихсэх тусам ChaCha20-н өсөлтийн хурд AES-256-тай ойртож, 0.6 MB-аас дээш хэмжээнд ChaCha20 нь AES-256-аас илүү удаашрах хандлагатай байна.

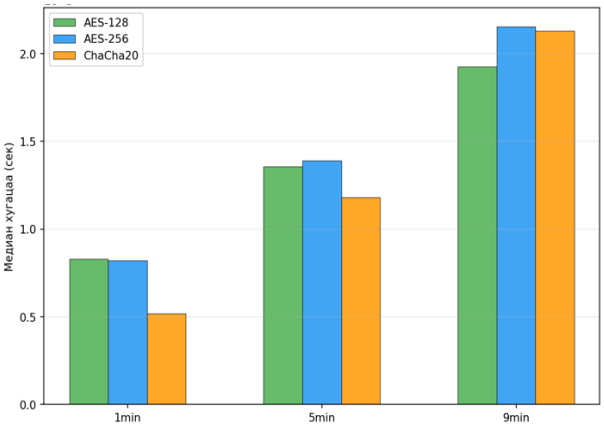


Зураг 10.

9 минутын файлын хэмжилтийн тархалтаар AES-128 болон ChaCha20-н box-ууд нягтарч хоорондоо ойртсон бол RSA-2048-н тархалт өмнөх шатны үзүүлэлттэй ижил түвшинд хадгалагдлаа. Дундаж хугацааны хувьд ChaCha20 анх удаа AES-128-аас удаан (0.0260 мс) гарсан бөгөөд RSA-2048 нь бусад алгоритмуудаасаа их хугацаа зарцуулсан хэвээр байна. Өгөгдлийн хэмжээ ихсэх тусам ChaCha20-н шугам AES-256-г давж, AES-128-тай ойртож байгаа нь урт өгөгдлийн дамжуулалтад ChaCha20-н хурдны давуу тал буурч байгааг илтгэж байна.



Зураг 11. Бүх алгоритмын лог масштаб Гурван өөр файлын хэмжээ дээрх медиан хоцрогдлыг log масштабаар харьцуулахад RSA-2048-ийн үзүүлэлт нь симметр алгоритмуудаас 10^4 дахин их буюу илүү өндөр хоцрогдолтой байна. Файлын хугацаа ихсэх тусам RSA-2048-ийн хугацаа 1.48 секундээс 0.30 секунд болж буурах хандлага ажиглагдсан бол симметр алгоритмууд 10-5 секундын түвшинд тогтвортой байна. Симметр алгоритмуудын медиан утга хоорондоо хэт ойрхон тул ерөнхий графикаас ялгаж харах боломжгүй бөгөөд тэдгээрийн нарийвчилсан харьцуулалтыг зураг 12-т тусгайлан үзүүлэв.



Зураг 12. Гурван симметр алгоритмын медиан хоцрогдлыг шугаман масштабаар харьцуулахад 1 болон 5 минутын файлууд дээр ChaCha20 хамгийн бага хоцрогдолтой буюу хамгийн хурдан байлаа. Харин 9 минутын файл дээр AES-128 ($1.925 * 10^{-5}$ сек) хамгийн хурдан болж, AES-256 болон ChaCha20-ийн үзүүлэлтүүд бараг тэнцэж байна. Энэхүү үр дүн нь өгөгдлийн хэмжээ ихсэх тусам алгоритмуудын гүйцэтгэлийн давуу тал өөрчлөгдөж, хоорондоо огтлолцож байгааг судалгааны үр дүн болгон харууллаа.

G. Хурдны харьцааны дүн шинжилгээ:

1 минутын файлд ChaCha20 нь AES-128-аас 1.60 дахин хурдан байсан боловч 9 минутад энэ харьцаа 0.90 дахин болж ChaCha20 AES-128-аас удаан болсон. Энэ огтлолцол нь бодит цагийн STT системд өгөгдлийн хэмжээ нь алгоритм сонголтод нөлөөлдгийг практикт харуулсан чухал олдвор юм.

RSA-2048-н медиан хоцрогдол нь 1 минутын файлд (1.481 сек) 5 болон 9 минутынхаас (0.343–0.303 сек) 4–5 дахин их байсан нь RSA-н түлхүүр үүсгэлтэд оролцдог анхны тоо хайх процессын санамсаргүй шинж чанараас үүдэлтэй хэлбэлзэлтэй холбоотой. Харин RSA-н хамгийн бага утга (0.048–0.175 сек) нь медианаас эрс бага байгаа нь тохиромжтой эхний тоо хурдан олдсон тохиолдлыг илтгэж байна.

RSA-2048 нь симметр алгоритмуудаас файл бүрт дараах хэмжээгээр удаан байсан нь тогтоогдов: 1 минутад AES-256-аас 180,666 дахин, 5 минутад 24,708 дахин, 9 минутад 14,080 дахин удаан байсан нь RSA-г STT өгөгдлийн шифрлэлтэд шууд ашиглах боломжгүй болохыг тодорхой харуулж байна. Практик хэрэглээнд RSA нь зөвхөн симметр түлхүүр солилцоонд ашиглах зориулалттай болохыг энэхүү судалгааны үр дүн нотолж байна.

Зураг 11 болон 12-д файл бүрийн нэгтгэсэн графикуудыг харуулав. Box plot (a) нь AES болон ChaCha20-н хоцрогдлын тархалт нь log масштабаар RSA-2048-нхаас эрс ялгаатай болохыг, bar chart (b) нь медиан хоцрогдол дээр суурилсан алгоритм хоорондын харьцуулалтыг, шугаман график (c) нь өгөгдлийн хэмжээ нэмэгдэхэд ChaCha20-н latency өсөлтийн хурд AES-нхаас их болохыг тус тус тодорхой дүрслэн харуулж байна.

V. ДҮГНЭЛТ

Энэхүү судалгаагаар бид Монгол хэлний ярианаас бичвэр буулгах системд төгсгөлөөс төгсгөл рүү нууцлах E2EE нэвтрүүлэхэд үүсэх техникийн хоцрогдол болон гүйцэтгэлийн жишиг үзүүлэлтийг тодорхойллоо.

Туршилтын үр дүнд үндэслэн дараах үндсэн дүгнэлтүүдийг гаргаж байна.

Өгөгдлийн хэмжээ бага байх үед ChaCha20 алгоритм нь AES-128-аас 1.6 дахин хурдан гүйцэтгэл үзүүлж байгаа бол, өгөгдөл томрох тусам AES-128 алгоритм тооцооллын хувьд илүү үр ашигтай болж байна. Энэ нь бодит цагийн STT системд өгөгдлийн урсгалын хурднаас хамаарч алгоритмыг сонгох шаардлагатайг харуулж байна.

RSA-2048 алгоритм нь симметр алгоритмуудаас дунджаар 14,000-аас 180,000 дахин удаан байгаа нь түүнийг STT-ийн их хэмжээний өгөгдлийг шууд шифрлэхэд ашиглах боломжгүйг туршилтаар баталлаа. RSA-ийг зөвхөн тэгш хэмт түлхүүр солилцох түвшинд хязгаарлах нь зохистой.

Туршилтаар хэмжигдсэн хамгийн их нууцлалын хоцрогдол нь ITU-T G.114 стандартын 150-400ms-ийн босгоос бага байна.

UTF-8 кодчилол бүхий Монгол бичвэрийн өгөгдөл нь алгоритмын ачаалалд нөлөөлж байгаа ч орчин үеийн симметр алгоритмууд үүнийг бодит цагийн горимд сааталгүй боловсруулах чадамжтай байна.

VI. Ном зүй

[1] Grand View Research, "Voice and Speech Recognition Market Size, Share & Trends Analysis Report By Technology, By End-use, By Region, and Segment Forecasts, 2024 - 2030," 2024. [Online]. Available: https://www.grandviewresearch.com/industry-analysis/voice-recognition-market.

[2] IBM Security, "2025 Cost of a Data Breach Report," IBM Corp., Armonk, NY, USA, 2025. [Online]. Available: https://www.ibm.com/reports/data-breach.

[3] ITU-T, "One-way transmission time," Recommendation G.114, International Telecommunication Union, May 2003. [Online]. Available: https://www.itu.int/rec/T-REC-G.114.

[4] H. Kheddar, M. Hemis, Y. Himeur, “Automatic speech recognition using advanced deep learning approaches: A survey,” Information Fusion, vol. 109, 2024, doi: 10.1016/j.inffus.2024.102422.

[5] A. Mehrish et al., “Automatic Speech Recognition: A survey of deep learning techniques and approaches,” ScienceDirect, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666307424000573

[6] A. Radford et al., “Robust speech recognition via large-scale weak supervision,” OpenAI, 2022. [Online]. Available: https://openai.com/index/whisper/

[7] R. Patel, S. Bhatt, A. Garg, “Text-to-Speech and Speech-to-Text Models: A Systematic Examination of Diverse Approaches,” IEEE Conference Publication, 2024.

[8] Y. Li, Y. Wang, Z. Gao, F. Bao, “Mongolian Speech Recognition Based on Semi-supervised Learning and Syllable Subword Modeling Units,” in Proc. NLPCC 2025, Lecture Notes in Comput. Sci., vol. 16105, Springer, 2026.

[9] K. Liang et al., “MnTTS2: An Open-Source Multi-Speaker Mongolian TTS Dataset,” in Proc. NCMMSC 2022, Commun. Comput. Inf. Sci., vol. 1765, Springer, 2023.

[10] Y. Zhang et al., “Frontier Research on Low-Resource Speech Recognition Technology,” Sensors, vol. 23, no. 22, p. 9096, 2023.

[11] AssemblyAI, “Speech-to-Text Privacy, Security, and Compliance,” 2024. [Online]. Available: https://www.assemblyai.com/blog/speech-to-text-security

[12] S. Shahid et al., “Real-Time Speech Privacy Protection,” in Proc. USENIX Security 2025, 2025.

[13] Fortanix, “Confidential Speech Recognition,” Fortanix Blog, 2023. [Online]. Available: https://www.fortanix.com/blog/confidential-speech-recognition

[14] A. Baevski et al., “wav2vec 2.0: A framework for self-supervised learning of speech representations,” in NeurIPS, 2020, vol. 33.

[15] Grand View Research, “Voice and Speech Recognition Market Size Report, 2030,” Grand View Research, 2023.

[16] AIML API, “The Ultimate 2026 Guide to Speech-to-Text (STT) APIs,” AI/ML API Blog, 2026. [Online]. Available: https://aimlapi.com/blog/introduction-to-speech-to-text-technology

[17] Picovoice, “Speech-to-Text Privacy, Security, and Compliance,” Picovoice Blog, 2026. [Online]. Available: https://picovoice.ai/blog/speech-to-text-privacy-security-compliance/

[18] A. B. Elminaam, H. M. A. Kader, and M. M. Hadhoud, "Performance Evaluation of Symmetric Encryption Algorithms," *Int. J. Comput. Sci. Netw. Secur.*, vol. 9, no. 3, pp. 249-264, Mar. 2009.

[19] W. Stallings, "Public-Key Cryptography and RSA," in *Cryptography and Network Security: Principles and Practice*, 7th ed. Upper Saddle River, NJ, USA: Pearson, 2017, ch. 9, pp. 277–290. https://www.uoitc.edu.iq/images/documents/informatics-institute/Competitive_exam/Cryptography_and_Network_Security.pdf

[20] W. Stallings, "Advanced Encryption Standard," in *Cryptography and Network Security: Principles and Practice*, 7th ed. Upper Saddle River, NJ, USA: Pearson, 2017, ch. 5, pp. 148–183. https://www.uoitc.edu.iq/images/documents/informatics-institute/Competitive_exam/Cryptography_and_Network_Security.pdf

[21] J.-P. Aumasson, "Stream Ciphers," in *Serious Cryptography: A Practical Introduction to Modern Encryption*. San Francisco, CA, USA: No Starch Press, 2017, ch. 5, pp. 71–88. https://www.kea.nu/files/textbooks/humblesec/seriouscrytography.pdf

[22] S. Gueron, "Intel Advanced Encryption Standard (AES) NewInstructions Set," Intel White Paper, 2010.

[23] D. J. Bernstein, "ChaCha, a variant of Salsa20," Citeseer, 2008.

[24] N. Nir and A. Langley, "ChaCha20 and Poly1305 for IETF Protocols," RFC 7539, May 2015

Монголын уртын дууны акустик онцлогт суурилсан хэв маягийн анализ

Норов Уранчимэг
Хөгжмийн урлагийн сургууль
Соёл Урлагийн Их сургууль
Улаанбаатар, Монгол
uranchimeg.norov@gmail.com

Баатархүү Амартүвшин
МТЭС
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
amartuvshin109@gmail.com

Адъяабат Энхжаргал*
МТЭС
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
adiyabat.e@num.edu.mn

Дагданпүрэв Сумъяаханд
МТЭС
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
sumiyakhand@num.edu.mn

Цэрэн Чулуунцэцэг
Театрын Урлагийн сургууль
Соёл Урлагийн Их сургууль
Улаанбаатар, Монгол
chuka0588@gmail.com

Дорж Бямбажав
МТЭС
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
byambajav@num.edu.mn

***Хураангуй*—Монгол уртын дуу нь нүүдэлчин монголчуудын соёл иргэншлээс үүсэлтэй, олон мянган жилийн түүхтэй уламжлалт дуулах урлагийн өвөрмөц хэлбэр бөгөөд хүний дуу хоолойн уян хатан байдал, акустик боломжийг дээд түвшинд илэрхийлдгээрээ онцлог юм. Энэхүү судалгаанд уртын дууны дуулах хэв маягийг акустик параметрүүдийн түвшинд тоон байдлаар тодорхойлох зорилго тавьсан. Судалгаанд дууны үндсэн давтамж (F_0), давтамжийн муруй, чичирхийллийн шинж (vibrato), болон Log-Mel спектрограм зэрэг параметрүүдийг ашиглан шинжилгээ хийсэн. Туршилтад Ж. Дорждавгын дуулсан "Өвгөн шувуу", "Эртний сайхан" дуунуудыг ашиглаж, аудио өгөгдлийг 30 секундын сегментүүдэд хуваан шинжилсэн. Статистик шинжилгээгээр F_0, vibrato rate, vibrato extent параметрүүдийн дундаж, стандарт хазайлт, 95%-ийн итгэлцлийн интервал, мөн хоёр сорьцын t-шалгалтыг ашиглан ялгааг үнэлсэн. Үр дүнгээс харахад уртын дуу нь өргөн давтамжийн муж, тогтвортой боловч динамик хэлбэлзэл, удаан чичирхийлэлтэйгээр тодорхойлогдож байгааг тогтоосон. Энэхүү судалгаа нь уртын дууны акустик "гарын үсэг"-ийг тодорхойлох боломжийг харуулж байгаа бөгөөд цаашид дуу таних, ангилах, машин сургалтын (SVM, Random Forest, CNN) загваруудад ашиглах суурь болно. Гэвч өгөгдлийн хэмжээ хязгаарлагдмал тул цаашид олон дуучин, олон дуу дээр туршилт хийх шаардлагатай.**

***Түлхүүр үг*—Монгол уртын дуу, акустик шинжилгээ, давтамжийн муруй (pitch contour), F_0, чичирхийллийн шинж (vibrato), Log-Mel спектрограм, дууны дохио боловсруулалт**

I. Удиртгал

Монгол уртын дуу нь нүүдэлчин монголчуудын ахуй, соёл, сэтгэлгээтэй салшгүй холбоотой, олон мянган жилийн түүхэн уламжлалтай өвөрмөц дуулах урлаг юм. Энэхүү урлаг нь 2005 онд ЮНЕСКО-гийн биет бус соёлын өвийн жагсаалтад бүртгэгдсэн бөгөөд хүн төрөлхтний дуу авианы урлагийн ховор, дахин давтагдашгүй хэлбэрүүдийн нэгэнд тооцогддог [1].

Уртын дууны онцлог нь нэг үгийн нэг эгшгийг маш удаан хугацаанд тасралтгүй, уянгалаг сунган дуулах явдал бөгөөд энэ нь хүний дуу хоолойн хяналт, амьсгалын зохицуулалт, акустик боломжийн дээд түвшний илэрхийлэл болдог. Дуулах явцад дуу хоолой нь тогтмол хэмнэл бүхий хэлбэлзэл буюу *чичирхийлэл* (vibrato) үүсгэж, мөн ам, хамар, залгиур зэрэг резонансын хөндийг оновчтой ашигласнаар үндсэн авианаас гадна нэмэлт давхар авиа (харь авиа, overtone) үүсгэдэг [2],[3]. Ийм олон давхар бүтэцтэй акустик шинж чанар нь уртын дууг судалгааны хувьд онцгой сонирхолтой объект болгодог.

Сүүлийн жилүүдэд яриа болон дууны дохионы боловсруулалтын чиглэлд хиймэл оюун ухаан, машин сургалтын аргууд эрчимтэй хөгжиж байгаа хэдий ч Монгол уртын дууны акустик онцлогийг тоон параметрээр системтэйгээр тодорхойлсон судалгаа харьцангуй хомс хэвээр байна. Иймд уртын дууны үндсэн акустик параметрүүд болох дууны үндсэн давтамж (F_0), давтамжийн муруй, спектрийн бүтэц, формант, мөн чичирхийллийн шинж (чичирхийллийн давтамж ба далайц) болон давхар авианы бүтэц зэрэг үзүүлэлтүүдийг нарийвчлан судлах шаардлага тулгарч байна [4],[5].

Уртын дууны акустик шинжийг тоон параметрээр тодорхойлох нь хэд хэдэн чухал ач холбогдолтой. Нэгдүгээрт, уламжлалт урлагийг дижитал хэлбэрт хадгалж, хойч үед өвлүүлэн үлдээх, соёлын өвийг хамгаалах үндэс суурийг бүрдүүлнэ [1]. Хоёрдугаарт, уртын дууны дуулах хэв маяг, арга барилыг автоматаар таних, ангилах, синтез хийх системүүдийг хөгжүүлэхэд шаардлагатай өгөгдөл, параметрийн суурийг бий болгоно [4]. Гуравдугаарт, дуучны дуулах онцлог, хэв маягийг ялган таних, үнэлэх, улмаар гүйцэтгэлийн ялгааг тодорхойлох боломжийг нэмэгдүүлнэ. Түүнчлэн энэ төрлийн судалгаа нь дуу авианы боловсруулалт, акустик шинжилгээ, соёлын өвийн дижитал хадгалал-

тын огтлолцолд шинэ судалгааны чиглэл хөгжихөд хувь нэмэр оруулах ач холбогдолтой юм.

Энэхүү судалгаанд нийт 20–30 орчим уртын дууны бичлэгийг (олон дуучин, олон бүтээл) ашиглан акустик шинжилгээ хийсэн бөгөөд энэ нь өмнөх хязгаарлагдмал өгөгдөлтэй судалгаануудтай харьцуулахад илүү өргөн хүрээний дүгнэлт гаргах боломжийг бүрдүүлж байна. Гэсэн хэдий ч өгөгдлийн төрөл, чанар, дуучдын ялгаа зэргээс шалтгаалан үр дүнг тайлбарлахдаа болгоомжтой хандах шаардлагатай бөгөөд цаашид илүү их хэмжээний өгөгдөл ашиглан судалгааг өргөжүүлэх шаардлагатай байна.

Иймээс энэхүү судалгааны зорилго нь Монгол уртын дууны акустик шинж чанарыг тоон параметрүүдээр тодорхойлох, тэдгээрийн хоорондын хамаарлыг судлах, цаашид автомат таних болон синтезийн системд ашиглах боломжийг бүрдүүлэхэд оршино.

Энэхүү өгүүлэлд Монгол уртын дууны акустик онцлогт суурилсан дууны хэв маягийн шинжилгээний үр дүнг танилцуулна. Хоёрдугаар хэсэгт дууны үндсэн давтамж, давтамжийн муруй, чичирхийллийн шинж чанар, Log-Mel спектрограмын онолын үндсийг авч үзнэ. Гуравдугаар хэсэгт статистик шинжилгээний арга зүй, дөрөвдүгээр хэсэгт туршилтын үр дүн, тавдугаар хэсэгт акустик шинжүүдийн харьцуулалтыг танилцуулж, эцэст нь дүгнэлтийг өгүүлнэ.

II. Акустик шинжүүдийн онолын үндэс

A. Дууны үндсэн давтамж (F_0) ба давтамжийн муруй (pitch contour)

Дууны үндсэн давтамж (F_0) нь хүний дууны хөвчний чичирхийллээс үүсэх хамгийн бага давтамж бөгөөд дуу авианы өндөр намыг тодорхойлох үндсэн параметр юм. Давтамжийн ойлголт (pitch) нь F_0-той шууд хамааралтай бөгөөд сонсогчийн хүлээн авах дууны өндөр намыг илэрхийлдэг субъектив хэмжигдэхүүн болно [6].

Дискрет дохионы хувьд үндсэн давтамжийг автокорреляцийн функц (autocorrelation)-ийн аргаар дараах байдлаар тодорхойлж болно:

$$R(\tau) = \sum_{n=0}^{N-1} x(n)x(n+\tau) \tag{1}$$

Энд $x(n)$ нь аудио дохио, τ нь саатлын индекс, $R(\tau)$ нь автокорреляцийн утга юм. Үндсэн давтамж нь $R(\tau)$-ийн хамгийн их утгад харгалзах τ-ийн урвуугаар тодорхойлогдоно:

$$F_0 = \frac{1}{\tau_{max}} \tag{2}$$

Давтамжийн муруй (pitch contour) нь F_0-ийн цаг хугацааны дагуух өөрчлөлтийг илэрхийлэх бөгөөд:

$$F_0(t) = \{F_0(t_1), F_0(t_2), \ldots, F_0(t_T)\} \tag{3}$$

хэлбэртэй байна. Энэ нь дууны аялгуу, интонац, дуулах хэв маягийг тодорхойлоход чухал үүрэгтэй [7],[8].

B. Чичирхийллийн шинж (Vibrato)

Чичирхийлэл (vibrato) нь үндсэн давтамжийн бага давтамжтай хэлбэлзэл бөгөөд дараах байдлаар илэрхийлэгдэнэ:

$$F_0(t) = F_c + A_v \sin(2\pi f_v t) \tag{4}$$

Энд:

- F_c – төв давтамж (carrier frequency)
- A_v – чичирхийллийн далайц (vibrato extent)
- f_v – чичирхийллийн давтамж (vibrato rate)

Чичирхийллийн давтамжийг:

$$f_v = \frac{N_{cycles}}{T} \tag{5}$$

Чичирхийллийн далайцыг:

$$A_v = \frac{F_{max} - F_{min}}{2} \tag{6}$$

мөн давтамжийн тогтвортой байдлыг (pitch stability) стандарт хазайлтаар:

$$\sigma_{F_0} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(F_0(i) - \mu)^2} \tag{7}$$

гэж тодорхойлно. Эдгээр үзүүлэлтүүд нь дууны өнгө, чанар, дуучны гүйцэтгэлийн онцлогийг илэрхийлэхэд чухал ач холбогдолтой [9].

C. Log-Mel спектрограм

Дууны дохиог богино хугацааны Фурьегийн хувиргал (Short-Time Fourier Transform, STFT)-аар задлан спектрийн шинжилгээ хийдэг:

$$X(n,k) = \sum_{m=0}^{M-1} x(m)w(n-m)e^{-j2\pi km/N} \tag{8}$$

Энд $w(n)$ нь цонхны функц (window function), $X(n,k)$ нь спектрийн утга юм.

Энергийн спектр:

$$S(n,k) = |X(n,k)|^2 \tag{9}$$

Хүний сонсголын логарифм шинжийг харгалзан Mel хэмжээсийг (Mel scale) дараах байдлаар тодорхойлно:

$$m = 2595 \log_{10}\left(1 + \frac{f}{700}\right) \tag{10}$$

Mel фильтр банк (Mel filter bank)-ийг ашиглан:

$$S_{mel}(i) = \sum_{k} S(n,k)H_i(k) \tag{11}$$

улмаар логарифм хувиргал хийснээр:

$$S_{logmel}(i) = \log(S_{mel}(i)) \tag{12}$$

Log-Mel спектрограм нь дууны спектрийн тархалтыг хүний сонсголд ойр хэлбэрээр илэрхийлдэг бөгөөд машин сургалтын (machine learning) системүүдэд өргөн хэрэглэгддэг [10],[11].

D. Параметрүүдийн уялдаа

Дууны акустик шинжүүдийг дараах үндсэн бүрэлдэхүүнүүдэд хувааж үзэж болно:

- F_0 – дууны үндсэн давтамж
- Чичирхийлэл (vibrato) – F_0-ийн микродинамик хэлбэлзэл
- Log-Mel спектрограм – спектрийн бүтцийн иж бүрэн дүрслэл

Эдгээр параметрүүдийг хамтад нь авч үзэх нь дууны хэв маягийг илүү нарийвчлалтай тодорхойлох боломжийг олгодог.

III. Статистик шинжилгээ

Энэхүү судалгаанд дууны акустик параметрүүд болох үндсэн давтамж (F_0), чичирхийллийн давтамж (vibrato rate) болон чичирхийллийн далайц (vibrato extent)-д статистик шинжилгээ хийж, үр дүнг тоон байдлаар үнэлсэн.

A. Тодорхойлох статистик (Descriptive Statistics)

Акустик параметрүүдийн тархалтыг тодорхойлохын тулд дундаж утга болон стандарт хазайлтыг дараах байдлаар тооцоолсон [20], [21]:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{13}$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \tag{14}$$

B. Итгэлцлийн интервал (Confidence Interval)

Дундаж утгын найдвартай байдлыг үнэлэх зорилгоор 95%-ийн итгэлцлийн интервалыг Student-ийн t тархалтаар тооцоолсон [21], [22]:

$$\mu \pm t_{\alpha/2,n-1} \cdot \frac{s}{\sqrt{n}} \tag{15}$$

C. Таамаглалын шалгалт (Hypothesis Testing)

Хоёр бүлгийн ялгааг шалгах зорилгоор хоёр сорьцын t-шалгалтыг ашигласан [22]:

$$t = \frac{\mu_1 - \mu_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{16}$$

Тэг таамаглал (H_0): ялгаа байхгүй Эсрэг таамаглал (H_1): ялгаа байна

D. Хамаарлын шинжилгээ (Correlation Analysis)

Параметрүүдийн хамаарлыг Пирсоны корреляцийн коэффициент (Pearson correlation coefficient)-оор үнэлсэн [23]:

$$r = \frac{\sum(x_i - \mu_x)(y_i - \mu_y)}{\sqrt{\sum(x_i - \mu_x)^2 \sum(y_i - \mu_y)^2}} \tag{17}$$

E. Нөлөөллийн хэмжээ (Effect Size)

Хоёр бүлгийн ялгааны хэмжээг Cohen-ийн d коэффициентоор тодорхойлсон [24]:

$$d = \frac{\mu_1 - \mu_2}{s_{\text{pooled}}} \tag{18}$$

$$s_{\text{pooled}} = \sqrt{\frac{s_1^2 + s_2^2}{2}} \tag{19}$$

- $d \approx 0.2$ — бага нөлөө
- $d \approx 0.5$ — дунд нөлөө
- $d \geq 0.8$ — их нөлөө

IV. Монгол уртын дууны акустик шинжилгээний хэрэглээ

Монгол уртын дууны акустик хэв маягийг тоон параметрүүдийн түвшинд тодорхойлох нь зөвхөн онолын ач холбогдлоор хязгаарлагдахгүй, практик хэрэглээний өргөн боломжийг нээж өгдөг. Тухайлбал, дууны үндсэн давтамж (F_0), давтамжийн муруй (pitch contour), чичирхийллийн шинж (vibrato), мөн Log-Mel спектрограм зэрэг акустик үзүүлэлтүүдийг ашиглан уртын дууны хэв маягийг автоматаар таних, ангилах, хадгалах, сэргээх зэрэг олон төрлийн систем хөгжүүлэх боломжтой.

Нэгдүгээрт, энэхүү шинжилгээ нь соёлын өвийг дижитал хэлбэрт хадгалах, архивлах системийн суурь болж чадна. Уртын дууны акустик параметрүүдийг тодорхойлсноор тухайн дууны “дижитал гарын үсэг” (acoustic signature)-ийг үүсгэж, уртын дууны өгөгдлийн сан бүрдүүлэх боломжтой. Энэ нь соёлын өвийг хамгаалах, цаашдын судалгаанд ашиглахад чухал ач холбогдолтой.

Хоёрдугаарт, дуулах хэв маяг таних (style recognition) болон дуучин таних (singer identification) системийг хөгжүүлэх боломж бүрдэнэ. Уртын дууны хувьд дуучин бүрийн давтамжийн муруй, чичирхийллийн хэв шинж, спектрийн бүтэц харилцан адилгүй байдаг тул эдгээрийг машин сургалтын (machine learning) аргаар боловсруулж, ангилах, ялган таних боломжтой.

Гуравдугаарт, дууны нийлэгжил (voice synthesis) болон хэв маяг дамжуулах (style transfer) системд ашиглах боломжтой. Акустик параметрүүдэд суурилан уртын дууны хэв маягийг загварчилснаар тухайн хэв маягийг дуурайлган гаргах, шинэ дуу үүсгэх, эсвэл өөр дуу хоолойд уртын дууны шинжийг шингээх боломж бүрдэнэ.

Мөн энэхүү судалгаанд тулгуурлан дуу таних систем (audio recognition) хөгжүүлэх боломжтой. Энэ нь Shazam зэрэг хөгжмийн таних системтэй төстэй зарчимд суурилдаг бөгөөд аудио дохионоос онцлог шинжүүдийг гарган авч (feature extraction), өгөгдлийн санд хадгалагдсан “акустик хурууны хээ” (audio fingerprint)-тэй харьцуулан таних үйлдэл гүйцэтгэнэ. Уртын дууны

хувьд урт хугацааны давтамжийн муруй болон чичиргийллийн бүтэц зэрэг онцлогийг тусгайлан харгалзан үзэх шаардлагатай.

Иймэрхүү системийг дараах үндсэн үе шатуудаар хэрэгжүүлж болно:

- Аудио дохионоос акустик шинжүүдийг гарган авах (feature extraction)
- Онцлог шинжүүдийг кодлох, шахах (feature encoding)
- Өгөгдлийн санд индексжүүлэх (database indexing)
- Оролтын дохиог өгөгдөлтэй харьцуулах (matching)

Монгол уртын дуу нь урт хугацааны үргэлжилсэн аялгуу, нарийн чичиргийллийн бүтэц, баялаг спектрийн шинж чанартай тул уламжлалт дуу таних системүүдээс ялгаатай сорилтуудыг үүсгэдэг. Иймээс цаашдын судалгаанд эдгээр онцлогийг харгалзан илүү оновчтой шинж төлөөлөл (feature representation) болон найдвартай харьцуулах алгоритмууд (robust matching algorithms) хөгжүүлэх шаардлагатай.

Иймд Монгол уртын дууны акустик шинжилгээ нь дуу таних, дуучин ялгах, дуу нийлэгжүүлэх, соёлын өвийг хамгаалах зэрэг олон талт хэрэглээний суурь болох боломжтой юм.

V. Холбогдох судалгаа

Яриа болон дууны дохионы акустик боловсруулалтын чиглэлд сүүлийн жилүүдэд олон чухал судалгаа хийгдсэн. Дууны үндсэн давтамж (F_0)-ийг нарийвчлалтай тодорхойлох нь энэ чиглэлийн суурь асуудлын нэг бөгөөд McLeod болон Wyvill [12] нар YIN алгоритмыг санал болгосноор давтамж илрүүлэх (pitch detection)-ийн нарийвчлал сайжирсан. Үүний дараа Mauch болон Dixon [13] нар магадлалд суурилсан pYIN алгоритмыг боловсруулж, дуутай болон дуугүй хэсгийг илүү найдвартай ялгах боломжийг бүрдүүлсэн.

Сүүлийн үед гүн сургалтын (deep learning) аргууд яриа таних болон дуу боловсруулах системүүдэд өргөн хэрэглэгдэж байна. Тухайлбал, Radford нарын боловсруулсан Whisper загвар [14] нь олон хэл дээр яриа таних өндөр чадвартай бол Pratap нарын MMS загвар [15] нь олон хэлний яриа таних, синтезийг хамарсан өргөн хүрээний шийдэл болсон. Эдгээр судалгаанууд нь дууны дохионоос семантик түвшний мэдээлэл гарган авах боломжийг харуулж байна.

Практик хэрэглээнд Python хэл дээрх librosa сан [16] өргөн ашиглагддаг бөгөөд богино хугацааны Фурьегийн хувиргал (STFT), Mel спектрограм, MFCC, pYIN зэрэг олон төрлийн акустик шинжүүдийг гарган авах боломжийг олгодог.

Нөгөө талаас уламжлалт хөгжмийн акустик шинжилгээний судалгаанууд нь тухайн ардын хөгжмийн онцлогийг тодорхойлоход чиглэсэн байдаг. Жишээлбэл, төв азийн уламжлалт дуулах хэлбэрүүдийн харь авианы (overtone) бүтэц, спектрийн шинжийг судалсан ажлууд бий [17]. Мөн этномузикологийн чиглэлд Монгол уртын дууны аялгуу, бүтэц, дуулах арга барилыг судалсан боловч ихэвчлэн чанарын шинжилгээнд тулгуурласан байдаг [18].

Монгол уртын дууны хувьд давтамжийн муруй, чичиргийллийн шинж, спектрийн бүтэц зэрэг акустик параметрүүдийг системтэйгээр тоон байдлаар судалсан ажил харьцангуй хомс хэвээр байна. Иймд орчин үеийн дохио боловсруулалт болон машин сургалтын аргуудыг ашиглан уртын дууны акустик хэв маягийг тоон параметрээр тодорхойлох нь судалгааны шинэ, ач холбогдолтой чиглэл болж байна.

VI. Монгол уртын дууны акустик хэв маягийн анализийн системийн бүтэц

Энэхүү судалгаанд Монгол уртын дууны акустик хэв маягийг тодорхойлох зорилгоор дууны дохионоос онцлог шинжүүдийг ялган авч, тэдгээрийг тоон параметр болгон илэрхийлэх системийг санал болгож байна. Тус систем нь орчин үеийн дууны дохио боловсруулалт болон аудио шинжилгээний онолуудад тулгуурласан [1],[5].

A. Системийн ерөнхий загвар

Системийн оролт нь аудио дохио $x(n)$ бөгөөд дараах хувиргалтаар акустик параметрүүдийг гарган авна:

$$\mathbf{F} = \Phi(x(n)) \tag{20}$$

Энд $\Phi(\cdot)$ нь онцлог шинж ялгах функц (feature extraction function) бөгөөд гаралт нь:

$$\mathbf{F} = \{F_0(t), V(t), S_{logmel}(t,f)\} \tag{21}$$

болно. Иймэрхүү төрлийн шинж төлөөлөл (feature representation) нь дуу болон яриа таних системүүдэд өргөн хэрэглэгддэг [6],[10].

B. Дууны дохио бэлтгэх (Pre-processing)

Аудио дохиог шинжилгээнд бэлтгэхийн тулд нормчлох болон цонхлох (windowing) үйлдлийг гүйцэтгэнэ:

$$x_w(n) = x(n) \cdot w(n) \tag{22}$$

Энд $w(n)$ нь Hamming цонхны функц (Hamming window):

$$w(n) = 0.54 - 0.46\cos\left(\frac{2\pi n}{N-1}\right) \tag{23}$$

Мөн дохиог богино хугацааны хэсгүүдэд (frame) хуваана:

$$x_k(n) = x(n + kH) \tag{24}$$

Эдгээр алхмууд нь хугацаа–давтамжийн шинжилгээ (time-frequency analysis)-ний суурь болдог [5].

C. Онцлог шинж ялгах (Feature Extraction)

1) Дууны үндсэн давтамж (F_0): Үндсэн давтамжийг автокорреляцийн функц (autocorrelation)-аар:

$$F_0(t) = \arg\max_{\tau} R(\tau) \tag{25}$$

эсвэл илүү нарийвчлалтайгаар pYIN алгоритмаар тодорхойлно [12],[13].

2) Чичирхийллийн шинж (Vibrato): Чичирхийллийг давтамжийн муруйгаас бага давтамжийн бүрэлдэхүүнийг ялган тодорхойлно:

$$V(t) = F_0(t) - \mu_{F_0}(t) \tag{26}$$

Чичирхийллийн давтамж (vibrato rate):

$$f_v = \frac{N_{cycles}}{T} \tag{27}$$

Чичирхийллийн далайц (vibrato extent):

$$A_v = \frac{\max(F_0(t)) - \min(F_0(t))}{2} \tag{28}$$

Эдгээр үзүүлэлтүүд нь дууны микродинамик хэлбэлзлийг тодорхойлоход чухал ач холбогдолтой [4].

3) Log-Mel спектрограм: Богино хугацааны Фурьегийн хувиргал (Short-Time Fourier Transform, STFT)-ыг ашиглан:

$$X(t,f) = \sum x(n) w(n-t) e^{-j2\pi f n} \tag{29}$$

Mel фильтр банк (Mel filter bank)-ийг ашиглан:

$$S_{mel}(i) = \sum_{f} |X(t,f)|^2 H_i(f) \tag{30}$$

улмаар логарифм хувиргал хийж:

$$S_{logmel}(i) = \log(S_{mel}(i)) \tag{31}$$

Log-Mel спектрограм нь дууны спектрийн бүтцийг хүний сонсголд ойр хэлбэрээр илэрхийлдэг бөгөөд машин сургалтын (machine learning) системүүдэд өргөн хэрэглэгддэг [10],[11].

D. Хэв маяг тодорхойлох (Pattern Analysis)

Гарган авсан шинжүүдийг дараах вектор болгон нэгтгэнэ:

$$\mathbf{z} = [\mu_{F_0}, \sigma_{F_0}, f_v, A_v] \tag{32}$$

Энэхүү вектор нь тухайн сегментийн акустик “гарын үсэг” (acoustic signature)-ийг илэрхийлнэ.

E. Дуу таних (Fingerprinting систем)

Санал болгож буй систем нь аудио хуруугны хээ (audio fingerprinting)-д суурилсан таних зарчмыг ашиглана. Оролтын дохионы шинжүүдийг өгөгдлийн санд хадгалагдсан векторуудтай харьцуулна:

$$D(\mathbf{z}_1, \mathbf{z}_2) = \|\mathbf{z}_1 - \mathbf{z}_2\| \tag{33}$$

Энэ зарчим нь орчин үеийн дуу таних системүүдэд өргөн хэрэглэгддэг [19].

F. Системийн онолын үндэслэл

Энэхүү систем нь дараах үндсэн зарчмуудад тулгуурлана:

- Дууны дохио хугацаа–давтамжийн мужид бүрэн илэрхийлэгдэнэ
- Үндсэн давтамж (F_0) нь дууны гол шинжийг илэрхийлнэ
- Чичирхийлэл (vibrato) нь микродинамик өөрчлөлтийг тодорхойлно
- Log-Mel спектрограм нь спектрийн бүтцийг хадгална

G. Системийн давуу тал

- Уртын дууны онцлогт тохирсон шинж ялгалт (feature extraction)
- Чичирхийллийг нарийвчлан тооцсон
- Аудио хуруугны хээ (audio fingerprinting)-д суурилсан таних боломж
- Машин сургалтын системд өргөтгөх боломж

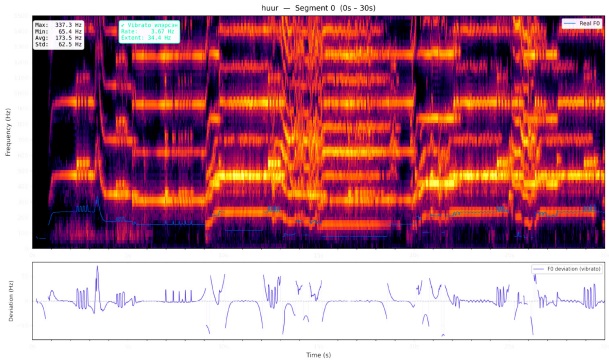
VII. Туршилтын үр дүн

Энэхүү судалгаанд нийт 25–30 орчим Монгол уртын дууны бичлэгийг ашиглан акустик шинжилгээ хийсэн. Туршилтад олон дуучин, олон бүтээл хамруулсан бөгөөд уртын дууны хэв маягийн ерөнхий акустик онцлогийг тодорхойлох зорилготой байв. Харин энэхүү өгүүлэлд үр дүнг илүү ойлгомжтой харуулах үүднээс Ж. Дорждагвагийн дуулсан “Эртний сайхан”, “Өвгөн шувуу” дуунуудыг төлөөлөл болгон сонгож, нарийвчилсан шинжилгээний үр дүнг үзүүлэв.

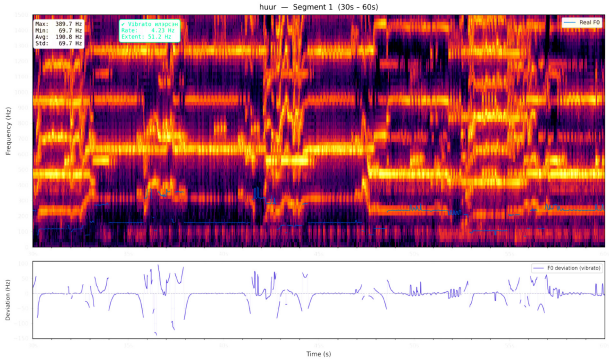
Аудио бичлэгийг 16 kHz дээж авах давтамжтайгаар боловсруулж, дууны дохиог 30 секундын урттай сегментүүдэд хуваасан. Нийт 5 сегмент (0–30s, 30–60s, 60–90s, 90–120s, 120–139s) дээр туршилтыг гүйцэтгэсэн.

Дууны дохионоос дараах акустик параметрүүдийг гарган авсан:

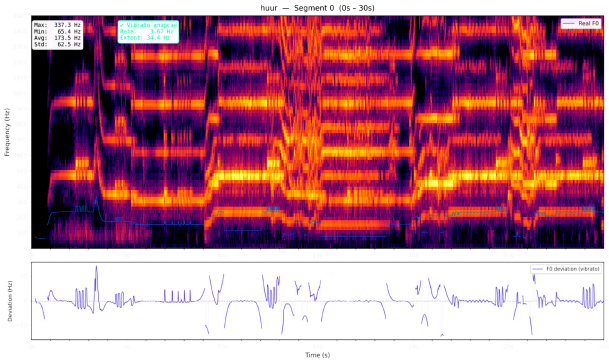
- Үндсэн давтамж (F_0) — pYIN алгоритм (pYIN algorithm)
- Давтамжийн муруй (pitch contour) болон давтамжийн тогтвортой байдал (pitch stability)
- Чичирхийллийн давтамж (vibrato rate) болон далайц (vibrato extent)
- Log-Mel спектрограм



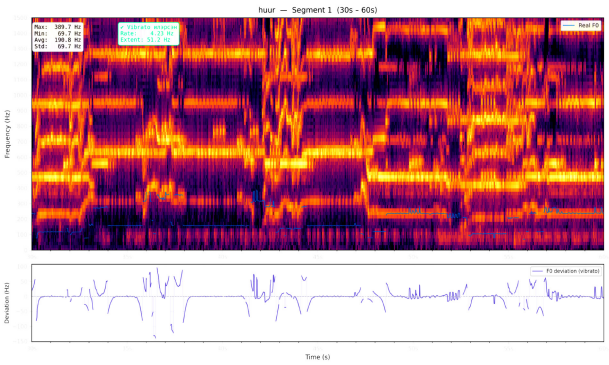
Зураг 1. Log-Mel спектрограм болон дууны үндсэн давтамжийн муруй (F_0) — “Эртний сайхан” уртын дуу (1-р сегмент)



Зураг 2. Log-Mel спектрограм болон F_0 муруй — “Эртний сайхан” уртын дуу (2-р сегмент)



Зураг 3. Log-Mel спектрограм болон F_0 муруй — “Өвгөн шувуу” уртын дуу (1-р сегмент)



Зураг 4. Log-Mel спектрограм болон F_0 муруй — “Өвгөн шувуу” уртын дуу (2-р сегмент)

Спектрийн шинжилгээнд богино хугацааны Фурьегийн хувиргал (Short-Time Fourier Transform, STFT)-ыг ашиглаж, Mel фильтр банк (Mel filter bank) болон логарифм хувиргал хийснээр Log-Mel спектрограм үүсгэсэн. Туршилтыг Python орчинд librosa санг ашиглан хэрэгжүүлсэн.

A. Спектрограм болон давтамжийн муруйн шинжилгээ

Зураг 1–4-д сегмент бүрийн спектрограм болон дууны үндсэн давтамжийн муруйг харуулсан. Спектрограм дээрх хэвтээ зурвасууд нь гармоник бүрэлдэхүүнүүд (харь авиа, overtone)-ийг илэрхийлж байгаа бөгөөд уртын дууны баялаг спектрийн бүтэц, олон давхар авианы онцлогийг харуулж байна.

Давтамжийн муруйгаас харахад F_0 нь ойролцоогоор 150–250 Hz мужид хадгалагдаж, аажим өөрчлөгдөж байна. Мөн чичирхийллийн хазайлт (vibrato deviation) нь тогтвортой, хэмнэлтэй хэлбэлзэлтэй байгаа нь уртын дууны чичирхийллийн онцлогийг илтгэнэ.

B. Давтамж болон чичирхийллийн параметрийн нэгдсэн шинжилгээ

Давтамжийн муж (F_0) 65–390 Hz хооронд хэлбэлзэж байгаа нь уртын дууны өргөн хүрээг илэрхийлнэ. Чичирхийллийн давтамж 3.3–4.2 Hz, далайц 30–50 Hz байгаа нь удаан, уянгалаг хэлбэлзлийг харуулж байна.

Ерөнхийдөө:

- өргөн давтамжийн муж
- тогтвортой боловч динамик F_0
- удаан, тасралтгүй чичирхийлэл

нь уртын дууны үндсэн шинжүүд юм.

C. Хоёр уртын дууны харьцуулалт

“Өвгөн шувуу” дуу нь илүү өргөн давтамжийн мужтай ($F_0 \approx 408$ Hz) бол “Эртний сайхан” нь харьцангуй бага мужтай ($F_0 \approx 341$ Hz) байна.

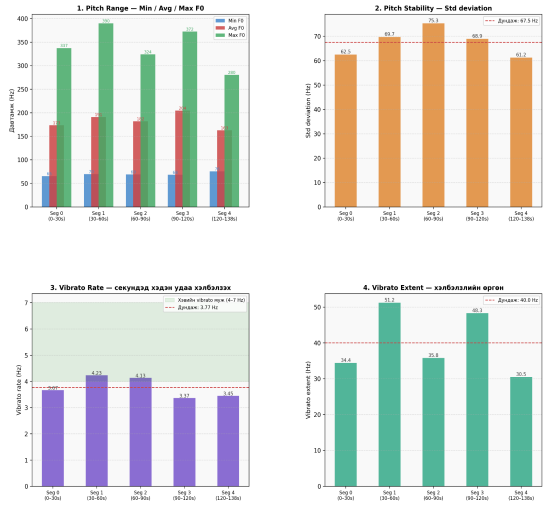
Чичирхийллийн давтамж хоёр дууны хувьд ойролцоо (3.6–3.8 Hz) байгаа боловч “Эртний сайхан” дуу илүү өргөн далайцтай байна.

Иймд:

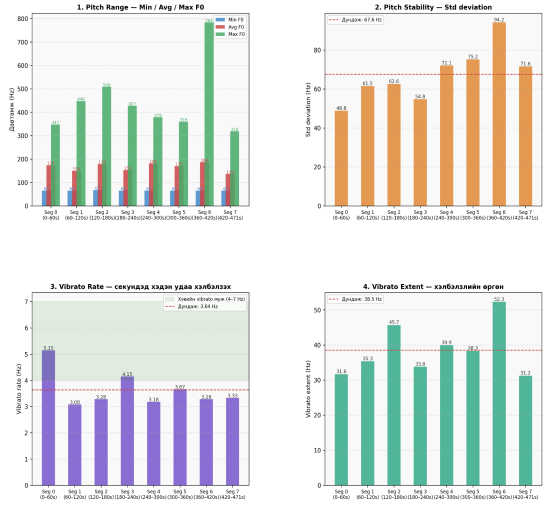
- “Өвгөн шувуу” — динамик өөрчлөлт их
- “Эртний сайхан” — илүү тогтвортой хэв маяг

гэсэн ялгаа ажиглагдав.

Эдгээр ялгаа нь уртын дууны хэв маяг, дуучны гүйцэтгэлийн онцлогийг илэрхийлж байна.



Зураг 5. Акустик параметрүүдийн харьцуулалт — “Эртний сайхан” уртын дуу



Зураг 6. Акустик параметрүүдийн харьцуулалт — “Өвгөн шувуу” уртын дуу

D. Акустик параметрүүдийн тоон шинжилгээ

Туршилтаар гарган авсан акустик параметрүүдийг тоон байдлаар үнэлж, уртын дууны онцлогийг тодорхойлов.

1) Дууны үндсэн давтамж (F_0): Шинжилгээний үр дүнгээс харахад дууны үндсэн давтамж (F_0) нь өргөн мужид хэлбэлзэж байна. Жишээлбэл, “Эртний сайхан” дууны хувьд:

- Хамгийн бага утга: ≈ 65 Hz
- Дундаж утга: ≈ 173 Hz
- Хамгийн их утга: ≈ 337 Hz
- Стандарт хазайлт: ≈ 62.5 Hz

Харин “Өвгөн шувуу” дууны хувьд F_0-ийн дээд утга илүү өндөр (ойролцоогоор 400 Hz-ээс дээш) байгаа нь илүү өргөн давтамжийн мужтай болохыг харуулж байна.

Энэ нь уртын дуу нь өргөн давтамжийн хүрээнд уян хатан өөрчлөгддөг болохыг илэрхийлнэ.

2) Чичирхийллийн шинж (Vibrato): Чичирхийллийн шинжилгээгээр дараах үр дүн гарсан:

- Чичирхийллийн давтамж (vibrato rate): ≈ 3.3–4.2 Hz
- Чичирхийллийн далайц (vibrato extent): ≈ 30–50 Hz

Жишээлбэл:

- “Эртний сайхан” дуу: $f_v \approx 3.67$ Hz, $A_v \approx 34.4$ Hz
- “Өвгөн шувуу” дуу: $f_v \approx 3.63$ Hz, $A_v \approx 38.5$ Hz

Эдгээр утгууд нь уртын дууны чичирхийлэл нь удаан, тогтвортой, харьцангуй өргөн далайцтай байдгийг харуулж байна.

3) Давтамжийн тогтвортой байдал (Pitch Stability): Давтамжийн тогтвортой байдлыг стандарт хазайлтаар үнэлэхэд:

- “Эртний сайхан”: $\sigma_{F_0} \approx 67.5$ Hz
- “Өвгөн шувуу”: $\sigma_{F_0} \approx 65.7$ Hz

Эдгээр утгууд нь F_0 нь тогтвортой боловч динамик хэлбэлзэлтэй байгааг илэрхийлж байна. Энэ нь уртын дууны онцлог болох уянгалаг, суналттай аялгууг илэрхийлнэ.

4) Сегмент хоорондын өөрчлөлт: Сегмент тус бүр дээрх дундаж F_0-ийн өөрчлөлтөөс харахад дууны явцад давтамжийн динамик шилжилт ажиглагдсан. Зарим сегментэд F_0 огцом өсөх (200 Hz $\rightarrow$ 300 Hz), буурах (180 Hz $\rightarrow$ 140 Hz) зэрэг өөрчлөлт илэрч байна.

Энэ нь уртын дуу нь тогтмол биш, харин цаг хугацааны явцад динамик өөрчлөлттэй, уян хатан бүтэцтэй болохыг харуулж байна.

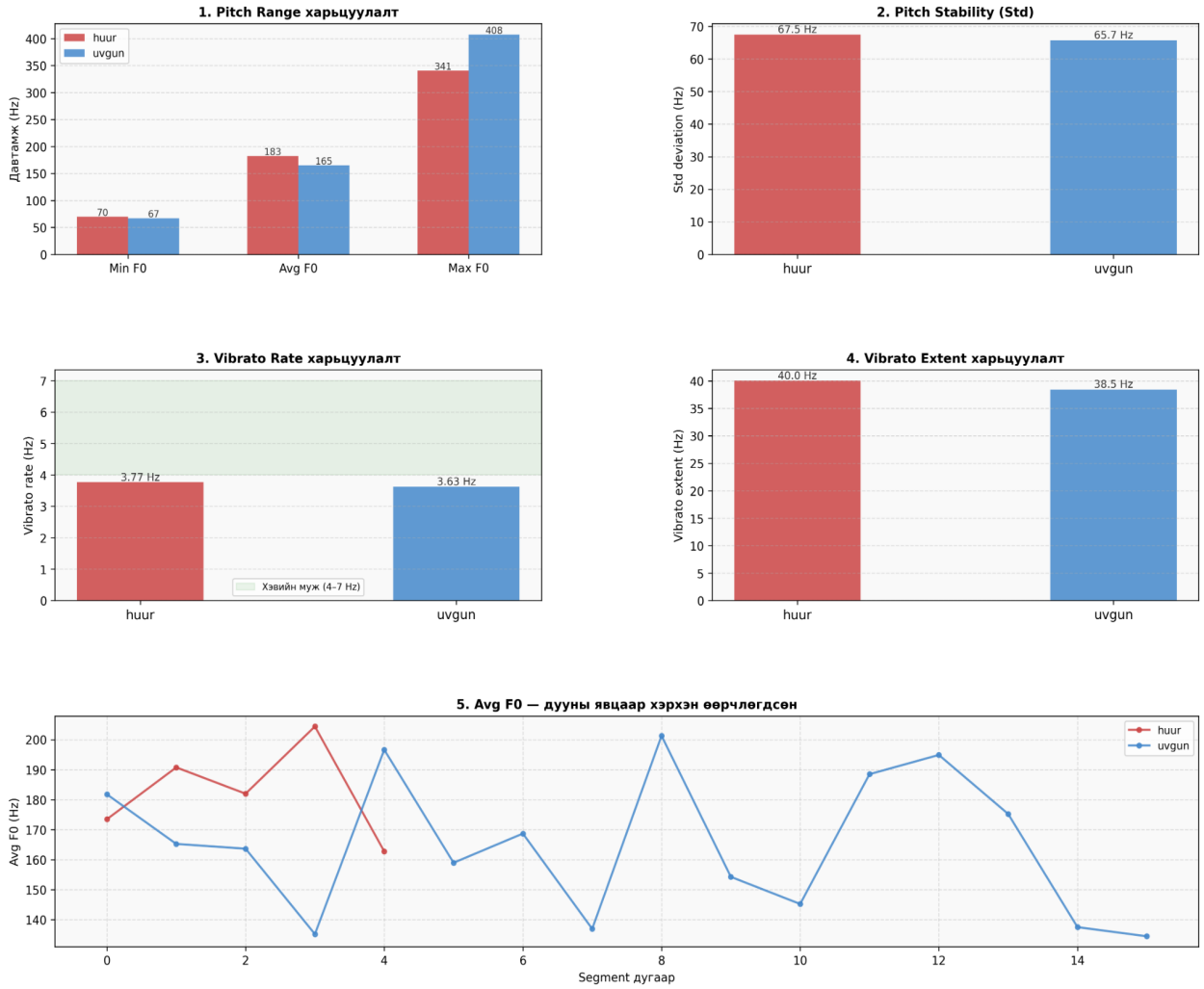
5) Спектрийн шинжилгээ: Log-Mel спектрограмын шинжилгээнээс дараах онцлогууд ажиглагдсан:

- Гармоник зурвасууд тод илэрсэн
- Давхар авианууд (overtone) олон давхаргатай
- Спектрийн энерги дунд болон өндөр давтамжийн мужид тархсан

Энэ нь уртын дуу нь баялаг спектрийн бүтэцтэй, тембрийн хувьд ялгарах шинжтэй болохыг илтгэнэ.

VIII. Дүгнэлт

Энэхүү судалгаанд Монгол уртын дууны акустик хэв маягийг тоон параметрээр тодорхойлох зорилгоор дууны үндсэн давтамж (F_0), давтамжийн муруй (pitch contour), чичирхийллийн шинж (vibrato), мөн Log-Mel спектрограмд суурилсан шинжилгээний аргыг санал болгож, туршилтаар харууллаа. Туршилтын үр дүнгээс харахад уртын дуу нь өргөн давтамжийн муж (ойролцоогоор 65–400 Hz), тогтвортой боловч динамик F_0 хэлбэлзэл, мөн тасралтгүй илрэх чичирхийллийн онцлог шинжээр тодорхойлогдож байна. Чичирхийллийн давтамж (vibrato rate) 3.3–4.2 Hz орчим байгаа нь уртын дууны удаан, уянгалаг дуулах хэв маягийг



Зураг 7. “Эртний сайхан” болон “Өвгөн шувуу” уртын дуунуудын акустик параметрийн харьцуулалт

илэрхийлж байгаа бол чичирхийллийн далайц (vibrato extent) 30–50 Hz-ийн хооронд хэлбэлзэж байгаа нь дууны илэрхийлэл, динамик шинжийг илтгэнэ.

Спектрийн шинжилгээний үр дүнгээр уртын дуу нь баялаг гармоник бүтэцтэй, олон давхар харь авиа (overtone) агуулсан болох нь тогтоогдсон бөгөөд энэ нь дууны тембр болон акустик онцлогийг тодорхойлох чухал үзүүлэлт болохыг харуулж байна.

Харьцуулсан шинжилгээний үр дүнгээс үзэхэд акустик параметрүүд нь дуулах хэв маяг, дуучны онцлог болон аялгууны ялгааг илэрхийлэх боломжтой болох нь тогтоогдлоо. Тухайлбал, давтамжийн муж болон чичирхийллийн далайц нь дууны илэрхийллийг, харин давтамжийн тогтвортой байдал болон чичирхийллийн давтамж нь дуулах техникийг тодорхойлоход чухал үүрэгтэй байна.

Иймээс энэхүү судалгаа нь Монгол уртын дуу нь өөрийн гэсэн өвөрмөц акустик “гарын үсэг” (acoustic signature)-тэй болохыг нотолж, уг гарын үсгийг F_0, чичирхийлэл болон спектрийн параметрүүдийн хослолоор илэрхийлэх боломжтойг харуулж байна.

Санал болгосон арга нь цаашид уртын дуу таних (audio recognition), дуучин ялгах (singer identification), дуулах хэв маяг ангилах (style classification), аудио хурууны хээ (audio fingerprinting) зэрэг олон төрлийн хэрэглээнд ашиглагдах боломжтой.

Гэвч энэхүү судалгаанд ашигласан өгөгдлийн хэмжээ хязгаарлагдмал тул гарсан үр дүнг ерөнхий дүгнэлт болгон авч үзэхэд тодорхой хязгаарлалт бий. Цаашдын судалгаанд илүү өргөн хэмжээний өгөгдөл (олон дуучин, олон төрөл жанр) ашиглан статистик баталгаажуулалт хийх, мөн машин сургалт (machine learning), гүн сургалт (deep learning)-д суурилсан автомат таних болон ангиллын систем хөгжүүлэх шаардлагатай байна.

Ашигласан ном

[1] UNESCO, “Urtiin Duu: Traditional Folk Long Song,” Intangible Cultural Heritage of Humanity, 2005.

[2] J. Sundberg, *The Science of the Singing Voice*. DeKalb, IL, USA: Northern Illinois University Press, 1987.

[3] I. Titze, *Principles of Voice Production*. Englewood Cliffs, NJ, USA: Prentice Hall, 1994.

[4] A. McLeod and G. Wyvill, “A smarter way to find pitch,” in *Proc. ICMC*, 2005.

[5] M. Mauch and S. Dixon, “pYIN: A fundamental frequency estimator using probabilistic threshold distributions,” in *Proc. ICASSP*, 2014.

[6] J. Sundberg, *The Science of the Singing Voice*. DeKalb, IL, USA: Northern Illinois University Press, 1987.

[7] A. McLeod and G. Wyvill, “A smarter way to find pitch,” in *Proc. Int. Computer Music Conf. (ICMC)*, 2005.

[8] M. Mauch and S. Dixon, “pYIN: A fundamental frequency estimator using probabilistic threshold distributions,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 659–663.

[9] I. R. Titze, *Principles of Voice Production*. Englewood Cliffs, NJ, USA: Prentice Hall, 1994.

[10] L. R. Rabiner and R. W. Schafer, *Digital Processing of Speech Signals*. Englewood Cliffs, NJ, USA: Prentice Hall, 1978.

[11] D. O’Shaughnessy, *Speech Communication: Human and Machine*. Piscataway, NJ, USA: IEEE Press, 2000.

[12] P. McLeod and G. Wyvill, "A smarter way to find pitch,"Proc. Int. Computer Music Conf. (ICMC), Barcelona, Spain, 2005.

[13] M. Mauch and S. Dixon, "pYIN: A fundamental frequency estimator using probabilistic threshold distributions,"Proc. IEEE ICASSP, Florence, Italy, pp. 659–663, 2014.

[14] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust Speech Recognition via Large-Scale Weak Supervision,"OpenAI Technical Report, 2022.

[15] V. Pratap, A. Tjandra, B. Shi et al., "Scaling Speech Technology to 1,000+ Languages,"Meta AI Research, arXiv:2305.13516, 2023.

[16] B. McFee, C. Raffel, D. Liang, D. P. W. Ellis, M. McVicar, E. Battenberg, and O. Nieto, "librosa: Audio and music signal analysis in Python,"Proc. 14th Python in Science Conf., Austin, TX, pp. 18–25, 2015.

[17] A. Bergevin, W. Zhou, and J. Reiss, “Overtone singing in Tibetan chanting and Mongolian khoomii,” J. Acoust. Soc. Am., vol. 127, no. 6, 2010.

[18] C. Pegg, *Mongolian Music, Dance, and Oral Narrative*. University of Washington Press, 2001.

[19] A. Wang, “An Industrial-Strength Audio Search Algorithm,” in Proc. ISMIR, 2003.

[20] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers*, 6th ed., Wiley, 2017.

[21] J. L. Devore, *Probability and Statistics for Engineering and the Sciences*, 9th ed., Cengage Learning, 2015.

[22] Student, “The probable error of a mean,” *Biometrika*, vol. 6, no. 1, pp. 1–25, 1908.

[23] K. Pearson, “Note on regression and inheritance in the case of two parents,” *Proc. Roy. Soc. London*, vol. 58, pp. 240–242, 1895.

[24] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed., Lawrence Erlbaum Associates, 1988.

САЛБАР ХУРАЛДААН – IV

Кибер аюулгүй байдал

№	**ҮБХИС-ийн хичээлийн I байр, 212 тоот-Сүбээдэй танхим** *Хуралдаан даргалагч: М.Золзаяа /ЗХЖШ/, Я.Дашдорж /ШУТИС/*
1	“Машин сургалтад суурилсан утасгүй сүлжээний халдлага илрүүлэх систем” *Шүрэнцэцэг Н, Дашдорж Ямхин*
2	“Хиймэл оюун ашигласан автоматжуулсан лог хяналт, шинжилгээний систем” *Zolboot Tsoodol, Erhem Otgonbayar, Bilguun E and Munkhbayar Bat-Erdene*
3	“Бага харилцан үйлчлэлтэй Honeypot-д суурилсан халдлага илрүүлэх систем хөгжүүлэх” *Munkhbayar Bat-Erdene, Dulguun Battur, Dulguun Batchuluun, Nominzul Bayantsatsral, and Ijiltsetseg Dawaajargal*
4	“IEEE 802.11 сүлжээ бүхий дроны сүргийн системийн auction-based даалгавар хуваарилах механизмын DDOS халдлагын нөлөөллийг симмуляциар үнэлэх нь” *Tuvshinbayar Purev, Enkhtur Ts, Dashdorj Ya, Uyangaakh Kh and Erdenebat Ch*
5	“Тайлбарлагдахуйц хиймэл оюуныг халдлага илрүүлэх системд хэрэглэх нь” *Usukhbayar Baldangombo, Enkhbayar Batjargal and Ugtakhbayar Naidansuren*
6	“Хиймэл оюунд суурилсан статик ба динамик шинжилгээг хослуулсан хортой программ илрүүлэх системийн загвар” *Oyungerel Sukhbaatar*

Машин сураглтанд суурилсан утасгүй сүлжээний халдлага илрүүлэх систем

Н.Шүрэнцэцэг
Монгол Улс, Улаанбаатар хот, ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль,
Мэдээллийн Сүлжээ Аюулгүй Байдлын Салбар
shurshuree959@gmail.com

Я.Дашдорж доктор (PhD) профессор
Монгол Улс, Улаанбаатар хот, ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль,
Мэдээллийн Сүлжээ Аюулгүй Байдлын Салбар
dashdorj@must.edu.mn

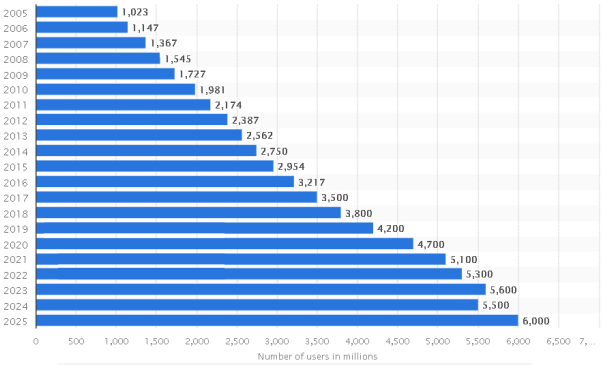
***Хураангуй*— Утасгүй сүлжээ нь нисгэгчгүй нисэх төхөөрөмж (UAV), IoT систем болон ухаалаг хотын дэд бүтцэд өргөн хэрэглэгдэж байгаа боловч нээлттэй дамжуулах орчин, динамик топологи болон төвлөрсөн бус архитектураас шалтгаалан кибер халдлагад өртөх өндөр эрсдэлтэй байдаг. Ялангуяа Blackhole, Wormhole, Sybil болон Flooding зэрэг халдлагууд нь сүлжээний найдвартай ажиллагаа болон өгөгдлийн бүрэн бүтэн байдалд сөргөөр нөлөөлдөг. Энэхүү судалгаанд утасгүй сүлжээний халдлагыг илрүүлэх зорилгоор машин сургалтад суурилсан intrusion detection system (IDS)-ийн загварыг боловсруулж, UAVIDS-2025 өгөгдлийн сан дээр туршив. Судалгаанд Logistic Regression, K-Nearest Neighbor, Decision Tree, Random Forest болон Support Vector Machine алгоритмуудыг ашиглан ангиллын загваруудыг байгуулж, гүйцэтгэлийг үнэлсэн. Судалгааны гол хувь нэмэр нь Chebyshev-ийн статистик тэнцэтгэлд суурилсан feature engineering аргачлалыг нэвтрүүлж, өгөгдлийн хазайлтыг илэрхийлэх шинэ шинж чанаруудыг үүсгэн, халдлагыг илрүүлэх чадварыг сайжруулсан явдал юм. Туршилтын үр дүнгээс статистик болон машин сургалтыг хослуулсан hybrid хандлага нь ангиллын алдааг бууруулж, илрүүлэлтийн үр дүнг сайжруулж байгааг харуулсан.**

Түлхүүр үг*— *Утасгүй сүлжээ, халдлага илрүүлэх систем (IDS), машин сургалт, Chebyshev feature engineering, кибер аюулгүй байдал

I. УДИРТГАЛ

Утасгүй сүлжээ (Wireless Networks) нь физик кабель ашиглахгүйгээр радио давтамжийн (RF) дохио болон цахилгаан соронзон долгионоор дамжуулан төхөөрөмжүүдийг холбох сүлжээний технологи юм [1]. Энэхүү технологи нь хөдөлгөөнт хэрэглэгчдэд уян хатан холболт олгох, сүлжээг өргөтгөхөд хялбар байх зэрэг давуу талтай байдаг.

Сүүлийн жилүүдэд утасгүй сүлжээ, IoT систем, ухаалаг хот, аж үйлдвэрийн автоматжуулалт болон нисгэгчгүй нисэх төхөөрөмж (UAV)-д суурилсан системүүдийн хэрэглээ эрчимтэй өсөж байна. Энэ өсөлт нь сүлжээний аюулгүй байдлыг судлах шаардлагыг улам нэмэгдүүлж байна. Олон улсын цахилгаан холбооны (ITU) 2024 оны тайланд дэлхий даяар интернэт хэрэглэгчийн тоо 5.5 тэрбум, 2025 онд 6.0 тэрбумд хүрч дэлхийн хүн амын 68%-ийг эзэлж байгаа нь сүлжээний хэрэглээ болон түүнтэй холбоотой аюулгүй байдлын эрсдэл нэмэгдэж байгааг харуулж байна.[1], [2]. Уг өсөлтийг Зураг 1-т харуулав.



Зураг 1. Утасгүй сүлжээний хэрэглэгчдийн өсөлт [16]

Зураг 1-ээс харахад интернэт хэрэглэгчдийн тоо тогтвортой өсөх хандлагатай байгаа нь утасгүй сүлжээний хэрэглээ нэмэгдэж байгааг илтгэнэ. Энэ нь сүлжээний аюулгүй байдлын асуудал улам чухал болж байгааг харуулж байна.

Утасгүй сүлжээний орчинд төрөл бүрийн кибер халдлагууд өргөн тархсан байдаг. Тухайлбал, Sybil халдлага нь нэг зангилаа олон хуурамч таних тэмдэг ашиглан сүлжээний итгэлцлийн механизмыг алдагдуулдаг бол Blackhole халдлага нь чиглүүлэлтийн протоколын сул талыг ашиглан пакет урсгалыг өөр дээрээ төвлөрүүлэн устгах замаар мэдээллийн алдагдал үүсгэдэг. Мөн Wormhole халдлага нь сүлжээний хоёр цэгийн хооронд нууц суваг үүсгэн өгөгдлийг богино замаар дамжуулах замаар routing системийг төөрөгдүүлдэг бөгөөд Flooding халдлага нь сүлжээний нөөцийг хэт ачааллуулж, үйлчилгээний саатал (DoS) үүсгэдэг [2], [3].

Эдгээр халдлагыг илрүүлэх, урьдчилан сэргийлэх зорилгоор халдлага илрүүлэх системийг өргөн хэрэглэгддэг. IDS нь сүлжээний урсгал болон системийн үйл ажиллагааг тасралтгүй хянаж, хэвийн бус үйлдлийг илрүүлэх хамгаалалтын чухал механизм юм. IDS-ийг ерөнхийд нь signature-based болон anomaly-based гэсэн үндсэн хоёр төрөлд ангилдаг бөгөөд эхнийх нь мэдэгдсэн халдлагын хэв маягийг илрүүлэхэд, харин хоёр дахь нь хэвийн бус үйлдлийг илрүүлэхэд чиглэдэг [4].

Гэвч орчин үеийн кибер халдлагууд улам бүр нарийн төвөгтэй, динамик шинжтэй болж байгаа тул уламжлалт дүрэмд суурилсан халдлага илрүүлэх системүүдийн үр нөлөө хязгаарлагдмал болж байна. Үүний улмаас машин сургалтад суурилсан халдлага илрүүлэх системүүдийг хөгжүүлэх судалгаа эрчимтэй явагдаж байгаа бөгөөд эдгээр аргууд нь их хэмжээний өгөгдлөөс халдлагын хэв маягийг автоматаар суралцаж, өндөр нарийвчлалтай илрүүлэх боломж олгодог [5], [6].

Иймээс энэхүү судалгааны ажил нь утасгүй сүлжээний орчинд үүсэх кибер халдлагыг машин сургалтын аргууд ашиглан илрүүлэх асуудлыг судлахад чиглэгдсэн болно. Тус судалгаанд Chebyshev-ийн статистик тэнцэтгэлд суурилсан шинж чанар боловсруулах аргачлалыг нэвтрүүлж, өгөгдлийн хазайлтыг илэрхийлэх шинэ шинж чанаруудыг үүсгэн, халдлага илрүүлэх гүйцэтгэлийг сайжруулах шинэ бөгөөд үр дүнтэй хосолмол хандлагыг боловсруулж, түүний үр нөлөөг туршилтаар баталсан.

II. ӨМНӨ СУДЛАГДСАН АЖИЛ

Утасгүй сүлжээ (Wireless Networks), IoT болон утасгүй мэдрэгчийн сүлжээ (Wireless Sensor Networks, WSN)-ийн хэрэглээ хурдацтай нэмэгдэж байгаа нь сүлжээний аюулгүй байдлын асуудлыг судалгааны чухал чиглэл болгож байна. Эдгээр системүүд нь ухаалаг хот, аж үйлдвэрийн автоматжуулалт, эрүүл мэндийн хяналт, нисгэгчгүй төхөөрөмжүүд зэрэг олон салбарт өргөн хэрэглэгдэж байгаа бөгөөд сүлжээний цар хүрээ, хэрэглээний эрчим нэмэгдэхийн хэрээр аюулгүй байдлын эрсдэл дагаад өсөж байна. Утасгүй сүлжээ нь нээлттэй дамжуулах орчин ашигладаг, зангилаануудын тооцоолох болон энергийн нөөц хязгаарлагдмал байдаг зэрэг онцлогоос шалтгаалан төрөл бүрийн кибер халдлагад өртөх өндөр эрсдэлтэй байдаг [8].

Ийм нөхцөлд халдлагыг илрүүлэх, урьдчилан сэргийлэх зорилгоор халдлага илрүүлэх системийн судалгаа өргөн хүрээнд хийгдэж байна. IDS нь сүлжээний урсгал, зан төлөв болон системийн үйл ажиллагааг тасралтгүй хянаж, хэвийн бус үйлдлийг илрүүлэх хамгаалалтын чухал механизм бөгөөд сүлжээний аюулгүй байдлыг хангахад чухал үүрэг гүйцэтгэдэг [9].

Судалгааны ажлуудад машин сургалтад суурилсан IDS системүүд нь уламжлалт танигдсан хэв маягт суурилсан аргуудаас илүү үр дүнтэй болохыг харуулж байна. Ялангуяа машин сургалтын алгоритмууд нь их хэмжээний өгөгдлөөс халдлагын хэв маягийг автоматаар суралцах, үл мэдэгдэх буюу шинэ төрлийн халдлагыг илрүүлэх боломжийг олгодог тул орчин үеийн кибер аюулгүй байдлын шийдлүүдэд өргөн хэрэглэгдэх болсон [10].

Random Forest, Decision Tree, Support Vector Machine зэрэг алгоритмууд intrusion detection асуудалд өргөн хэрэглэгдэж байгаа бөгөөд ангиллын өндөр нарийвчлал, тогтвортой гүйцэтгэл үзүүлдгээрээ онцлогтой. Pandey [11] Random Forest алгоритмд суурилсан IDS загварыг optimization техникээр сайжруулж, утасгүй мэдрэгчийн сүлжээний халдлагыг илрүүлэх үр дүнг мэдэгдэхүйц нэмэгдүүлсэн бол Arabiat [12] олон алгоритмыг харьцуулсан судалгаагаар Random Forest болон AdaBoost загварууд илүү өндөр гүйцэтгэлтэй байгааг тогтоосон байна.

Тодорхой төрлийн халдлагуудыг илрүүлэх чиглэлээр мөн олон судалгаа хийгдсэн. Тухайлбал, Alshehri [13] машин сургалтын арга ашиглан wormhole халдлагыг илрүүлэх загвар боловсруулж, өндөр нарийвчлалтай илрүүлэлт хийх боломжтойг харуулсан. Үүнээс гадна DoS болон flooding төрлийн халдлагыг илрүүлэхэд машин сургалтын аргууд үр дүнтэй болохыг Al-Sukkar [15] судалгаандаа баталсан байна.

Сүүлийн үеийн судалгаанд халдлага илрүүлэх системийн гүйцэтгэлийг сайжруулах зорилгоор олон төрлийн алгоритм, аргачлалыг нэгтгэсэн хосолмол хандлагыг хөгжүүлэхэд онцгой анхаарал хандуулж байна. Zhang [14] олон эх үүсвэрийн өгөгдлийг нэгтгэн боловсруулдаг SC-MLIDS framework боловсруулж, халдлага илрүүлэх үр дүнг сайжруулсан байна. Иймэрхүү төрлийн хосолмол хандлага нь олон төрлийн өгөгдлийн шинж чанарыг хамтад нь ашигласнаар илрүүлэлтийн нарийвчлал, найдвартай байдлыг нэмэгдүүлэх давуу талтай.

Эдгээр судалгааны үр дүнгээс үзэхэд машин сургалт болон хиймэл оюун ухаанд суурилсан халдлага илрүүлэх системүүд нь утасгүй сүлжээний халдлагыг илрүүлэхэд өндөр үр дүнтэй болох нь тогтоогдсон. Гэвч өгөгдлийн тэнцвэргүй байдал, оновчтой шинж чанар сонголт, бодит сүлжээний орчинд турших боломж, мөн загварын ерөнхийшүүлэх чадварыг сайжруулах зэрэг асуудлууд бүрэн шийдэгдээгүй хэвээр байна. Иймээс илүү үр дүнтэй, найдвартай, бодит орчинд хэрэгжүүлэх боломжтой халдлага илрүүлэх загвар боловсруулах шаардлага нь судалгааны чухал чиглэл хэвээр байна.

III. САНАЛ БОЛГОХ МЕХАНИЗМ

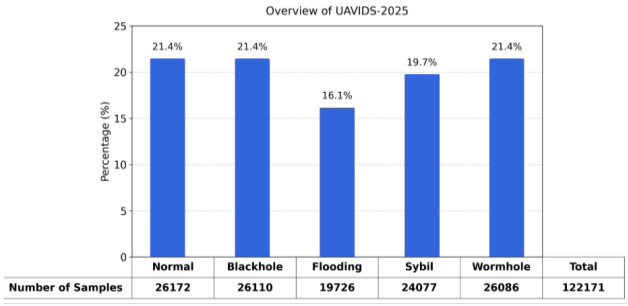
UAVIDS-2025 өгөгдлийн сан нь утасгүй UAV болон ad-hoc сүлжээний орчинд үүссэн сүлжээний урсгал мэдээллийг агуулсан бөгөөд халдлага илрүүлэх системийн (IDS) судалгаанд өргөн ашиглагддаг орчин үеийн өгөгдлийн сан юм. Энэхүү өгөгдөл нь сүлжээний урсгалын түвшинд цуглуулсан тул сүлжээний зан төлөв болон халдлагын хэв маягийг илүү нарийвчлалтай илэрхийлэх боломжтой.

Өгөгдлийн сан нь нийт 122,171 сүлжээний урсгал болон 23 баганаас бүрдэх бөгөөд үүнд 22 ширхэг шинж чанар болон 1 ангиллын шошго багтана. Шинж чанарууд нь сүлжээний урсгалын төлөв байдлыг тодорхойлох зориулалттай бөгөөд пакетын тоо, байтын хэмжээ, дамжуулах хурд, пакет алдагдал, саатал, хэлбэлзэл, нэвтрүүлэх чадвар зэрэг статистик үзүүлэлтүүдийг агуулдаг. Эдгээр үзүүлэлтүүд нь хэвийн болон хэвийн бус урсгалыг ялган танихад чухал үүрэг гүйцэтгэдэг.

Ангиллын хувьд өгөгдөл нь хэвийн урсгал, Blackhole халдлага, Wormhole халдлага, Sybil халдлага болон Flooding халдлага гэсэн таван төрлийн урсгалыг агуулдаг. Эдгээр халдлагууд нь утасгүй сүлжээний орчинд түгээмэл тохиолддог бөгөөд чиглүүлэлтийн үйл ажиллагаа, нөөцийн ашиглалт болон сүлжээний итгэлцлийн механизмыг алдагдуулах замаар системийн хэвийн ажиллагаанд сөргөөр нөлөөлдөг. Иймээс энэхүү өгөгдлийн сан нь халдлага илрүүлэх алгоритмуудыг турших, харьцуулахад тохиромжтой жишиг өгөгдөл болдог.

Мөн энэхүү өгөгдлийн сан нь симуляцлагдсан хэдий ч бодит орчинтой ойролцоо сүлжээний нөхцөл, тухайлбал хөдөлгөөнт зангилаа, утасгүй дамжуулалт, динамик чиглүүлэлтийн орчинд тулгуурлан үүсгэсэн тул халдлагын зан төлөвийг бодит нөхцөлд дөхүүлэн илэрхийлэх давуу талтай.

UAVIDS-2025 өгөгдлийн сан дахь урсгалын ангиллын тархалтыг Зураг 2-т харуулав.



Зураг 2. UAVIDS-2025 өгөгдлийн сан дахь урсгалын ангиллын тархалт

Энэхүү тархалтаас харахад ангиллуудын хооронд тодорхой хэмжээний тэнцвэргүй байдал ажиглагдаж байгаа бөгөөд энэ нь машин сургалтын загварын гүйцэтгэлд нөлөөлөх чухал хүчин зүйл болдог. Ийм төрлийн тархалтын онцлог нь өгөгдлийн хазайлтыг статистик аргаар үнэлэх боломжийг бүрдүүлдэг. Ялангуяа Chebyshev-ийн статистик хандлагыг ашигласнаар дундаж утгаас хэт хазайсан урсгалуудыг илрүүлж, хэвийн бус төлөвийг тодорхойлох боломжтой.

A. Өгөгдөл цэвэрлэх ба урьдчилсан боловсруулалт

Машин сургалтын алгоритмуудыг үр дүнтэй ашиглахын тулд өгөгдлийг урьдчилан боловсруулж, чанарыг сайжруулах шаардлагатай байдаг. Энэхүү судалгаанд UAVIDS-2025 өгөгдлийн санд дараах үндсэн preprocessing алхмуудыг хэрэгжүүлсэн.

- Missing value шалгах.
  Өгөгдлийн сан дахь дутуу утгуудыг шалгаж, шаардлагатай тохиолдолд устгах эсвэл тохирох статистик утгаар нөхөх боловсруулалтыг хийсэн. Энэ алхам нь өгөгдлийн бүрэн бүтэн байдлыг хангаж, загварын алдааг бууруулахад чухал үүрэгтэй.
- Кодчилол
  Категори болон текстэн хэлбэртэй шинж чанаруудыг машин сургалтын алгоритмд ашиглах боломжтой болгох зорилгоор тоон хэлбэрт хөрвүүлэв. Үүнд ангиллын шошгын кодчилол болон категори утгын кодчилолын аргуудыг ашигласан. Ингэснээр өгөгдлийг алгоритмд шууд оруулж боловсруулах боломж бүрдсэн.
- Normalization.
  Шинж чанаруудын хэмжээсийн ялгааг багасгах зорилгоор өгөгдлийг normalization хийж, нэг ижил масштабт оруулсан. Энэ нь алгоритмуудын сургалтын тогтвортой байдлыг сайжруулж, гүйцэтгэлд эерэг нөлөө үзүүлдэг.
- Train/Test split.
  Машин сургалтын загварыг сургах болон үнэлэх зорилгоор өгөгдлийг сургалтын (training) болон туршилтын (testing) хэсэгт хуваасан. Нийт өгөгдлийн 80%-ийг сургалтын өгөгдөл, 20%-ийг туршилтын өгөгдөл болгон ашигласан бөгөөд үүний үр дүнд сургалтын өгөгдөл 97,736 бичлэг, туршилтын өгөгдөл 24,435 бичлэг болсон. Сургалтын өгөгдлийг загвар сургахад, туршилтын өгөгдлийг гүйцэтгэлийг үнэлэхэд ашигласан.

. Өгөгдлийн хуваарилалтыг Хүснэгт 1-д үзүүлэв.

ХҮСНЭГТ 1. TRAIN/TEST SPLIT ХИЙСНИЙ ДАРААХ ӨГӨГДЛИЙН ХУВААРИЛАЛТ

Dataset хэсэг	**Бичлэгийн тоо**	**Хувь (%)**
Training dataset	97,736	80%
Testing dataset	24,435	20%
Нийт	122,171	100%

Мөн боловсруулсан өгөгдлийг дахин ашиглах боломжийг хангах үүднээс **train_processed.csv** болон **test_processed.csv** файлууд болгон хадгалсан

B. Feature сонголт ба шинж чанарын шинжилгээ

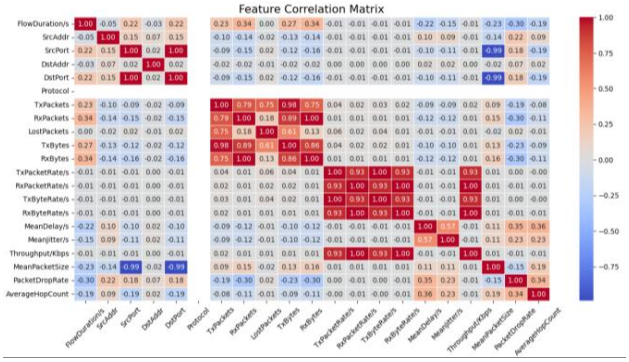
Машин сургалтын загварын гүйцэтгэлийг сайжруулахын тулд өгөгдлийн шинж чанаруудын нөлөөллийг шинжилж, хамгийн чухал үзүүлэлтүүдийг сонгон ашиглах шаардлагатай. Feature selection нь илүүдэл болон хамаарал багатай шинж чанаруудыг бууруулснаар загварын сургалтын хурд, тогтвортой байдал болон ангиллын нарийвчлалыг нэмэгдүүлэхэд чухал үүрэг гүйцэтгэдэг.

- *Feature Selection*

UAVIDS-2025 өгөгдлийн сан нь сүлжээний урсгалын олон төрлийн статистик шинж чанаруудыг агуулдаг бөгөөд эдгээр үзүүлэлтүүд нь халдлага илрүүлэхэд харилцан адилгүй нөлөө үзүүлдэг. Иймээс шинж чанаруудын ач холбогдлыг тодорхойлж, илүү нөлөө бүхий feature-үүдийг сонгон ашиглах зорилгоор feature selection хийсэн.

- *Correlation Analysis*

Шинж чанаруудын хоорондын хамаарлыг тодорхойлох зорилгоор корреляцийн шинжилгээ хийж, үр дүнг heatmap хэлбэрээр Зураг 3-т харуулав.



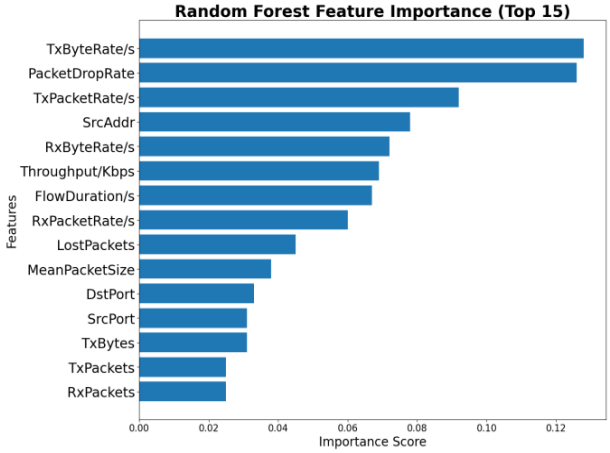
Зураг 3. UAVIDS-2025 өгөгдлийн сан дахь шинж чанаруудын корреляцийн матриц

Зураг 3-аас харахад зарим шинж чанарууд хоорондоо өндөр хамааралтай байгааг ажиглаж болно. Ялангуяа TxPacketRate/s, RxPacketRate/s, TxByteRate/s, RxByteRate/s зэрэг урсгалын хурдтай холбоотой үзүүлэлтүүд хоорондоо хүчтэй корреляцитай байна. Мөн TxPackets, RxPackets, TxBytes, RxBytes зэрэг пакет дамжуулалтын хэмжээг илэрхийлэх үзүүлэлтүүд харилцан хамааралтай байгаа нь эдгээр нь ижил төрлийн мэдээлэл агуулж байгааг илтгэнэ.

Ийм шинжилгээ нь илүүдэл шинж чанаруудыг тодорхойлох, давхардсан мэдээллийг багасгах, улмаар загварын гүйцэтгэлийг сайжруулахад чухал ач холбогдолтой.

- *Random Forest feature importance*

Шинж чанаруудын халдлага илрүүлэхэд үзүүлэх нөлөөллийг тодорхойлох зорилгоор Random Forest алгоритмыг ашиглан feature importance шинжилгээ хийсэн. Энэхүү арга нь decision tree бүрт хуваалт хийх үед үүсэх impurity бууралтыг тооцоолж, бүх модны үр дүнг нэгтгэн шинж чанар бүрийн ач холбогдлын оноог тодорхойлдог. Feature importance-ийн үр дүнг Зураг 4-т харуулав.



Зураг 4. Random Forest алгоритмаар тодорхойлсон чухал шинж чанарууд

Энэ нь халдлагын үед сүлжээний урсгалын параметрүүд мэдэгдэхүйц өөрчлөгдөж байгааг илтгэж байгаа бөгөөд дээрх шинж чанарууд intrusion detection системийн ангиллын загварт голлох үүрэг гүйцэтгэж байгааг харуулж байна [18].

Иймд feature importance шинжилгээ нь илүүдэл шинж чанаруудыг бууруулах, загварын сургалтын үр ашгийг нэмэгдүүлэх, цаашлаад intrusion detection системийн гүйцэтгэлийг сайжруулахад чухал ач холбогдолтой юм.

Эдгээр өндөр ач холбогдол бүхий шинж чанаруудыг Chebyshev-ийн статистик хандлагад ашигласнаар өгөгдлийн хазайлтыг хэмжих, хэвийн бус урсгалыг илрүүлэх боломжийг нэмэгдүүлнэ.

C. Chebyshev-based Feature Engineering

Энэхүү судалгааны гол шинэлэг хувь нэмэр нь Chebyshev-ийн статистик тэнцэтгэлд суурилсан feature engineering аргачлалыг утасгүй сүлжээний халдлага илрүүлэх асуудалд нэвтрүүлсэн явдал юм. Chebyshev-ийн тэнцэтгэл нь санамсаргүй хувьсагчийн утга дундаж утгаас тодорхой хэмжээнээс хэтрэх магадлалыг дараах байдлаар тодорхойлдог [19].

$$P(|X-\mu| \geq k\sigma) \leq 1-\frac{1}{k^2} \qquad (1)$$

Энд μ нь дундаж утга, σ нь стандарт хазайлт, k нь эерэг тогтмол утга юм. Энэхүү тэнцэтгэл нь тархалтын хэлбэрээс үл хамааран өгөгдлийн хазайлтыг үнэлэх боломж олгодог бөгөөд Anomaly detection-д өргөн хэрэглэгддэг[20], [21].

1) Онолын үндэслэл ба хэрэглээ

Chebyshev-ийн тэнцэтгэлээс үзэхэд өгөгдлийн ихэнх утгууд дундаж утгын орчимд, (2)-р томьёонд үзүүлсэн интервалд байрлах хандлагатай байна:

$$\mu - k\sigma \leq X \geq \mu + k\sigma \qquad (2)$$

(2)-р томьёогоор тодорхойлогдох энэхүү интервалаас гадуур байрлах утгууд нь статистикийн хувьд бага магадлалтай тул хэвийн бус төлөв гэж үзэх боломжтой [22]. Иймээс (2)-р томьёог сүлжээний урсгалын хэвийн болон хэвийн бус төлөвийг ялгах статистик шалгуур болгон ашиглав.

2) Feature Engineering хийх аргачлал

UAVIDS-2025 өгөгдлийн сан дахь халдлагыг илрүүлэх үр ашгийг нэмэгдүүлэх зорилгоор Random Forest алгоритмын ач холбогдлын үнэлгээнд үндэслэн дараах өндөр нөлөө бүхий шинж чанаруудыг сонгон ашиглав. Үүнд:

- TxByteRate/s
- PacketDropRate
- TxPacketRate/s
- RxByteRate/s
- Throughput/Kbps
- FlowDuration

Сонгогдсон шинж чанар бүрийн хувьд Chebyshev-ийн статистик шалгуурт суурилсан шинж чанарын инженерчлэлийн дараах үе шаттайгаар хэрэгжүүлэв.

1. Сургалтын өгөгдлийн багцаас тухайн шинж чанар бүрийн дундаж утга μ болон стандарт хазайлт σ-г тооцоолов.
2. Хэвийн бус хазайлтыг тодорхойлох босго утга болгон k=2 болон k=3-ыг туршилтаар сонгон ашиглав.
3. Өгөгдлийн тухайн утга дундаж утгаас хэдэн стандарт хазайлтаар зөрж байгааг (3)-р томьёогоор үнэлэв.

$$\frac{|x_i-\mu|}{\sigma} > k \qquad (3)$$

Энд x_i нь тухайн урсгалын шинж чанарын утга μ нь дундаж утга, σ нь стандарт хазайлт, k нь босго параметр болно.
(3)-р томьёоны нөхцөл биелсэн тохиолдолд тухайн ажиглалтыг хэвийн бус төлөвтэй гэж тэмдэглэв. Ингэснээр үндсэн урсгалын шинж чанарууд дээр статистик хазайлтыг илэрхийлэх нэмэлт шинжүүд үүсч, машин сургалтын загварын ангиллын гүйцэтгэлийг сайжруулах боломж бүрдсэн.

3) Шинэ Feature үүсгэх

Chebyshev-ийн статистик хандлагад үндэслэн сүлжээний урсгалын хэвийн төлөвөөс хазайх түвшинг тоон утгаар илэрхийлэх шинэ шинж чанарууд үүсгэв. Эдгээрээс голчлон Chebyshev Score шинжийг ашигласан.

- Chebyshev Score

Тухайн урсгалын шинж чанарын утга дундаж утгаас хэдэн стандарт хазайлтаар зөрж байгааг (4)-р томьёогоор тооцоолов.

$$C_i = \frac{|x_i - \mu|}{\sigma} \qquad (4)$$

(4)-р томьёогоор тооцогдох утга их байх тусам тухайн өгөгдөл хэвийн төлөвөөс илүү их хэмжээгээр хазайсан гэж үзнэ. Харин бага утгатай тохиолдолд тухайн ажиглалт хэвийн төлөвт ойр байгааг илэрхийлнэ.

- Aggregated Chebyshev Score

Олон шинж чанарын нийлбэр хазайлтыг нэгтгэн үнэлэх зорилгоор Aggregated Chebyshev Score шинжийг тодорхойлов. Энэхүү хэмжигдэхүүн нь сонгогдсон бүх шинж чанарын дундажаас хазайсан түвшинг нэгтгэн илэрхийлэх бөгөөд тухайн урсгалын ерөнхий хэвийн бус байдлыг хэмжихэд ашиглагдана.

Тухайн нийлбэр хазайлтыг (5)-р томьёогоор тооцоолов.

$$C_{total} = \frac{1}{n}\sum_{j=1}^{n} \frac{|x_{ij} - \mu_j|}{\sigma_j} \qquad (5)$$

j нь сонгогдсон шинж чанаруудыг илэрхийлнэ.

- Хэвийн бус төлөвийн үзүүлэлт

Chebyshev Score-д үндэслэн хэвийн болон хэвийн бус төлөвийг ялгах binary шинж чанарыг дараах байдлаар тодорхойлсон:

$$A_i = \begin{cases} 1, & \text{if } C_i > k \\ 0, & \text{otherwise} \end{cases} \qquad (6)$$

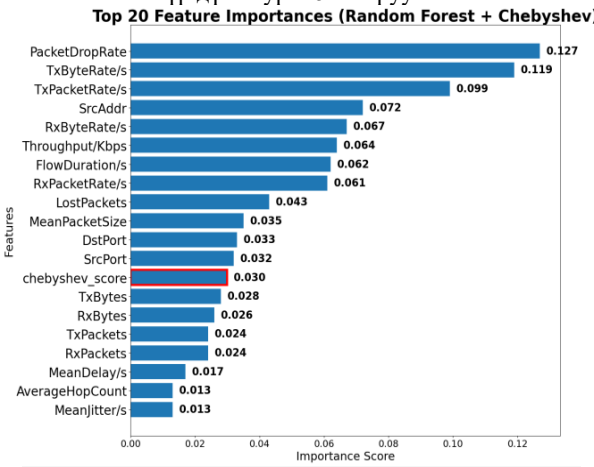
Энд A_i нь хэвийн бус төлөвийн үзүүлэлт, C_i нь Chebyshev-ийн үнэлгээ, k нь урьдчилан сонгосон босго утга болно. $A_i = 1$ бол тухайн урсгал хэвийн бус төлөвтэй, харин $A_i = 0$ бол хэвийн төлөвтэй гэж үзнэ.
Энэхүү шинэ шинж чанар нь дараах давуу талтай.

- Сүлжээний урсгалын статистик хазайлтыг энгийн хэлбэрээр илэрхийлнэ.
- Халдлагын үед үүсэх огцом өөрчлөлтийг илрүүлэх боломж олгоно.
- Ангиллын загварт нэмэлт ялгах чадвар өгч, ангиллын гүйцэтгэлийг сайжруулахад дэмжлэг үзүүлнэ.

Ингэснээр үндсэн урсгалын шинж чанарууд дээр нэмэлтээр хэвийн бус төлөвийг шууд илэрхийлэх хоёр утгат шинж үүсэж, халдлагын урсгалыг илүү үр дүнтэй таних боломж бүрдсэн.

4) Машин сургалттай хослуулсан хэрэглээ

Үүсгэсэн Chebyshev-д суурилсан шинж чанаруудыг машин сургалтын алгоритмуудтай хослуулснаар feature space-ийг өргөжүүлж, ангиллын загварын ялгах чадварыг сайжруулах боломж бүрдсэн. Энэхүү хандлага нь статистик болон машин сургалтыг нэгтгэсэн hybrid statistical–ML аргачлалд хамаарах бөгөөд өгөгдлийн хазайлтыг илэрхийлэх нэмэлт шинж чанарууд нь загварын гүйцэтгэлд эерэг нөлөө үзүүлдэг [25], [26].
Ялангуяа Chebyshev-based feature-үүд нь сүлжээний урсгалын хэвийн төлөвөөс хазайсан өөрчлөлтүүдийг илүү мэдрэмжтэй илрүүлэх боломж олгодог. Үүний үр дүнд халдлагын үед илэрдэг огцом өөрчлөлтүүдийг илүү тодорхой барьж авч, ангиллын нарийвчлалыг нэмэгдүүлэхэд чухал үүрэг гүйцэтгэж байна.
Random Forest алгоритмд Chebyshev-based feature-үүдийг нэгтгэн ашигласны дараах feature importance шинжилгээний үр дүнг Зураг 5-т харуулав.



Зураг 5. Chebyshev + Random Forest алгоритмаар тодорхойлсон чухал шинж чанарууд

Зураг 5-аас харахад Chebyshev-д суурилсан шинж чанаруудыг нэмснээр feature importance-ийн тархалт өөрчлөгдөж, статистик хазайлтыг илэрхийлэх шинэ шинж чанарууд илүү өндөр ач холбогдолтой болсон байна. Энэ нь Chebyshev-based feature-үүд нь сүлжээний урсгалын хэвийн бус өөрчлөлтийг илүү сайн илэрхийлж, ангиллын загварын ялгах чадварыг нэмэгдүүлж байгааг харуулж байна.

D. Машин сургалтын алгоритмуудыг сургах

Утасгүй сүлжээний халдлагыг илрүүлэх зорилгоор хэд хэдэн supervised learning алгоритмуудыг ашиглан ангиллын загваруудыг сургаж туршсан. Энэхүү судалгаанд Logistic Regression, K-Nearest Neighbor (KNN), Decision Tree, Random Forest, Support Vector Machine (SVM) зэрэг өргөн хэрэглэгддэг алгоритмуудыг ашигласан.

Эдгээр алгоритмууд нь UAVIDS-2025 өгөгдлийн сан дахь шинж чанарууд (жишээлбэл, TxByteRate/s, PacketDropRate, TxPacketRate/s, Throughput/Kbps, FlowDuration гэх мэт)-ыг ашиглан сүлжээний урсгалыг Normal traffic болон халдлагын төрлүүд (Blackhole, Wormhole, Sybil, Flooding)-д ангилах зорилготой.
Загваруудыг сургалтын өгөгдөл дээр сургаж, туршилтын өгөгдөл ашиглан гүйцэтгэлийг үнэлсэн.

1)Logistic Regression

Logistic Regression нь ангиллын асуудлыг шийдвэрлэхэд өргөн хэрэглэгддэг шугаман ангиллын алгоритм бөгөөд оролтын шинж чанаруудын жигнэсэн нийлбэр дээр sigmoid функц ашиглан тухайн өгөгдөл тодорхой ангилалд хамаарах магадлалыг тооцоолдог. Энэхүү арга нь энгийн бүтэцтэй, хурдан сурдаг, мөн шинж чанар бүрийн нөлөөллийг тайлбарлах боломжтой давуу талтай. Тухайн өгөгдөл эерэг ангилалд хамаарах магадлалыг тооцох Logistic Regression алгоритмын математик загварыг (7)-р томьёогоор илэрхийлэв.

$$P(y = 1/x) = \frac{1}{1+e^{-(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_n x_n)}} \quad (7)$$

- $x_1, x_2, \ldots, x_n$ нь UAVIDS-2025 өгөгдлийн сангийн шинж чанарууд (TxByteRate/s, PacketDropRate, Throughput/Kbps гэх мэт)
- β_i ухайн шинж чанар бүрийн жин болно.

Энэ алгоритм нь feature бүрийн нөлөөллийг шугаман байдлаар үнэлдэг.

2)K-Nearest Neighbor (KNN)

KNN нь жишээнд суурилсан сургалтын алгоритм бөгөөд шинэ өгөгдлийн ангиллыг сургалтын багц дахь хамгийн ойр kkk хөрш өгөгдлүүдийн ангилалд үндэслэн тодорхойлдог. Энэхүү арга нь урьдчилан загвар байгуулах шаардлагагүй бөгөөд өгөгдлийн ижил төстэй байдалд тулгуурлан ангилалт хийдгээрээ онцлог юм. KNN алгоритмд хоёр өгөгдлийн хоорондох зайг ихэвчлэн Евклидийн зайгаар хэмждэг бөгөөд үүнийг (8)-р томьёогоор илэрхийлнэ.

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \quad (8)$$

- Энд x_i болон y_iнь харьцуулж буй хоёр сүлжээний урсгалын i-р шинж чанарын утга, n нь нийт шинж чанарын тоо болно.

(8)-р томьёогоор тооцогдох зай бага байх тусам тухайн хоёр урсгал хоорондоо илүү төстэй гэж үзнэ. Харин зай их байх нь урсгалуудын төлөв ялгаатай байгааг илэрхийлнэ.

3)Decision Tree

Decision Tree нь өгөгдлийг шаталсан мод бүтэц ашиглан ангилах хяналттай сургалтын алгоритм юм. Энэхүү арга нь өгөгдлийг тодорхой нөхцөл, дүрэмд үндэслэн дэд бүлгүүдэд хуваах, эцэст нь ангиллын шийдвэр гаргадаг. Модны зангилаа бүрт хамгийн их ялгах чадвартай шинж чанарыг сонгон салбарлуулдаг. Шинж чанар сонгох үед зангилааны цэвэршилтийн хэмжүүр болгон Gini impurity-г өргөн хэрэглэдэг бөгөөд үүнийг (9)-р томьёогоор илэрхийлнэ.

$$Gini = 1 - \sum_{i=1}^{n} p_i^2 \quad (9)$$

- Энд p_i нь тухайн зангилаанд харгалзах i-р ангиллын магадлал, nnn нь нийт ангиллын тоо болно.

(9)-р томьёогоор тооцогдох утга бага байх тусам тухайн зангилааны өгөгдөл нэг ангилалд төвлөрсөн буюу илүү цэвэр байна гэж үзнэ. Иймээс алгоритм нь Gini impurity хамгийн их буурах шинж чанарыг сонгон өгөгдлийг хуваадаг

4)Support Vector Machine (SVM)

SVM алгоритм нь өгөгдлийг ангилахдаа ангиллуудын хооронд хамгийн их зайтай шийдвэрийн хил тогтоодог хяналттай сургалтын алгоритм юм. Энэхүү арга нь өөр өөр ангиллын өгөгдлүүдийг хамгийн сайн тусгаарлах хавтгайг олж, шинэ өгөгдлийг тухайн хил дээр үндэслэн ангилдаг. SVM алгоритмын шийдвэрийн хил буюу ангилах хавтгайг (10)-р томьёогоор илэрхийлнэ.

$$w * x + b = 0 \quad (10)$$

- x нь шинж чанарын вектор (TxByteRate/s, PacketDropRate гэх мэт) , w нь жингийн вектор, b нь тогтмол коэффициент болно.

(10)-р томьёогоор тодорхойлогдох хавтгай нь хоёр ангиллын хоорондох зайг хамгийн их байлгах байдлаар сонгогддог. Иймээс хэвийн (Normal) болон халдлагат (Attack) урсгалуудыг хамгийн сайн ялгах шийдвэрийн хил тогтоох боломжтой.

5)Random Forest

Random Forest нь олон decision tree-ийн нийлбэрт суурилсан хяналттай сургалтын арга юм. Энэхүү алгоритм нь bootstrap sampling болон санамсаргүй шинж чанар сонголт ашиглан олон мод сургаж, тэдгээрийн хамтарсан шийдвэрээр эцсийн ангиллыг тодорхойлдог. Ийм бүтэц нь нэг модтой харьцуулахад илүү тогтвортой, нарийвчлал өндөртэй байдаг.
Random Forest алгоритмын эцсийн ангиллын шийдвэрийг (11)-р томьёогоор илэрхийлнэ.

$$\hat{y} = mode\{T_1(x), T_2(x), \cdots, T_m(x)\} \quad (11)$$

- $T_i(x) \rightarrow i$-р decision tree-ийн ангиллын үр дүн, mнь нийт модны тоо, y нь эцсийн ангиллын шийдвэр болно.

(11)-р томьёонд үзүүлснээр бүх модны гаргасан ангиллын үр дүнгээс хамгийн олон санал авсан ангиллыг эцсийн үр дүн болгон сонгодог. Иймээс нэг модны алдаанаас хамаарах нөлөө багасч, загварын ерөнхийшүүлэх чадвар сайжирдаг. UAVIDS-2025 өгөгдлийн сан дээр TxByteRate/s, PacketDropRate, Throughput зэрэг шинж чанарууд ангилалд өндөр ач

холбогдолтой нөлөө үзүүлсэн. Иймээс эдгээр шинж чанарууд халдлагын урсгалыг ялгахад чухал мэдээлэл агуулдаг гэж үзэв.

IV. ТУРШИЛТЫН ҮР ДҮН БА ХЭЛЭЛЦҮҮЛЭГ

Туршилтын орчин

Утасгүй сүлжээний орчин дахь халдлагыг илрүүлэх машин сургалтын загваруудын гүйцэтгэлийг үнэлэх зорилгоор UAVIDS-2025 өгөгдлийн санг ашигласан. Тус өгөгдлийн сан нь UAV (Unmanned Aerial Vehicle) сүлжээний орчинд бүртгэгдсэн хэвийн болон халдлагат сүлжээний урсгалуудыг агуулдаг бөгөөд Blackhole attack, Flooding attack, Sybil attack, Wormhole attack болон Normal traffic гэсэн таван ангиллаас бүрдэнэ.

Туршилтын өмнө өгөгдөлд урьдчилсан боловсруулалт хийж, шаардлагатай feature-үүдийг сонгон авч, өгөгдлийг сургалтын болон туршилтын багц болгон хуваасан. Сургалтын өгөгдлийг загвар сургахад, туршилтын өгөгдлийг загварын ерөнхийшүүлэх чадвар болон бодит нөхцөл дэх гүйцэтгэлийг үнэлэхэд ашигласан.

Logistic Regression, K-Nearest Neighbor (KNN), Decision Tree, Random Forest болон Support Vector Machine (SVM) алгоритмуудыг ашигласан. Туршилтыг Python орчинд хэрэгжүүлж, өгөгдөл боловсруулахад Pandas болон NumPy, загвар сургах болон үнэлэхэд Scikit-learn, үр дүнг дүрслэхэд Matplotlib болон Seaborn сангуудыг ашигласан.

Мөн өгөгдлийн статистик шинж чанарыг илүү тодорхой илэрхийлэх зорилгоор Chebyshev-ийн статистик тэнцэтгэлд суурилсан feature engineering аргачлалыг ашигласан. Уг аргачлал нь feature-үүдийн дундаж болон стандарт хазайлтад үндэслэн хэвийн бус утгуудыг илрүүлэх бөгөөд сүлжээний урсгалын хазайлтыг илүү сайн тодорхойлох боломж олгодог. Үүний үр дүнд машин сургалтын загваруудад илүү ялгарахуйц feature space үүсэж, халдлагын илрүүлэлтийн гүйцэтгэлийг сайжруулах нөхцөл бүрдсэн.

A. Үнэлгээний үзүүлэлтүүд

Машин сургалтын загваруудын халдлага илрүүлэх гүйцэтгэлийг иж бүрэн үнэлэх зорилгоор ангиллын загварын үр ашгийг тодорхойлох түгээмэл хэрэглэгддэг хэд хэдэн үнэлгээний үзүүлэлтүүдийг ашигласан. Эдгээр үзүүлэлтүүд нь загварын гүйцэтгэлийг олон талаас нь үнэлэх боломж олгодог бөгөөд ялангуяа кибер аюулгүй байдлын орчинд буруу ангилалтын нөлөөллийг тодорхойлоход онцгой ач холбогдолтой байдаг. Тухайлбал, хуурамч эерэг (False Positive) болон хуурамч сөрөг (False Negative) ангиллууд нь халдлага илрүүлэх системийн найдвартай байдалд шууд нөлөөлдөг. False Positive өндөр байх нь хэвийн урсгалыг халдлага гэж буруу ангилах эрсдэлийг нэмэгдүүлдэг бол False Negative өндөр байх нь бодит халдлагыг илрүүлж чадахгүй байх аюултай нөхцөлийг үүсгэнэ.

Ашигласан үнэлгээний хэмжүүрүүдийн үндэс нь андуурaлтын матриц (confusion matrix) юм. Андуурaлтын матриц нь ангиллын загварын таамагласан үр дүн болон бодит ангиллын хоорондын хамаарлыг харуулдаг бөгөөд дараах дөрвөн үндсэн бүрэлдэхүүн хэсгээс тогтоно.

- True Positive (TP): халдлага бодитоор байсан бөгөөд систем зөв илрүүлсэн тохиолдол
- True Negative (TN): халдлага байгаагүй бөгөөд систем зөв ангилсан тохиолдол
- False Positive (FP): халдлага байгаагүй боловч систем халдлага гэж буруу ангилсан тохиолдол
- False Negative (FN): халдлага бодитоор байсан боловч систем илрүүлж чадаагүй тохиолдол

Уг судалгаа нь олон ангилалт (multi-class classification) хэлбэрийн асуудлыг авч үзэж байгаа тул дээрх хэмжигдэхүүнүүдийг ангилал бүрийн хувьд one-vs-rest зарчмаар тооцоолсон.

Accuracy нь нийт өгөгдлийн дундаас зөв ангилсан тохиолдлын эзлэх хувийг илэрхийлдэг.. Accuracy хэмжүүрийн математик илэрхийллийг (12)-р томьёогоор үзүүлэв.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \qquad (12)$$

Precision нь загварын халдлага гэж ангилсан тохиолдлуудын дундаас үнэхээр халдлага байсан тохиолдлын хувийг илэрхийлдэг бөгөөд хуурамч эерэг ангиллыг бууруулах чадварыг үнэлэхэд чухал ач холбогдолтой.

$$Precision = \frac{TP}{TP+FP} \qquad (13)$$

Recall буюу Sensitivity нь бодит халдлагуудын дундаас хэдэн хувийг систем зөв илрүүлж чадсаныг хэмждэг. Энэ нь халдлагыг алдалгүй илрүүлэх чадварыг үнэлэхэд ашиглагдана. Precision хэмжүүрийн математик илэрхийллийг (14)-р томьёогоор үзүүлэв.

$$Recall = \frac{TP}{TP+FN} \qquad (14)$$

Recall өндөр байх нь бодит халдлагыг илрүүлэх чадвар сайн байгааг илтгэнэ. Кибер аюулгүй байдлын нөхцөлд энэхүү үзүүлэлт онцгой ач холбогдолтой бөгөөд илрүүлж чадаагүй халдлага (FN) нь бодит системд ноцтой эрсдэл дагуулж болзошгүй.

F1-score нь Precision болон Recall хэмжүүрүүдийн гармоник дундаж бөгөөд ангиллын загварын ерөнхий гүйцэтгэлийг нэгтгэн үнэлэхэд өргөн хэрэглэгддэг.

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision+Recall} = \frac{2TP}{2TP+FP+FN} \qquad (15)$$

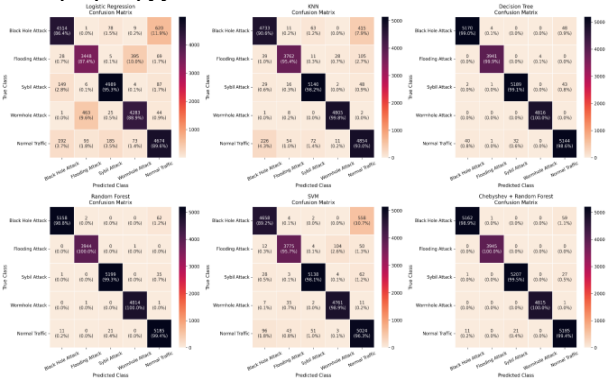
F1-score өндөр байх нь загвар халдлагыг зөв илрүүлэхийн зэрэгцээ буруу ангилалтын түвшинг харьцангуй бага байлгаж чадсаныг илтгэнэ.

B. Confusion Matrix

Confusion Matrix (андуурaлтын матриц) нь ангиллын (classification) загварын гүйцэтгэлийг үнэлэхэд өргөн хэрэглэгддэг хүснэгтэн хэлбэрийн арга бөгөөд бодит ангилал болон загварын таамагласан ангиллын хоорондын хамаарлыг нарийвчлан харуулдаг. Энэхүү матриц нь зөв болон буруу ангилалтын тархалтыг илэрхийлж, загварын алдааны бүтэц болон аль ангиллууд хоорондоо андуурагдаж байгааг тодорхойлох боломжийг олгодог.

Confusion matrix-д мөрүүд нь бодит (true) ангиллыг, баганууд нь загварын таамагласан (predicted) ангиллыг илэрхийлдэг бөгөөд ингэснээр ангиллын загварын гүйцэтгэлийг илүү гүнзгий түвшинд шинжлэх боломж бүрддэг.

5-р зурагт Logistic Regression, KNN, Decision Tree, Random Forest, SVM болон Chebyshev + Random Forest загваруудын ангиллын үр дүнг confusion matrix хэлбэрээр харуулав.



Зураг 6. Машин сургалтын алгоритмуд болон Chebyshev + Random Forest загварын андууралтын матрицын харьцуулалт

Зурагт харагдаж байгаагаар алгоритмуудын гүйцэтгэл хоорондоо мэдэгдэхүйц ялгаатай байна. Logistic Regression болон SVM алгоритмуудын хувьд Normal traffic болон Blackhole ангиллууд дээр харьцангуй их андуурал ажиглагдаж байгаа бөгөөд энэ нь эдгээр аргуудын шугаман шийдвэрийн хил нь өгөгдлийн нарийн бүтэц, тархалтыг бүрэн илэрхийлж чадахгүй байгаатай холбоотой юм.

KNN алгоритмын хувьд ихэнх халдлагын ангиллууд дээр дунджаас дээш гүйцэтгэл үзүүлсэн боловч Normal traffic болон Blackhole ангиллуудын хооронд тодорхой хэмжээний андуурал ажиглагдаж байна. Энэ нь feature space дахь өгөгдлийн тархалтын давхцалтай холбоотой байж болохыг илтгэнэ.

Харин Decision Tree болон Random Forest алгоритмуудын хувьд бараг бүх ангилал дээр өндөр нарийвчлалтай үр дүн гарсан бөгөөд ялангуяа Wormhole болон Flooding халдлагын ангилал дээр бараг төгс ангилалт хийж байгаа нь ажиглагдаж байна. Энэ нь модонд суурилсан алгоритмууд нь өгөгдлийн шугаман бус хамаарлыг илүү үр дүнтэй суралцаж байгааг харуулж байна.

Онцгойлон Chebyshev-based feature engineering ашигласан Random Forest загварын хувьд андууралтын матриц илүү цэвэр (clean) болж, буруу ангилалтын тоо цөөрсөн байна. Ялангуяа:

- Sybil халдлагын ангиллын алдаа буурсан
- Flooding болон Normal traffic хоорондын андуурал багассан
- Blackhole ангиллын зөв илрүүлэлт нэмэгдсэн

Мөн Normal traffic ангиллын зөв ангилалтын түвшин нэмэгдсэн нь Chebyshev transformation нь хэвийн болон хэвийн бус өгөгдлийн ялгааг илүү тодорхой болгож байгааг харуулж байна.

Эдгээр үр дүнгээс харахад Chebyshev-based feature engineering нь өгөгдлийн статистик хазайлтыг илүү сайн илэрхийлж, машин сургалтын загварын ялган ангилах чадварыг сайжруулж байгааг баталж байна. Түүнчлэн False Positive болон False Negative утгууд буурсан нь intrusion detection системийн найдвартай байдлыг нэмэгдүүлж байгааг илтгэнэ.

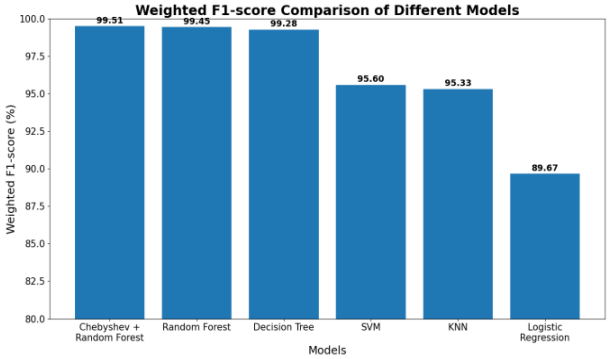
C. Алгоритмуудын гүйцэтгэлийн харьцуулалт

Зураг 6 болон Хүснэгт 2-т UAVIDS-2025 өгөгдлийн сан дээр туршсан машин сургалтын алгоритмуудын гүйцэтгэлийг харьцуулан үзүүлэв. Загваруудын гүйцэтгэлийг Accuracy, Precision, Recall болон F1-score хэмжүүрүүдээр үнэлсэн бөгөөд эдгээр нь ангиллын нарийвчлал, илрүүлэлтийн чадвар болон загварын ерөнхийшүүлэх боломжийг тодорхойлох үндсэн үзүүлэлтүүд юм.

2-Р ХҮСНЭГТ. АЛГОРИТМУУДЫН ХАРЬЦУУЛАЛТ

Алгоритм	**Accuracy**	**Precision**	**Recall**	**F1-score**
Logistic Regression	0.8966	0.8975	0.8966	0.8967
KNN	0.9533	0.9537	0.9533	0.9533
Decision Tree	0.9928	0.9928	0.9928	0.9928
Random Forest	0.9945	0.99945	0.9945	0.9945
SVM	0.9558	0.9577	0.9558	0.9560
Chebyshev + Random Forest	**0.9950**	**0.9951**	**0.9950**	**0.9951**

Туршилтын үр дүнгээс харахад Chebyshev-based feature engineering ашигласан Random Forest загвар хамгийн өндөр гүйцэтгэл үзүүлж, Accuracy = 0.9950, Precision = 0.9951, Recall = 0.9950, F1-score = 0.9951 үзүүлэлтүүдэд хүрсэн байна. Энгийн Random Forest загвартай (Accuracy = 0.9945, F1-score = 0.9945) харьцуулахад Chebyshev-based feature нэмснээр гүйцэтгэл нэмэгдэж байгаа нь статистик хазайлтыг илэрхийлэх нэмэлт шинж чанарууд нь загварын ялгах чадварыг сайжруулж байгааг харуулж байна.



Зураг 7. Алгоритмуудын Weighted F1-score үзүүлэлтийн харьцуулалт

Зураг 7-д алгоритмуудын Weighted F1-score үзүүлэлтийг харьцуулан харуулсан бөгөөд Chebyshev + Random Forest загвар хамгийн өндөр утга үзүүлж байгааг

ажиглаж болно. Энэ нь статистик feature engineering ашигласнаар ангиллын ялгаралт нэмэгдэж байгааг илтгэнэ.

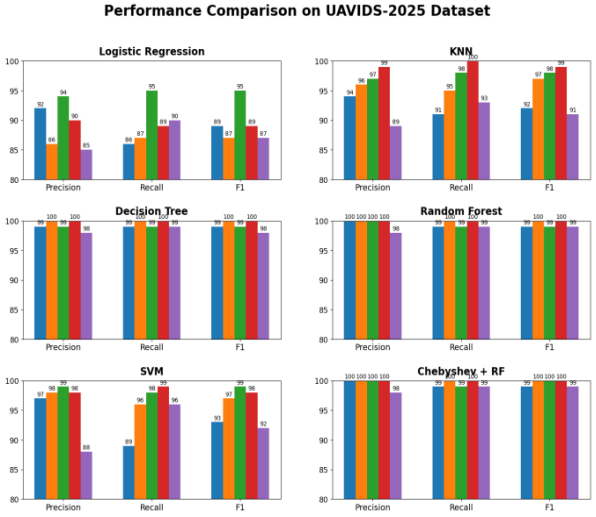
Энгийн Random Forest алгоритм өндөр гүйцэтгэл үзүүлсэн (Accuracy = 0.9945, F1-score = 0.9945) бөгөөд ensemble learning арга нь олон decision tree-ийн нийлбэр дээр суурилан өгөгдлийн шугаман бус, олон хэмжээст хамаарлыг үр дүнтэй суралцаж чаддагтай холбоотой юм. Гэсэн хэдий ч Chebyshev-based feature нэмэгдсэн тохиолдолд гүйцэтгэл цаашид сайжирч (Accuracy = 0.9950, F1-score = 0.9951), статистик feature engineering нь ангиллын үр дүнд эерэг нөлөө үзүүлж байгааг харуулж байна.

Decision Tree алгоритм мөн өндөр гүйцэтгэл үзүүлсэн (Accuracy = 0.9928, F1-score = 0.9928) бөгөөд өгөгдлийн логик хамаарлыг сайн илэрхийлж чадсан. Гэсэн хэдий ч ганц мод ашигладаг тул Random Forest-той харьцуулахад ерөнхийшүүлэх чадвар харьцангуй хязгаарлагдмал байна.

SVM алгоритм дундаас дээш гүйцэтгэл үзүүлсэн (Accuracy = 0.9558), харин KNN алгоритм (Accuracy = 0.9533) нь feature space дахь давхцал ихсэх үед гүйцэтгэл харьцангуй буурах хандлагатай байгааг ажиглаж болно. Logistic Regression алгоритм хамгийн бага гүйцэтгэл үзүүлсэн (Accuracy = 0.8966) нь өгөгдлийн шугаман бус, өндөр хэмжээсийн шинж чанартай холбоотой юм.

Эдгээр үр дүнгээс харахад Chebyshev-based feature engineering нь өгөгдлийн статистик хазайлтыг илүү нарийвчлалтай илэрхийлж, intrusion detection системийн ангиллын гүйцэтгэлийг сайжруулах үр дүнтэй арга болохыг харуулж байна.

Алгоритмуудын гүйцэтгэлийн ялгааг илүү дэлгэрэнгүй харуулах зорилгоор Precision, Recall болон F1-score үзүүлэлтүүдийг ангилал тус бүрээр Зураг 7-т үзүүлэв. Энэхүү зураг нь алгоритм бүрийн гүйцэтгэлийн онцлог, ялгааг илүү тодорхой харуулах боломж олгож байна.



Зураг 8. UAVIDS-2025 өгөгдлийн сан дээрх машин сургалтын алгоритмуудын гүйцэтгэлийн харьцуулалт

Зурагт алгоритм бүрийн Precision, Recall болон Зураг 8-д алгоритм бүрийн Precision, Recall болон F1-score үзүүлэлтүүдийг харьцуулан харуулсан бөгөөд Random Forest болон Chebyshev + Random Forest загварууд тогтвортой өндөр гүйцэтгэл үзүүлж байгааг ажиглаж болно.

Эдгээр үр дүнгээс харахад Chebyshev-based feature engineering нь өгөгдлийн статистик хазайлтыг илүү нарийвчлалтай илэрхийлж, intrusion detection системийн ангиллын гүйцэтгэлийг сайжруулах үр дүнтэй арга болохыг харуулж байна.

Өндөр гүйцэтгэлийн үзүүлэлтүүд нь өгөгдлийн онцлог болон feature engineering-ийн нөлөөг илэрхийлж байгаа боловч цаашид загварын ерөнхийшүүлэх чадварыг бодит орчинд турших шаардлагатай.

V. Дүгнэлт

Энэхүү судалгаанд утасгүй сүлжээний халдлагыг илрүүлэх машин сургалтад суурилсан intrusion detection system (IDS)-ийн загварыг UAVIDS-2025 өгөгдлийн сан дээр боловсруулж, туршив. Судалгаанд Logistic Regression, K-Nearest Neighbor (KNN), Decision Tree, Random Forest болон Support Vector Machine (SVM) алгоритмуудыг ашиглан гүйцэтгэлийг харьцуулан үнэлсэн. Туршилтын үр дүнгээс Random Forest алгоритм хамгийн өндөр суурь гүйцэтгэл үзүүлж байгааг тогтоосон.

Мөн өгөгдлийн статистик хазайлтыг илэрхийлэх зорилгоор Chebyshev-д суурилсан feature engineering аргачлалыг нэвтрүүлсэн. Үүний үр дүнд Chebyshev + Random Forest загвар нь ангиллын нарийвчлалыг нэмэгдүүлж, алдааг бууруулсан нь ажиглагдсан.

Иймээс статистик болон машин сургалтыг хослуулсан hybrid хандлага нь утасгүй сүлжээний халдлага илрүүлэх системийн гүйцэтгэлийг сайжруулах үр дүнтэй арга болохыг энэхүү судалгаа харуулж байна.

Ном зүй

[1] R. Sudha and M. Premkumar, “Securing wireless sensor networks: A survey of challenges and innovations,” Franklin Open, vol. 2025, 2025.

[2] E. Alotaibi, “Wireless Sensor Networks Under Cyber Attacks,” Journal of Cyber Security and Risk Auditing, vol. 2025, no.1, pp. 1–12, 2024.

[3] U. Ghugar and J. Pradhan, “Survey of Wormhole Attack in Wireless Sensor Networks,” Computer Science and Information Technologies, vol. 2, no.1, pp. 33-42, 2021.

[4] B. Devanathan, “A Comparison Study of Various Wireless Intrusion Detection Systems,” International Journal of Scientific and Applied Technology, 2025.

[5] A. M. Arabiat et al., “Intrusion Detection in Wireless Sensor Networks Using Machine Learning Techniques,” Journal of Communications, vol.20, no.4, 2025.

[6] N. R. Saadallah et al., “Enhancing Wireless Security Through Deep Learning: Network Attack Detection and Advanced Classification,” Journal of Modern Technology and Engineering, vol.10, no.1, 2025.

[7] “Intrusion detection system,” Wikipedia, 2024. [Online]. Available: https://en.wikipedia.org/wiki/Intrusion_detection_system. [Accessed: 20-Mar-2026].

[8] H. Mliki, A. Kaceam, and L. Chaari, “A Comprehensive Survey on Intrusion Detection based Machine Learning for IoT Networks,” EAI Endorsed Transactions on Internet of Things, 2021.

[9] T. M. Nguyen et al., “Enhancing Intrusion Detection in Wireless Sensor Networks Using Machine Learning Techniques,” Sensors, vol. 24, 2024.

[10] M. A. Talukder et al., “Machine Learning-based Intrusion Detection for Wireless Sensor Networks using SMOTE-Tomek,” International Journal of Information Security, 2024

[11] V. K. Pandey et al., “Enhancing Intrusion Detection in Wireless Sensor Networks Using Random Forest and Optimization Techniques,” Scientific Reports, 2025.

[12] A. M. Arabiat et al., “Intrusion Detection in Wireless Sensor Networks Using Machine Learning Techniques,” Journal of Communications, vol. 20, no.4, 2025.

[13] A. H. Alshehri et al., “Wormhole Attack Detection and Mitigation Model for IoT and Wireless Sensor Networks Using Machine Learning,” PeerJ Computer Science, 2024.

[14] H. Zhang et al., “SC-MLIDS: Fusion-based Machine Learning Framework for Intrusion Detection in Wireless Sensor Networks,” Ad Hoc Networks, 2025.

[15] G. Al-Sukkar et al., “Machine Learning-based DoS Attack Detection in Wireless Sensor Networks,” Engineering, Technology & Applied Science Research, 2025.

[16] Statista Research Department, “Number of internet users worldwide from 2000 to 2025,” *Statista*, 2025. [Online]. Available: https://www.statista.com/statistics/273018/number-of-internet-users-worldwide/. Accessed: Mar. 16, 2026.

[17] H. Zhang et al., “SC-MLIDS: Fusion-based Machine Learning Framework for Intrusion Detection in Wireless Sensor Networks,” Ad Hoc Networks, vol. 157, 2025.

[18] T. M. Nguyen et al., “Enhancing Intrusion Detection in Wireless Sensor Networks Using Machine Learning Techniques,” Sensors, vol. 24, 2024.

[19] P. Chebyshev, “Des valeurs moyennes,” Journal de Mathématiques Pures et Appliquées, vol. 12, pp. 177–184, 1867.

[20] P. Billingsley, Probability and Measure, 3rd ed. New York, NY, USA: Wiley, 1995.

[21] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” ACM Computing Surveys, vol. 41, no. 3, pp. 1–58, 2009.

[22] D. M. Hawkins, Identification of Outliers. London, U.K.: Chapman and Hall, 1980.

[23] M.Goldstein and S. Uchida, “A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data,” PLoS ONE, vol. 11, no. 4, 2016.

[24] H. Liu, M. Chen, and C. Zhang, “Statistical feature engineering for anomaly detection in network traffic,” IEEE Access, vol. 8, pp. 123456–123468, 2020

[25] N. Moustafa and J. Slay, “UNSW-NB15: A comprehensive data set for network intrusion detection systems,” in Proc. IEEE Military Communications Conf. (MILCOM), 2015, pp. 1–6.

[26] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.

Хиймэл Оюун Ашиглан Автоматжуулсан Лог Хяналт, Шинжилгээний Систем

Цоодолын ЗОЛБООТ
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн оюутан
ШУТИС, Мэдээлэл, холбооны технологийн сургууль
Улаанбаатар, Монгол улс
tszolboot99@gmail.com

Отгонбаярын ЭРХЭМ
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн оюутан
ШУТИС, Мэдээлэл, холбооны технологийн сургууль
Улаанбаатар, Монгол улс
Erkhemot@gmail.com

Эрдэнэсүхийн БИЛГҮҮН
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн оюутан
ШУТИС, Мэдээлэл, холбооны технологийн сургууль
Улаанбаатар, Монгол улс
B221870045@must.edu.mn

Бат-Эрдэнийн МӨНХБАЯР
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн эрхлэгч,
Доктор(Ph.D.), дэд профессор
ШУТИС, Мэдээлэл, холбооны технологийн сургууль
Улаанбаатар, Монгол улс
munkhbayar.b@must.edu.mn

***Хураангуй*—Орчин үеийн цахим дэд бүтэц асар их хэмжээний лог өгөгдөл үүсгэж буй бөгөөд уг өгөгдлийг гараар шинжлэх боломжгүй болжээ. Энэхүү судалгааны ажлын зорилго нь веб серверийн лог өгөгдөл дээр суурилан кибер халдлагыг автоматаар илрүүлэх, мөн илрүүлэлтийн үр дүнг тайлбарлах боломжтой хиймэл оюунд суурилсан систем боловсруулахад оршино. Судалгаанд Isolation Forest алгоритм болон SHAP (SHapley Additive exPlanations) тайлбарлах аргыг хослуулан хэрэглэсэн болно. Туршилтыг синтетик Apache лог өгөгдөл болон UNSW-NB15, CIC-IDS2018 датасет дээр явуулсан бөгөөд систем нь 93.7% accuracy, 84.8% F1-score үзүүлсэн үзүүлэлттэй гарсан. SHAP тайлбарлах механизм нь загварын шийдвэрийг SOC аналистад ойлгомжтой хэлбэрт оруулж, кибер аюулгүй байдлын практикт ашиглах боломжтойг харуулжээ. Уг ажил нь тайлбарлах боломжтой хиймэл оюуныг (XAI) кибер аюулгүй байдлын салбарт нэвтрүүлэх чиглэлд хувь нэмэр оруулсан юм.**

***Түлхүүр үг*—Аномали илрүүлэх, Explainable AI, Isolation Forest, SHAP, Веб лог шинжилгээ**

I. Удиртгал

Сүүлийн жилүүдэд кибер аюулгүй байдлын асуудал дэлхий даяар хурцдаж, халдлагын тоо эрс нэмэгдэж байна. Монгол Улсын хувьд өдөрт дунджаар 40,000–50,000 кибер халдлага бүртгэгдэж байгаагаас 1.9 сая нь фишинг төрлийн халдлага эзэлдэг бөгөөд нийт кибер халдлагын хохирол 2.4 тэрбум төгрөгт хүрчээ [1].

Орчин үеийн системүүдээс маш их хэмжээний лог өгөгдөл үүсдэг бөгөөд эдгээрийг гараар шинжлэх боломжгүй болсон. Сервер, галт хана (firewall), IDS/IPS болон аппликейшн зэрэг олон эх үүсвэрээс үүсэх лог өгөгдлийг бодит цаг хугацаанд шинжлэх нь аюулыг эрт илрүүлэхэд зайлшгүй шаардлагатай [2].

Гэсэн хэдий ч одоо байгаа шийдлүүдийн ихэнх нь “хар хайрцаг” (black-box) шинжтэй байдаг тул SOC (Security Operations Center) аналист тэдгээрийн шийдвэрт итгэж хэрэглэхэд хүндрэлтэй байдаг [3]. Тайлбарлах боломжтой хиймэл оюун (Explainable AI, XAI) нь энэ асуудлыг шийдэх үндсэн хэрэгсэл болж байна [4].

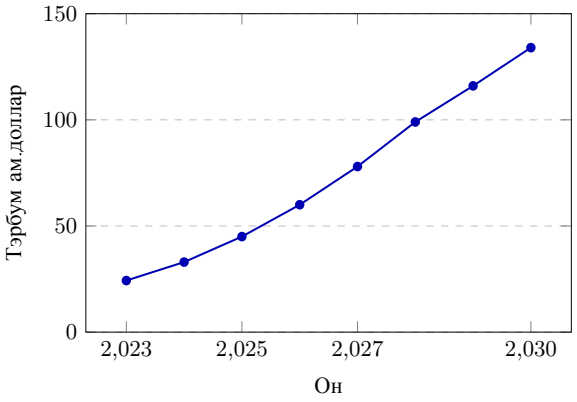
Энэхүү судалгааны ажлын хүрээнд бид дараах зорилтуудыг тавин ажиллав: (1) Unsupervised аргаар веб лог дахь аномалийг илрүүлэх; (2) Isolation Forest ашиглан шинэ халдлагын хэв шинжийг таних; (3) SHAP тайлбарлах аргаар шийдвэрийн үндэслэлийг ил тод болгох; (4) SOC аналистад ойлгомжтой байгалийн хэлний тайлбар үүсгэх; (5) CPU дээр ажиллах хөнгөн архитектур бүхий систем боловсруулах.

II. Сэдэв сонгосон үндэслэл

Кибер аюулгүй байдлын хиймэл оюун ухааны зах зээл 2023 онд 24.3 тэрбум ам.доллараар үнэлэгдсэн бөгөөд 2030 он гэхэд 134 тэрбум ам.долларт хүрэх төлөвтэй байна [5]. Уг өсөлтийн хандлагыг Зураг 1-д харуулав.

Монгол Улсад өдөрт 40,000–50,000 кибер халдлага бүртгэгддэг бөгөөд халдлагын шугамд фишинг хамгийн өндөр тохиолдолтой (1,920,000), дараа нь DDoS (92,759) болон Malware (23,307) орно [1]. Веб серверийн лог нь эдгээр халдлагын үндсэн ул мөрийг агуулдаг бөгөөд уламжлалт дүрэмд суурилсан (rule-based) илрүүлэлт шинэ халдлагыг илрүүлэхэд хангалтгүй тул AI болон Machine Learning ашиглах шаардлага улам бүр нэмэгдсээр байна.

Энэхүү судалгааг ШУТИС-ийн мэдээлэл холбооны технологийн сургуулийн кибер аюулгүй байдлын судалгааны багийн дэмжлэгтэйгээр гүйцэтгэв.



Зураг 1. AI зах зээлийн өсөлтийн хандлага.

III. Өмнө судлагдсан ажил

III-A. DeepLog (Du et al., 2017)

Du нар [3] LSTM ашиглан дараалын загвар бүтээж, системийн лог дотроос аномали илрүүлэхэд өндөр гүйцэтгэл үзүүлсэн. Гэсэн хэдий ч уг загвар нь тайлбарлах боломжгүй "хар хайрцаг" шинжтэй бөгөөд heterogeneous лог датад уялдаа дутмаг байдаг.

III-B. Farzad and Gulliver (2020)

Farzad, Gulliver нар [6] хяналтгүй аномали илрүүлэлт хийж, label шаардлагагүй практик шийдлийг санал болгосон. Autoencoder болон Isolation Forest-ийн хослолоор 86–92% F1-score үзүүлсэн хэдий ч тайлбарлах механизм дутмаг байна.

III-C. Chua (2024)

Chua [7] Isolation Forest ашиглан веб траффикийн аномали илрүүлэхэд ∼92% F1-score үзүүлсэн. Хурдан бөгөөд масштаблах боломжтой ч тайлбарлах чадвар байхгүй.

Өмнөх гол судалгааг арга, гүйцэтгэл, тайлбарлах чадварын хувьд харьцуулсан үр дүнг Хүснэгт I-д үзүүлэв.

Хүснэгт I
Өмнөх судалгаануудтай харьцуулалт

Судлаач	**Арга**	**F1**	**XAI**
Du 2017	LSTM	∼90%	Үгүй
Farzad 2020	AE+IF	86–92%	Хэсэг
Chua 2024	IF	∼92%	Үгүй
Манай	IF+SHAP	84.8%	**Тийм**

Дүгнэж хэлбэл өмнөх судалгаанууд гүйцэтгэлийн хувьд сайн үр дүн үзүүлсэн боловч explainability дутмаг байдаг. Манай судалгаа нь Isolation Forest болон SHAP-ийн хослолоор энэ цоорхойг нөхөхийг зорьсон болно.

IV. Судалгааны асуудал ба зорилго

IV-A. Тулгарч буй асуудлууд

Black-box загвар: AI загварын шийдвэрийг тайлбарлах боломжгүй тул SOC аналист итгэж хэрэглэх хүндрэлтэй. **Гараар шинжлэх боломжгүй:** Өдөрт сая лог мөр үүсдэг тул хүний хяналтаар боловсруулах бололцоогүй. **Labeled data шаардлага:** Хяналттай аргаар сургахад labeled dataset олох, үүсгэх зардалтай. **Zero-day халдлага:** Урьд мэдэгдэж байгаагүй шинэ төрлийн халдлагыг уламжлалт арга илрүүлэхэд бэрхшээлтэй.

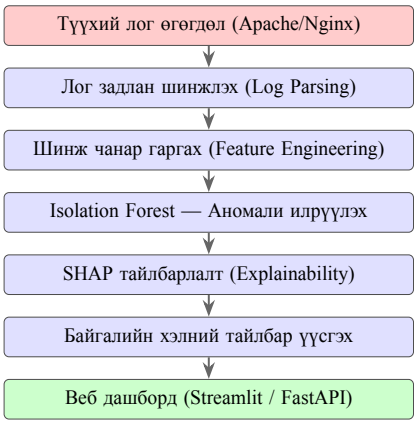
IV-B. Судалгааны таамаглал

H1: Isolation Forest нь веб лог дээр үр дүнтэй аномали илрүүлэх чадвартай.

H2: SHAP ашигласнаар загварын тайлбарлах боломж мэдэгдэхүйц сайжирна.

V. Санал болгох механизм ба архитектур

Системийн бүрэлдэхүүн хэсгүүд болон өгөгдлийн урсгалыг Зураг 2-д дэс дараалан харуулав.



Зураг 2. Санал болгож буй системийн архитектур.

Системийн ажиллагааны үе шатууд: лог мөрүүд орж ирэх; parsing хийж бүтэцтэй өгөгдөл болгох; feature extraction гүйцэтгэх; Isolation Forest загварт дамжуулан аномали оноо тооцох; SHAP утга тооцоолон feature-ийн нөлөөг тодорхойлох; байгалийн хэлний тайлбар автоматаар үүсгэх; эцэст нь веб дашборд дээр үр дүнг харуулах.

VI. Ашигласан өгөгдөл ба шинж чанар гаргах

VI-A. Өгөгдлийн багц

Туршилтын датасет нь Apache/Nginx access log форматтай синтетик өгөгдлөөс бүрдсэн бөгөөд нийт ∼10,000 лог мөр агуулна. Үүнээс 85% хэвийн (normal), 15% аномали (attack) өгөгдөл эзэлнэ. Жишээ лог мөрүүд:

```
192.168.1.10 "GET /index.html" 200
185.220.101.45 "GET /?id=1' OR '1'='1" 500
```

Судалгаанд SQL Injection, XSS (`<script>alert()</script>`), Directory Traversal (`../../etc/passwd`), Command Injection, Brute Force зэрэг халдлагын төрлүүдийг авч үзсэн. Харьцуулалтын зорилгоор UNSW-NB15 [8] болон CIC-IDS2018 [9] датасетийг нэмэлтээр ашиглав.

VI-B. Шинж чанар гаргах

Лог өгөгдлөөс гаргасан 10 шинж чанарыг Хүснэгт II-д жагсаав.

Хүснэгт II
Гаргасан шинж чанарууд

Шинж чанар	**Тайлбар**
query_length	URL-ын нийт урт
has_sql_keyword	'OR','UNION','SELECT' агуулах
has_script_tag	<script> тег
path_depth	URL замын гүн
status_code	HTTP хариу код
request_method	GET/POST/PUT/DELETE
request_hour	Хүсэлт гарсан цаг
special_chars	@, %, - тусгай тэмдэгт
has_traversal	../ агуулах эсэх
payload_size	Хүсэлтийн хэмжээ (bytes)

VII. Ашигласан алгоритм

VII-A. Isolation Forest

Isolation Forest [10] нь хяналтгүй (unsupervised) машин сургалтын алгоритм бөгөөд аномали өгөгдлийг "тусгаарлах" (isolate) зарчмаар ажилладаг. Санамсаргүй шинж чанар болон хуваалтын утга сонгон isolation tree үүсгэж, path length тооцдог. Аномали өгөгдөл хэвийн өгөгдлөөс хэд дахин амархан тусгаарлагдах шинжтэй байдаг.

Аномали оноог (1)-аар тооцно:

$$s(x,\, n) = 2^{-\dfrac{h(x)}{c(n)}} \tag{1}$$

энд $h(x)$ нь x цэгийн path length, $c(n)$ нь n өгөгдлийн дундаж path length юм. $s \approx 1$ бол аномали, $s \approx 0.5$ бол хэвийн гэж үзнэ.

VII-B. SHAP тайлбарлалт

SHAP (SHapley Additive exPlanations) [4] нь тоглоомын онолын Shapley утганд суурилсан feature attribution арга юм. Загварын таамаглалыг задлан, feature бүрийн хувь нэмрийг (2)-аар илэрхийлнэ:

$$f(x) = \phi_0 + \sum_{i=1}^{M} \phi_i \tag{2}$$

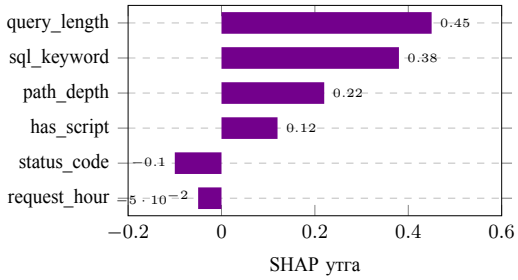
энд ϕ_i нь i-р feature-ийн SHAP утга, ϕ_0 нь суурь утга (baseline) юм.

SHAP утгаар feature бүрийн аномали илрүүлэлтэд оруулах хувь нэмрийг Зураг 3-д харуулав.

VIII. Туршилт ба үр дүн

VIII-A. Туршилтын орчин

Туршилтыг Python 3.10, scikit-learn Isolation Forest, SHAP 0.42, Streamlit/FastAPI орчинд явуулсан бөгөөд тоног төхөөрөмжийн хувьд ердийн CPU ашигласан (GPU шаардлагагүй).



Зураг 3. SHAP feature importance дүрслэл.

Хүснэгт III
Гүйцэтгэлийн үзүүлэлтүүд

Үзүүлэлт	**Утга**
Accuracy	93.7%
Precision	81.5%
Recall	88.3%
F1-score	84.8%
False Positive Rate	5.0%
False Negative Rate	11.7%

VIII-B. Туршилтын үр дүн

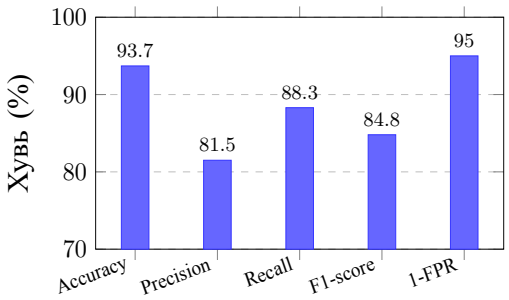
Загварын гүйцэтгэлийн үндсэн үзүүлэлтүүдийг Хүснэгт III-д нэгтгэн харуулав.

2,000 лог мөр бүхий тестийн өгөгдөл дээрх ангилалтын үр дүнг Хүснэгт IV-д харуулав.

Хүснэгт IV
Confusion matrix

	Таам: Хэвийн	**Таам: Аномали**
Бодит: Хэвийн	TN = 1540	FP = 75
Бодит: Аномали	FN = 35	TP = 350

Accuracy, Precision, Recall, F1-score болон 1-FPR үзүүлэлтүүдийн харьцуулсан дүрслэлийг Зураг 4-д үзүүлэв.



Зураг 4. Гүйцэтгэлийн үзүүлэлтүүд.

Систем нь 93.7% accuracy үзүүлсэн нь нийт зөв ангилалтын өндөр хувийг харуулж байна. Recall (88.3%) өндөр байгаа нь халдлагыг илрүүлэх чадвар сайн болохыг баталж байгаа юм. Харин False Positive Rate (5.0%) харьцангуй бага байгаа нь практик ашиглалтад чухал давуу тал болно.

IX. Бусад аргуудтай харьцуулалт

Манай системийг өмнөх судалгаануудтай F1-score, XAI болон GPU шаардлагын хувьд Хүснэгт V-д харьцуулан үзүүлэв.

Хүснэгт V
Аргуудын харьцуулалт

Судалгаа	Арга	F1	XAI	GPU
DeepLog 2017	LSTM	∼90%	Үгүй	Тийм
Farzad 2020	AE+IF	86–92%	Хэсэг	Үгүй
Chua 2024	IF	∼92%	Үгүй	Үгүй
Манай	IF+SHAP	84.8%	**Тийм**	**Үгүй**

Манай систем F1-score-ийн хувьд DeepLog болон Chua-гийнхаас бага байгаа боловч explainability-ийн хувьд тэдгээрээс давуу. Кибер аюулгүй байдлын салбарт загварын шийдвэрийг тайлбарлах чадвар практик ашиглалтад маш чухал бөгөөд IF + SHAP хослол нь detection болон interpretability хоёрын тэнцвэрийг хангасан практик шийдэл болж чаддаг.

X. Давуу тал ба хязгаарлалт

X-A. Давуу тал

Unsupervised: Labeled dataset шаардлагагүй тул хэрэгжүүлэхэд хялбар. **Explainable:** SHAP нь feature-ийн нөлөөг тоогоор гаргаж SOC аналистад тайлбарлах боломж олгоно. **Lightweight:** CPU дээр ажилладаг тул GPU-тай сервер шаардлагагүй. **Real-time:** Isolation Forest-ийн хурдан inference бодит цагийн хяналтад зориулагдсан. **Шинэ халдлага:** Урьд танигдаагүй шинэ халдлагын хэв шинжийг илрүүлэх боломжтой.

X-B. Хязгаарлалт

Feature dependency: Feature engineering-ийн чанар шууд нарийвчлалд нөлөөлнө. **Дарааллын хамаарал дутмаг:** LSTM шиг temporal dependency ойлгохгүй. **Синтетик датасет:** Туршилтын датасет synthetic тул бодит орчны variation бага. **FNR 11.7%:** Нийт аномалийн 11.7%-г алгасах эрсдэлтэй.

XI. Ёс зүйн асуудал

Туршилтын явцад лог өгөгдөл боловсруулагдах тул мэдээллийн нууцлалыг хамгаалах шаардлагатай. AI загварын автоматжуулсан шийдвэр нь хэрэглэгчдийн эрхэд нөлөөлж болзошгүй тул шударга байдлыг хангах нь чухал юм. Систем нь зөвхөн зөвшөөрөлтэй, лабораторийн хяналтан орчинд хэрэглэгдэх ёстой. Black-box асуудлыг SHAP-ийн тусламжтайгаар хэсэгчлэн шийдвэрлэсэн боловч бүрэн тайлбарлалт өгөхгүй байгааг хүлээн зөвшөөрч байна.

XII. Дүгнэлт

Энэхүү судалгаагаар веб лог өгөгдөл дээр суурилсан аномали илрүүлэх системийг боловсруулж, Isolation Forest алгоритм ашиглан хэвийн бус зан төлөвийг илрүүлэв. Мөн SHAP тайлбарлах механизмыг хэрэглэн загварын шийдвэрийг ил тод болгосон юм. Туршилтын үр дүнгээр систем нь 93.7% accuracy, 84.8% F1-score үзүүлж, кибер аюулгүй байдлын практикт ашиглах боломжтойг харуулсан.

Судалгааны гол хувь нэмрүүд: (1) Unsupervised Isolation Forest-ийг SQL injection болон traversal халдлагыг илрүүлэхэд ашигласан; (2) SHAP feature attribution нь SOC аналистад ойлгомжтой шийдвэрийн үндэслэл гаргаж өгсөн; (3) GPU, LLM шаардлагагүйгээр ажиллах хөнгөн архитектур бүтээсэн; (4) Олон мянган лог мөрийг автоматаар шинжлэх боломж олгосон.

Цаашдын судалгаанд LSTM-тэй хослуулах, бодит датасет дээр туршилт явуулах, real-time deployment-ийг сайжруулах чиглэлүүдийг авч үзэхийг санал болгож байна.

Ашигласан материал, ном зүй

[1] Монгол Улсын Цахим Гэрчилгээний Үндэсний Төв, “PUBCERT,” [Цахим]. Холбоос: https://www.pubcert.mn/mn/. [Үзсэн: 2025 оны 3-р сар].

[2] V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” *ACM Comput. Surv.*, vol. 41, no. 3, pp. 1–58, 2009.

[3] M. Du, F. Li, G. Zheng, and V. Srikumar, “DeepLog: Anomaly detection and diagnosis from system logs through deep learning,” in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur. (CCS)*, 2017, pp. 1285–1298.

[4] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, 2017, pp. 4765–4774.

[5] All About AI, “AI statistics in cybersecurity,” 2024. [Online]. Available: https://www.allaboutai.com/

[6] A. Farzad and T. A. Gulliver, “Unsupervised log message anomaly detection,” *ICT Express*, vol. 6, no. 3, pp. 229–237, 2020.

[7] W. Chua, “Web traffic anomaly detection using isolation forest,” *Future Internet*, vol. 16, no. 4, p. 83, 2024.

[8] N. Moustafa and J. Slay, “UNSW-NB15: A comprehensive data set for network intrusion detection systems,” in *Proc. Mil. Commun. Inf. Syst. Conf. (MilCIS)*, 2015, pp. 1–6.

[9] I. Sharafaldin, A. H. Lashkari, and A. A. Ghorbani, “Toward generating a new intrusion detection dataset and intrusion traffic characterization,” in *Proc. 4th Int. Conf. Inf. Syst. Secur. Privacy (ICISSP)*, 2018, pp. 108–116.

[10] F. T. Liu, K. M. Ting, and Z.-H. Zhou, “Isolation forest,” in *Proc. 8th IEEE Int. Conf. Data Mining (ICDM)*, 2008, pp. 413–422.

[11] M. T. Ribeiro, S. Singh, and C. Guestrin, ““Why should I trust you?”: Explaining the predictions of any classifier,” in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining (KDD)*, 2016, pp. 1135–1144.

[12] Z. Guan et al., “LogLLM: Log-based anomaly detection using large language models,” *arXiv:2411.08561*, 2024.

Бага харилцан үйлчлэлтэй Honeypot-д суурилсан халдлага илрүүлэх систем хөгжүүлэх

Баттөрийн ДӨЛГӨӨН
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн оюутан
ШУТИС, Мэдээлэл холбооны технологийн сургууль
Улаанбаатар, Монгол улс
dulguunnbattur@gmail.com

Баянцацралын НОМИНЗУЛ
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн оюутан
ШУТИС, Мэдээлэл холбооны технологийн сургууль
Улаанбаатар, Монгол улс
bayntsatsralnominzul@gmail.com

Батчулууны ДӨЛГӨӨН
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн оюутан
ШУТИС, Мэдээлэл холбооны технологийн сургууль
Улаанбаатар, Монгол улс
enkhamgalandulguun7@gmail.com

Даваажаргалын ИЖИЛЦЭЦЭГ
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн оюутан
ШУТИС, Мэдээлэл холбооны технологийн сургууль
Улаанбаатар, Монгол улс
beomijlee@gmail.com

Бат-Эрдэнийн МӨНХБАЯР
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн эрхлэгч, доктор (Ph.D), дэд профессор
ШУТИС, Мэдээлэл холбооны технологийн сургууль
Улаанбаатар, Монгол улс
munkhbayar.b@must.edu.mn

***Хураангуй* — Кибер халдлагын тоо хэмжээ тасралтгүй нэмэгдэж буй өнөө үед халдлагын урьдчилсан шинж тэмдгийг эрт илрүүлэх нь чухал асуудал болж байна. Энэхүү судалгаанд Dionaea honeypot болон Suricata IDS - ийг хослуулсан идэвхтэй илрүүлэлтийн системийг санал болгосон. Систем нь honeypot, IDS, хяналтын гэсэн гурван модулиас бүрдэж, тусгаарлагдсан виртуал орчинд ажиллана. Туршилтыг CIC-IDS2018 дата багц болон халдлагын симуляци ашиглан гүйцэтгэсэн. Үр дүнд нь хосолсон систем 85.4%-ийн илрүүлэлтийн түвшинд хүрч, худал эерэг түвшинг 1.6% хүртэл бууруулж, тэг-өдөр халдлагыг амжилттай илрүүлсэн. Энэхүү судалгаа нь уламжлалт IDS-ийг сайжруулж, халдлагыг эрт илрүүлэхэд ач холбогдолтой.**

Түлхүүр үг — Кибер аюулгүй байдал, Халдлага илрүүлэх систем, Идэвхтэй илрүүлэлт, Honeypot, Dionaea

I. Удиртгал

Сүүлийн жилүүдэд кибер халдлагын тоо, төрөл, арга техник эрчимтэй нэмэгдэж, байгууллагын сүлжээ болон мэдээллийн системийн аюулгүй байдалд учруулах эрсдэл улам бүр өсөж байна. Ялангуяа тэг өдөр болон үл мэдэгдэх шинэ халдлагууд нэмэгдэж байгаа нь уламжлалт гарын үсэгт суурилсан илрүүлэлт болон галт ханын үр нөлөөг хязгаарлаж байна [4], [5]. Иймээс халдлагыг зөвхөн илрүүлэх бус, түүний урьдчилсан шинж тэмдгийг эрт таних идэвхтэй хамгаалалтын арга зүй шаардлагатай болж байна.

Олон улсын судалгаагаар өдөр бүр 450,000 орчим шинэ хортой код илэрдэг нь кибер аюулын түвшин өндөр байгааг харуулж байна [3], [14]. Honeypot зэрэг өгөөшинд суурилсан технологи халдагчийг тусгаарлагдсан орчинд ажиглаж урьдчилан илрүүлэх боломж олгодог ч өндөр харилцан үйлчлэлтэй системүүд нөөц их зарцуулдаг, өндөр зардалтай сул талтай [6], [17].

Энэхүү судалгаанд бид нөөцийн хэмнэлттэй атлаа халдлагыг эрт үе шатанд илрүүлэх боломжтой Dionaea бага харилцан үйлчлэлтэй honeypot болон Suricata халдлага илрүүлэх системийг хослуулсан архитектурыг санал болгож байна. Тус системийг CIC-IDS2018 дата багц болон халдлагын симуляци ашиглан туршиж, уламжлалт IDS-ээс илүү үр дүнтэй илрүүлэлт болон бага худал эерэг түвшин үзүүлж байгааг баталсан.

II. Ижил төстэй судалгаа

Сүүлийн жилүүдэд honeypot болон халдлага илрүүлэх системийг хослуулсан олон төрлийн судалгааны ажлууд эрчимтэй хийгдэж байна. Эдгээр судалгаанууд нь халдлагыг илрүүлэх нарийвчлал, бодит цагийн хяналт, мөн нөөцийн хэрэглээг оновчлох зэрэг асуудлуудыг шийдвэрлэхэд чиглэж байна.

Тухайлбал, [7]-д үүлэн орчинд өндөр харилцан үйлчлэлтэй (HIH) болон бага харилцан үйлчлэлтэй (LIH) honeypot-уудыг харьцуулан судалж, LIH систем нь техник хангамжийн нөөцийг 90%-иар бага зарцуулдаг болохыг тогтоосон. Гэвч honeypot-уудыг тусдаа виртуал машин дээр байршуулсан нь системийн өргөжих боломжийг хязгаарлах сул талтай байв.

Мориц болон бусад (2024) [13]-ийн судалгаанд honeypot систем нь хиймэл оюун ухаан болон гарын үсгийн санд огтхон ч тулгуурлахгүйгээр шинэ төрлийн халдлагыг амжилттай илрүүлэх чадвартайг онолын хувьд баталсан байна. Түүнчлэн энэ арга нь маш бага худал эерэг дохио өгдөг давуу талтайг онцолсон. Үүнтэй төстэйгээр, [1]-д олон бага харилцан үйлчлэлтэй honeypot-уудыг нэгтгэн ажиллуулснаар халдагчийн үйлдлийг илүү өргөн хүрээнд хамрах боломж бүрдэж, илрүүлэлтийн нарийвчлал мэдэгдэхүйц нэмэгдэж, илрүүлэлтийн зөрүү ихээхэн багассан болохыг туршилтаар харуулсан.

Мөн [9]-д идэвхтэй хамгаалалтын архитектурыг судалж, хиймэл оюун ухаанд суурилсан илрүүлэлтийн системүүдийн өндөр нөөцийн хэрэглээг honeypot технологиор нөхөх боломжтойг үнэлсэн бол [5]-д төвлөрсөн удирдлагатай олон давхаргат honeypot архитектур санал болгож, системийн удирдлагын зардлыг бууруулах боломжийг тодорхойлсон. Эдгээр судалгаанууд нь honeypot технологийн давуу талыг баталсан боловч бодит орчинд хэрэгжүүлэхэд зарим хязгаарлалтууд байсаар байна.

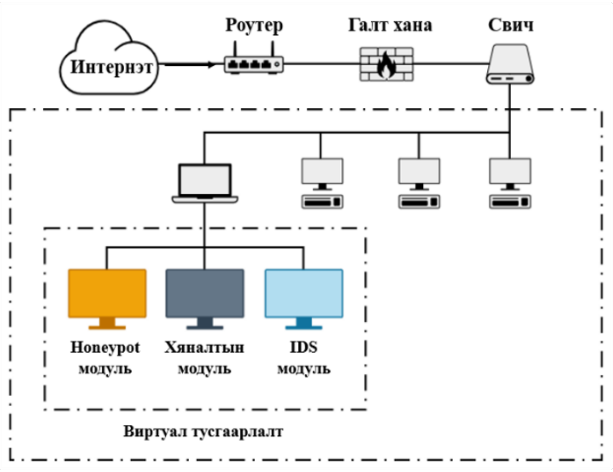
Гэвч дээрх судалгаанууд нь ихэвчлэн өндөр нөөц шаардсан бүтэцтэй, эсвэл олон виртуал орчинд тархсан байдлаар хэрэгжсэн нь системийн уян хатан байдал болон өргөжүүлэх боломжийг хязгаарлаж байна. Мөн бодит цагийн орчинд бага нөөц ашиглан тогтвортой, үр ашигтай илрүүлэлт хийх асуудлыг бүрэн шийдээгүй байна.

Иймээс энэхүү судалгаанд бид Dionaea бага харилцан үйлчлэлтэй honeypot болон Suricata IDS-ийг нэгтгэсэн, нэг виртуал орчинд олон үйлчилгээ зэрэг ажиллуулах боломжтой хосолмол архитектурыг санал болгож байна. Энэхүү шийдэл нь хоёр технологийн гол давуу талуудыг хослуулж, системийн уян хатан байдал, найдвартай байдлыг мэдэгдэхүйц нэмэгдүүлэх зорилготой юм. Тус систем нь хиймэл оюун ухаан ашиглахгүйгээр нөөцийн хэрэглээг бага түвшинд барьж, тэг-өдөр болон шинэ төрлийн халдлагыг илрүүлэх чадварыг мэдэгдэхүйц сайжруулснаараа өмнөх судалгаануудаас ялгаатай юм.

III. Санал болгож буй механизм

Санал болгож буй систем нь тусгаарлагдсан виртуал сүлжээний орчинд байрлуулсан Honeypot модуль, IDS модуль, хяналтын модуль гэсэн гурван үндсэн модулиас бүрдэнэ.

Зураг 1-т уг системийн сүлжээний архитектурыг харуулав. Тус систем нь халдлагыг илрүүлэх, хянах, хариу арга хэмжээ авах үүрэг бүхий харилцан хамааралтай гурван үндсэн модулиас бүрдэнэ. Систем нь бүхэлдээ виртуал тусгаарлагдсан орчинд ажиллах бөгөөд энэ орчин нь халдагчийг бодит сүлжээ болон мэдээлэлд нэвтрэхээс хамгаалах аюулгүйн бүсийг бүрдүүлдэг.

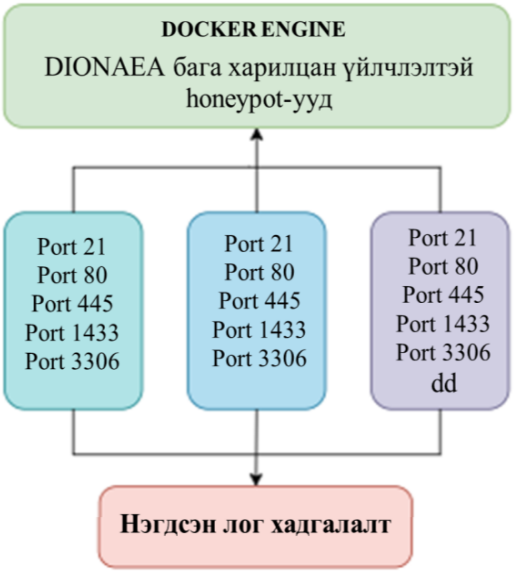


Зураг 1. Системийн сүлжээний архитектур

A. Honeypot модуль

Honeypot модуль нь системийн урд эгнээний хамгаалалт бөгөөд бага харилцан үйлчлэлтэй Dionaea ашиглан хуурамч IP болон FTP, SSH, HTTP, HTTPS, SMB, MySQL зэрэг үйлчилгээг дуурайлган халдагчийг тусгаарлагдсан орчинд татан оруулдаг. Honeypot руу чиглэсэн бүх урсгал анхнаасаа сэжигтэй тул гарын үсэггүйгээр халдлагыг илрүүлэх боломжтой. Ирж буй холболтын IP, порт, нэвтрэх оролдлого, payload зэрэг мэдээллийг SQLite санд бүртгэж, Python-оор боловсруулан статистик болон зан төлвийн шинжилгээ хийдэг. Энэ модуль нь Backdoor, Shellcode зэрэг гарын үсэгт бүртгэгдээгүй тэг-өдөр халдлагыг илрүүлэх чадвараараа давуу талтай.

Зураг 2-т үзүүлсний дагуу Docker Engine дотор Dionaea honeypot нь зэрэгцэн ажиллах гурван жишээгээр дамжуулан хуурамч үйлчилгээ үзүүлж буй портуудыг нэгэн зэрэг хянаж, цугласан бүх лог мэдээллийг нэгдсэн лог хадгалалтын системд автоматаар бүртгэдэг.



Зураг 2. Honeypot модулийн бүтэц

B. IDS модуль

Энэхүү модуль нь системийн шинжилгээний үндсэн давхарга бөгөөд гарын үсэгт суурилсан Suricata IDS болон машин сургалтад суурилсан Isolation Forest алгоритмыг дараалсан байдлаар хослуулан ашигладаг. Эхний шатанд Suricata нь сүлжээний урсгалыг бодит цаг хугацаанд тасралтгүй хянаж, өөрийн гарын үсгийн сантай харьцуулан мэдэгдэж буй халдлагыг хурдан илрүүлэх үүрэгтэй. Хэрэв тухайн урсгалд ямар нэгэн тохирол илрээгүй тохиолдолд, хоёр дахь шатанд Isolation Forest алгоритм автоматаар идэвхжиж, хэвийн бус зан төлөвийг илрүүлэх замаар илүү нарийвчилсан шинжилгээ хийж, тэг-өдөр болон шинэ төрлийн халдлагыг илрүүлэх боломжийг олгодог.

Ийнхүү хоёр шатлалт энэхүү зохион байгуулалт нь системийн илрүүлэлтийн үр ашиг болон уян хатан байдлыг нэмэгдүүлэхэд чиглэгдсэн. Энэ нь системийн илрүүлэлтийн чадварыг мэдэгдэхүйц сайжруулдаг.

1) Гарын үсэгт суурилсан илрүүлэлт

Suricata нь honeypot орчин руу чиглэсэн бүх сүлжээний урсгалыг бодит цаг хугацаанд хянаж, гарын үсгийн сантай тулган шалгана. Энэ модуль нь зөвхөн honeypot трафикийг шинжилдэг тул хэвийн хэрэглэгчийн урсгал оролцохгүй, ингэснээр хуурамч дохио эрс багасдаг. Гарын үсэг таарсан тохиолдолд тухайн урсгалыг “Known Alert” болгон бүртгэж, дараагийн модульд дамжуулна.

2) Машин сургалтад суурилсан илрүүлэлт

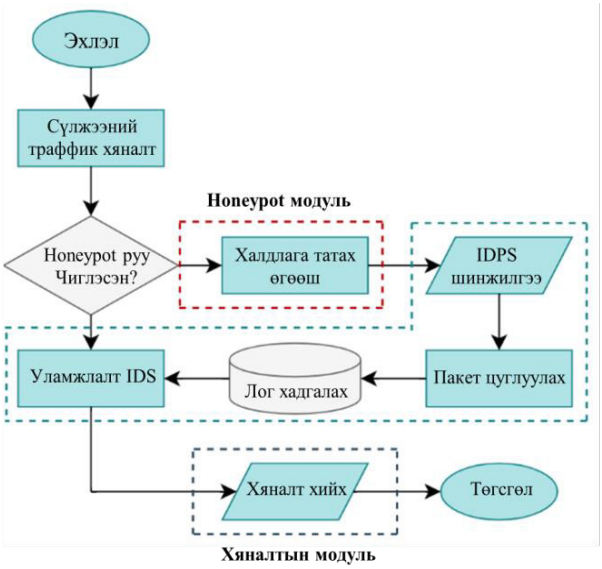
Шинэ болон гарын үсэгт бүртгэгдээгүй халдлагыг илрүүлэхийн тулд Isolation Forest алгоритмыг хоёр дахь шатанд ашигласан. Энэ нь зөвхөн хэвийн урсгалын өгөгдлөөр сургалт хийдэг тул тэг-өдөр халдлагад тохиромжтой. Honeypot трафикаас CICFlowMeter ашиглан 23 онцлогийг гаргаж авч, моделийг CIC-IDS2018 дата багцын 50,000 хэвийн урсгал дээр сурган, contamination = 0.05 гэж тохируулсан бөгөөд гажуудал илэрвэл “Тэг-Өдөр” болгон ангилна.

C. Хяналтын модуль

Хяналтын модуль нь honeypot болон халдлага илрүүлэх системээс ирсэн өгөгдлийг нэгтгэн боловсруулж, аюулгүй байдлын мэдээллийг бодит цаг хугацаанд харуулна. Dionaea-ийн SQLite сан болон Suricata-ийн eve.json файлаас өгөгдөл авч, нэг эх IP хаягтай холбоотой бүх үйлдлүүдийг цагийн дарааллаар нэгтгэн нарийвчилсан шинжилгээ хийдэг. Үүний үр дүнд халдлагын тархалт, цаг хугацааны өөрчлөлт, протоколын статистик болон идэвхтэй халдагчдыг нэгдсэн самбарт тодорхой, ойлгомжтойгоор харуулна.

D. Системийн ажиллах алгоритм

Тусгаарлагдсан орчинд халдлагыг илрүүлж, тэдгээрт тохирох хариу арга хэмжээг хэрэгжүүлэх үйл явц нь урьдчилан тодорхойлогдсон дарааллын дагуу бүрэн автоматжсан системтэйгээр явагдах бөгөөд уг процесс нь зураг 3-т үзүүлсэн алгоритмийн дагуу шат дараатай, тодорхойлогдсон алхмуудаар явагдана.



Зураг 3. Системийн архитектур

Сүлжээний урсгалыг роутер/свитч түвшинд шүүж honeypot руу чиглэсэн трафикийг тусгаарлагдсан виртуал орчинд оруулснаар халдагчийн үйлдэл, ашигласан хэрэгсэл болон payload мэдээллийг саадгүй цуглуулах боломж бүрдэнэ. Цугласан өгөгдөлд IDPS шинжилгээ хийж лог файлд хадгалснаар дараагийн дүн шинжилгээнд ашиглагдана. Үүнд тулгуурлан хяналтын модуль нь IP блоклох, порт хаах, IDS/IPS дүрэм шинэчлэх зэрэг хариу арга хэмжээг автоматаар хэрэгжүүлж, системийг идэвхтэй хамгаалалтын хэлбэрээр ажиллуулна.

IV. Туршилтын арга зүй

A. Туршилтын орчин

Туршилтын орчныг гадны сүлжээний нөлөөллөөс ангид байлгах үүднээс "host-only" горим бүхий тусгаарлагдсан виртуал сүлжээнд гурван виртуал машин ашиглан бүрдүүлэв. Халдлага илрүүлэх систем бүхий сервер нь Ubuntu 24.04 LTS үйлдлийн систем дээр суурилагдсан бөгөөд үүнд гарын үсэг дээр суурилсан Suricata системийг суурилуулсан болно. Honeypot бүхий сервер дээр Docker контейнер дотор байршуулсан Dionaea бага харилцан үйлчлэлтэй Honeypot-ийг хослуулан ажиллуулсан болно. Харин халдлага гүйцэтгэгч талын машиныг нэвтрэх шалгалтын стандарт хэрэгслүүдээр тоноглогдсон, аюулгүй байдлын судалгаанд зориулагдсан Kali Linux 2024 систем дээр тохируулав. Туршилтын орчны иж бүрэн Хүснэгт 1-т нэгтгэн харуулав.

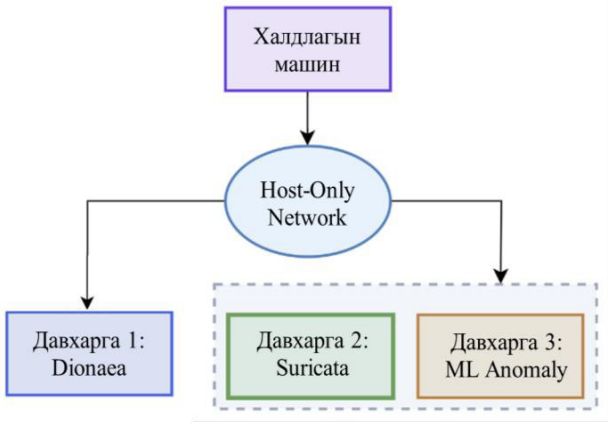
ХҮСНЭГТ 1. ТУРШИЛТЫН ОРЧИН

Бүрэлдэхүүн	**Тодорхойлолт**
IDS машин	Ubuntu server 24.04 LTS
Honeypot машин	Ubuntu server 24.04 LTS
Хяналтын машин	Ubuntu 24.04 LTS
Халдлагын машин	Kali Linux
Уламжлалт IDS	Suricata 8.0.3
Honeypot	Dionaea (Docker контейнер)
Машин сургалтын илрүүлэлт	Isolation Forest (sckit-learn)
Хортой код шалгагч	ClamAV

Зураг 4-т хосолсон халдлага илрүүлэх системийн гурван давхаргат архитектурыг дүрслэн харуулав. Халдлагын машин нь Host-Only сүлжээгээр дамжин системд холбогдож, орж ирсэн урсгал нь дараалсан гурван шатанд шинжлэгдэнэ.

Эхний давхаргад Dionaea honeypot нь ирж буй бүх холболтыг бүртгэн авч, халдагчийн үйл ажиллагаа, ашигласан арга техник, payload-ийн дэлгэрэнгүй мэдээллийг цуглуулж, хадгална. Хоёр дахь давхаргад Suricata нь сүлжээний урсгалыг бодит цаг хугацаанд тасралтгүй хянаж, гарын үсгийн сантай тулган мэдэгдэж буй халдлагыг өндөр нарийвчлалтайгаар илрүүлнэ. Харин гурав дахь давхаргад Isolation Forest алгоритм нь Suricata-аар илрээгүй хэвийн бус зан төлөвийг статистик аргаар нарийвчилсан шинжилгээ хийж, тэг-өдөр болон шинэ төрлийн халдлагын боломжит тохиолдлыг тодорхойлдог. Иймхүү гурван давхаргын уялдаа холбоо

нь мэдэгдэж буй болон шинэ төрлийн халдлагыг хамтад нь үр дүнтэй илрүүлэх боломжийг бүрдүүлдэг.



Зураг 4. Хосолсон IDS системийн архитектур

B. Дата багц болон халдлагын симуляци

Энэхүү судалгаанд сүлжээний халдлага илрүүлэх судалгааны салбарт өргөнөөр ашиглагддаг Канадын Кибер аюулгүй байдлын хүрээлэнгийн "CIC IDS2018" жишиг дата багцийг ашигласан. Тус дата багцийг B-Profile систем болон CICFlowMeter-ийг ашиглан хяналттай сүлжээний орчинд бүрдүүлсэн бөгөөд таван өдрийн хугацаанд хэвийн сүлжээний урсгал болон долоон төрлийн халдлагын урсгалыг бүртгэн авсан байна. Хүснэгт 2-т энэхүү үнэлгээнд ашигласан CIC-IDS2018 дата багцийн статистик задаргааг, тэр дундаа халдлагын ангилал бүрт харгалзах сүлжээний урсгалын тоог нэгтгэн харуулав.

ХҮСНЭГТ 2. АШИГЛАГДСАН ХАЛДЛАГЫН АНГИЛАЛ

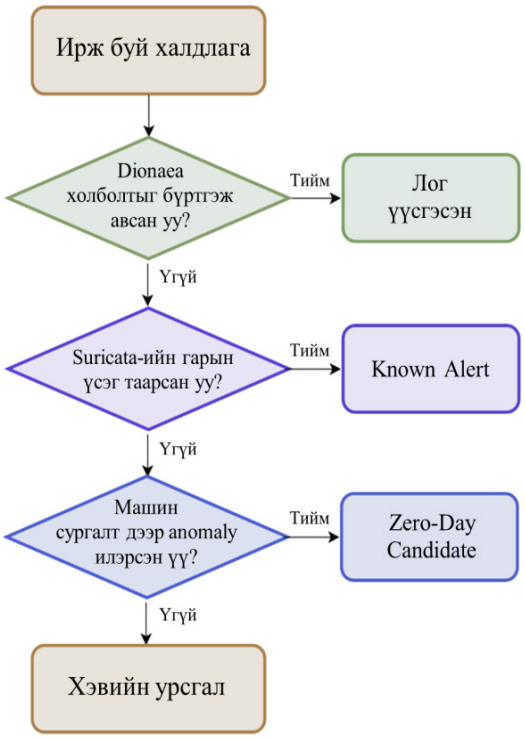
Ангилал	Нийт урсгал	Халдлагын урсгал	Халдлагын хувь
Brute Force	221,020	13,835	6.3%
Dos / DDoS	1,048,575	693,329	66.1%
Web Attacks	170,366	2,180	1.3%
Infiltration	161,934	93	0.1%
Bot	286,191	286,191	100%
Port Scan	286,467	158,930	55.5%
Нийт	2,400,298	1,282,583	53.4%

C. Сүлжээний урсгалыг загварчлах аргачлал

Туршилтын орчны бодит байдлыг хангах үүднээс CICIDS2018 дата багцийн PCAP файлуудыг tcpreplay хэрэгслийг ашиглан туршилтын сүлжээнд дахин дамжуулсан болно. Энэхүү процесс нь Honeypot машин руу чиглэсэн бодит халдлагын нөхцөл байдлыг лабораторийн орчинд аль болох нарийвчлалтайгаар дуурайлган үүсгэх боломжийг олгосон.

Туршилтын үр дүнгийн үнэн зөв байдлыг хангах, өөр өөр тохиргоо болон халдлагын төрлүүд хоорондоо холилдохоос сэргийлэх зорилгоор жишиг үнэлгээний гүйцэтгэл бүрийн өмнө илрүүлэлтийн бүх лог файлуудыг системээс бүрэн цэвэрлэж, шинээр эхлүүлж байв.

Пакет бүрийг хүлээн авах, шинжлэх болон ангилах үе шатуудыг Зураг 5 дээр системчилсэн байдлаар дүрслэн харуулсан бөгөөд энэ нь пакет бүр үнэлгээний явцад дамжин өнгөрөх илрүүлэлтийн логик урсгалыг тодорхойлж байгаа юм.



Зураг 5. Сүлжээний урсгалын илрүүлэлт, ангиллын урсгал

D. Үнэлгээний үзүүлэлтүүд

IDS-ийн гурван өөр тохиргооны илрүүлэлтийн гүйцэтгэлийг бодитой харьцуулахын тулд туршилт бүрт зургаан тоон үзүүлэлтийг тодорхойлж, тооцоолсон болно. Бүх үзүүлэлтүүдийг энэхүү судалгаанд зориулан хөгжүүлсэн жишиг үнэлгээний скриптээр автоматаар тооцоолж, дахин ашиглах боломжтой JSON форматаар хадгалсан.

Энд TP - үнэн эерэг, FP - худал эерэг, FN - худал сөрөг, TN - үнэн сөрөг, DR – илрүүлэлтийн түвшин утгуудыг илэрхийлэх ба үүнд:

- **Үнэн эерэг (TP):** CIC-IDS2018 өгөгдлийн сан дахь халдлагын урсгалыг IDS систем зөв илрүүлж, анхааруулга өгсөн тохиолдол.
- **Худал эерэг (FP):** Дата багц дахь хэвийн урсгалыг IDS систем халдлага хэмээн алдаатай тодорхойлсон тохиолдол.
- **Худал сөрөг (FN):** Дата багц дахь халдлагын урсгалыг IDS систем илрүүлж чадаагүй тохиолдол.
- **Үнэн сөрөг (TN):** Хэвийн урсгалыг аюулгүй хэмээн зөв ангилсан тохиолдол.
- **Илрүүлэлтийн түвшин (DR):** Систем нь дата багц дахь халдлагын урсгалыг илрүүлэх чадварыг илтгэнэ.

3) Илрүүлэлтийн түвшин:

Илрүүлэлтийн түвшин нь систем нийт халдлагын хэдэн хувийг амжилттай илрүүлснийг хэмжинэ:

$$DR = \frac{TP}{TP + FN} \times 100\% \quad (1)$$

Илрүүлэлтийн түвшин өндөр байх нь халдлагыг хамрах хүрээ өргөн байгааг илтгэх бөгөөд энэ нь IDS-ийн тохиргоонуудыг харьцуулах үндсэн үзүүлэлт юм.

4) Б. Худал эерэг түвшин

Худал эерэг түвшин нь хэвийн урсгалыг халдлага гэж алдаатай мэдээлсэн хувийг хэмждэг:

$$FPR = \frac{FP}{FP + TN} \times 100\% \quad (2)$$

Худал эерэг түвшин бага байх нь үндэслэлгүй анхааруулга цөөн гарч, аюулгүй байдлын шинжээчдийн ачааллыг бууруулан, үйл ажиллагааны найдвартай байдлыг дээшлүүлдэг.

5) В. Худал сөрөг түвшин

Худал сөрөг түвшин нь системээс орхигдуулсан бодит халдлагын хувийг хэмждэг:

$$FNR = \frac{FN}{FN + TP} \times 100\% \quad (3)$$

Илрүүлээгүй орхисон халдлага нь системд ихээхэн хохирол учруулж болох тул энэ үзүүлэлт маш чухал бөгөөд үүнийг аль болох бага байлгахыг эрмэлздэг.

6) Г. Илрүүлэлтийн зөрүү

Илрүүлэлтийн зөрүү нь нэг халдлагын эх үүсвэрийн хувьд Suricata-ийн анхны анхааруулга болон Dionaea-ийн анхны холболтын лог хоорондын цаг хугацааны зөрүүг хэмждэг. Үүнийг зөвхөн хоёр систем зэрэг ажиллаж буй "хосолсон" горимд хэмжсэн:

$$Detection\ Gap = | T_{suricata} - T_{dionaea} | \times 1000мс \quad (4)$$

Илрүүлээгүй орхисон халдлага нь системд ихээхэн хохирол учруулж болох тул энэ үзүүлэлт маш чухал бөгөөд үүнийг аль болох бага байлгахыг эрмэлздэг.

Хүснэгт 3-т IDS системийн гүйцэтгэлийг үнэлэхэд ашигласан үндсэн үзүүлэлтүүдийг нэгтгэн харуулав. Үүнд илрүүлэлтийн түвшин, худал эерэг болон худал сөрөг түвшин, илрүүлэлтийн зөрүү, мөн нөөцийн ашиглалт зэрэг хэмжигдэхүүнүүд багтсан бөгөөд эдгээр нь системийн нарийвчлал, найдвартай байдал, гүйцэтгэлийн үр ашгийг цогцоор нь үнэлэх боломжийг бүрдүүлдэг. Тус хүснэгтэд үзүүлсэн оновчтой утгууд нь системийн ажиллагааг сайжруулах зорилтот түвшинг илэрхийлж, халдлагыг өндөр нарийвчлалтай илрүүлэхийн зэрэгцээ худал дохиог бууруулах шаардлагыг тодорхойлж байна.

ХҮСНЭГТ 3. ҮНЭЛГЭЭНИЙ ҮЗҮҮЛЭЛТИЙН ХУРААНГУЙ

Үзүүлэлт	**Оновчтой утга**	**Ач холбогдол**
Илрүүлэлтийн түвшин (DR)	90-100%	Халдлагыг хамрах хүрээ
Худал эерэг түвшин (FPR)	0%	Үндэслэлгүй анхааруулгыг бууруулах
Худал сөрөг түвшин (FNR)	0%	Орхигдуулсан халдлагыг багасгах
Илрүүлэлтийн зөрүү (DG)	Хамгийн бага	Цаг хугацааны хамаарлын боломж
Нөөцийн ашиглалт	Хамгийн бага	Практикт нэвтрүүлэх боломж

Е. Тэг-өдөр халдлагыг илрүүлэх аргачлал

Тэг-өдөр халдлагыг үнэлэх зорилгоор CIC-IDS2018 дата багцад агуулагдаагүй, Suricata-ийн гарын үсгийн санд бүртгэлгүй 70 ширхэг симуляцлагдсан payload ашигласан.

1) Тэг-өдөр халдлагын өгөгдөл

Payload-уудыг Metasploit exploit, msfvenom-оор үүсгэсэн полиморф payload, мөн Python/Scapy ашиглан үүсгэсэн протоколын зөрчилтэй пакет гэсэн гурван аргаар бэлтгэсэн. Бүх payload-уудыг Kali Linux машинаас Dionaea honeypot-ийн нээлттэй портууд-руу host-only сүлжээгээр дамжуулан илгээсэн. Хяналттай орчинд үүсгэсэн payload, Dionaea-ийн лог бүртгэл, мөн VirusTotal шалгалт зэрэг дээр тулгуурлан үнэхээр тэг-өдөр халдлага мөн эсэхийг баталгаажуулсан.

Тэг-өдөр туршилтын өгөгдлийн олон талт байдлыг хангах үүднээс payload-уудыг өөр өөр аргаар үүсгэсэн бөгөөд тэдгээрийн задаргааг хүснэгт 4-т нэгтгэн харуулав.

ХҮСНЭГТ 4. ТЭГ-ӨДӨР ХАЛДЛАГЫГ ТОДОРХОЙЛОХ ХҮЧИН ЗҮЙЛС

Аргачлал	**Тоо**	**Халдлагын төрөл**	**Хэрэгсэл**
Metasploit exploit модуль	25	Алсын код гүйцэтгэл, Service exploit	Metasploit
Полиморф payload	25	Backdoor, Shellcode	Msfvenom, Netcat
Үүсгэсэн скрипт	20	Гажуудсан пакет, Протоколын зөрчил	Python, Scapy
Нийт	70		

2) Тэг-өдөр халдлагыг илрүүлэх онооны загвар

Гарын үсэгт суурилсан хэрэгслүүд үл мэдэгдэх шинэ халдлагыг илрүүлэх боломжгүй тул "Тэг-өдөр" сэжигтэй үйлдлийг ялгах зорилгоор олон түвшинт зан төлөвийн шинжилгээний дамжлагыг боловсруулсан. Энэхүү аргачлалаар халдлагын эх үүсвэр бүрт зан төлөв, үйл ажиллагааны онцлог, ашигласан порт, payload-ийн шинж чанар зэрэг олон талын нотлох хүчин зүйлсэд тулгуурлан итгэлцлийн оноог оноож, тооцон үнэлгээ хийдэг.

ХҮСНЭГТ 5. ТЭГ-ӨДӨР ХАЛДЛАГЫГ ТОДОРХОЙЛОХ ХҮЧИН ЗҮЙЛС

Хүчин зүйлс	Оноо	Үндэслэл
Suricata гарын үсэгтэй тараагүй	+3	Санамжийн өгөгдлийн санд бүртгэгдээгүй
ML Isolation Forest Anomaly	+3	Статистикийн хувьд хэвийн бус зан төлөв
Протоколын стандарт зөрчил	+2	Хүлээгдэж буй протокол ашиглалтын загвараас өөр
Үл мэдэгдэх протокол	+2	Аль ч мэдэгдэж буй клиент ийм өгөгдөл илгээдэггүй
Эксплойт хэлбэрийн	+2	Шууд эксплойт хийх оролдлого
Удаан үргэлжилсэн	+1	Урт хугацааны, зорилтот зан төлөв
Өндөр энтропи бүхий binary	+3	Пакетлагдсан эсвэл далдлагдсан өгөгдөл
VirusTotal-д үл мэдэгдэх бинар	+3	70 гаруй антивирусны системээр илрээгүй
Мэдэгдэж буй сканнер	-4	Сканнераас үүссэн худал эерэг түвшинг бууруулах
Масс Brute-Force (>50 нэвтрэлт)	-2	Автоматжуулсан шуугиан үүсгэгч байх магадлалтай

Бие даасан аргуудаар 6-аас дээш оноо авсан IP хаягийг “Тэг-өдөр” халдлагад тооцно. Энэхүү үнэлгээ нь олон хүчин зүйлд суурилсан бөгөөд тухайн IP хаяг нь дор хаяж хоёр өөр илрүүлэлтийн механизм дээр зэрэг баталгаажсан байх шаардлагатай. Иймээс хосолмол шалгуур нь худал эерэг илрүүлэлтийг бууруулахад чухал үүрэгтэй.

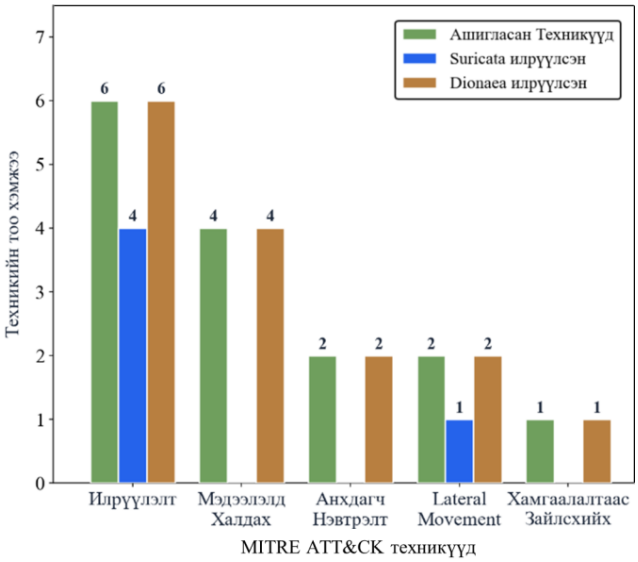
Мөн Honeypot болон Docker орчны дотоод IP хаягуудыг тооцооллоос хасах замаар дотоод холболтоос үүдэлтэй алдаа багасгаж, гадаад халдлагын урсгалыг илүү бодитой ялган таних боломжийг бүрдүүлсэн.

F. MITRE ATT&CK хамрах хүрээ

Энэхүү судалгаанд боловсруулсан хосолсон систем нь халдлагын амьдралын мөчлөгийн янз бүрийн үе шатуудыг илрүүлэх чадвартай бөгөөд үүнийг MITRE ATT&CK матрицаар дараах байдлаар ангилав:

- **Тагнуул, тандалт:** Suricata-ийн хяналтын дүрмүүд болон nmap сканнерын идэвхжүүлсэн дохиоллууд.
- **Анхдагч нэвтрэлт:** Dionaea honeypot-д бүртгэгдсэн үйлчилгээний мөлжлөг болон brute-force оролдлогууд.
- **Гүйцэтгэл:** Бүртгэгдсэн хортой бинар файлууд болон тэдгээрийн VirusTotal дээрх шинжилгээ.
- **Илрүүлэлт, хайгуул:** Сүлжээний дотоод портуудыг шалгах болон үйлчилгээний мэдээлэл цуглуулах оролдлогыг Dionaea-ийн тусламжтайгаар бүртгэх.

Зураг 6 дээр тактик бүрт хамаарах техникүүдийн тархалт болон систем бүрийн илрүүлэлтийн хамрах хүрээг харуулав. Энэхүү зураг нь MITRE ATT&CK тактикийн бүлэг тус бүрт ашиглагдсан техникүүд болон системүүдийн илрүүлэлтийн хамрах хүрээг тодорхой харуулж байна.



Зураг 6. Suricata болон Dionaea системүүдийн харьцуулалт

V. ТУРШИЛТЫН ҮР ДҮН

Энэхүү хэсэгт гурван төрлийн тохиргоо болох уламжлалт IDS, өндөр харилцан үйлчлэлтэй honeypot-д суурилсан IDS болон LIHP-д суурилсан IDS-г CIC-IDS2018 өгөгдлийн сан дээр хийсэн туршилтын үр дүнг танилцуулна.

Үр дүнг нийт таван гол хэмжүүрээр илэрхийлсэн бөгөөд үүнд илрүүлэлтийн хувь, хуурамч дохионы тоо, тэг-өдөр халдлага илрүүлэлтийн чадвар, нөөцийн хэрэглээ зэрэг үзүүлэлтүүд багтана. Эдгээр хэмжүүрүүд нь системийн илрүүлэлтийн үр ашиг, найдвартай байдал болон практик хэрэглээний талаас нь цогцоор үнэлэх боломжийг олгодог.

A. Илрүүлэлтийн үр дүнгийн тойм

Хүснэгт 5 дээр CIC-IDS2018 өгөгдлийн сан дахь 1,282,583 халдлагын урсгалд үндэслэн гурван төрлийн IDS тохиргооны илрүүлэлтийн үр дүнгийн тоймыг харуулсан.

ХҮСНЭГТ 6. ИЛРҮҮЛЭЛТИЙН ҮР ДҮНГИЙН ТОЙМ

Үзүүлэлт	Уламжлалт IDS	HIHP IDS	Хосолсон IDS
Халдлагын урсгал	1,282,583	1,282,583	**1,282,583**
Үнэн эерэг (TP)	892,451	1,198,734	**1,095,325**
Худал эерэг (FP)	48,721	31,456	**18,234**
Худал сөрөг (FN)	390,132	83,849	**187,258**
Илрүүлэлтийн түвшин (DR)	69.6%	93.5%	**85.4%**
Худал сөрөг түвшин (FNR)	30.4%	6.5%	**14.6%**
Худал эерэг түвшин (FPR)	4.4%	2.8%	**1.6%**
Нарийвчлал	94.8%	97.4%	**98.5%**

1) Тэг-өдөр халдлагын өгөгдөл

Илрүүлэлтийн тоймоос харахад хосолсон халдлага илрүүлэх систем нь 1,095,325 'үнэн эерэг' үр дүнтэйгээр 85.4%-ийн илрүүлэлтийн түвшинд хүрч, уламжлалт системээс 15.8%-иар илүү гүйцэтгэл үзүүлсэн. Хэдийгээр бага харилцан үйлчлэлтэй honeypot-д суурилсан хандлага нь илрүүлэлтийг сайжруулж байгаа ч өндөр харилцан үйлчлэлтэй honeypot системтэй харьцуулахад бага зэрэг сул байна. Учир нь өндөр харилцан үйлчлэлтэй honeypot нь бодит орчныг гүнзгий дуурайснаар халдагчийн үйлдлийг удаан хугацаанд ажиглах, илүү их өгөгдөл цуглуулах боломжийг олгодог нь худал сөрөг үзүүлэлтийг бууруулдаг байна.

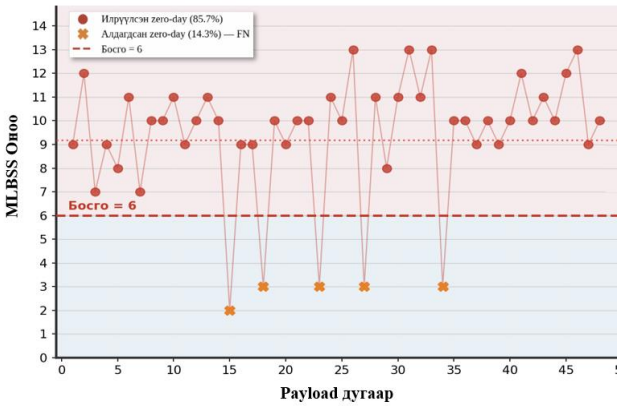
2) Нөөцийн үр ашиг болон нарийвчлал

Хэдийгээр хосолсон системийн илрүүлэлтийн түвшин өндөр харилцан үйлчлэлтэй системээс 8.1%-иар бага байгаа ч CPU-ийн зарцуулалтыг 46.1%-аас 2.8% хүртэл, RAM-ийн хэрэглээг 2,214 MB-аас 456 MB хүртэл (16 дахин болон 5 дахин) бууруулсан нь хязгаарлагдмал техник хангамжтай бодит орчинд илрүүлэлтийн өндөр нарийвчлалыг хадгалан үлдээж, практикт өргөн хүрээнд хэрэгжүүлэх боломжтой маш үр дүнтэй, практик шийдэл болохыг тодорхой нотолж байна. Энэхүү нөөцийн хэмнэлт нь ялангуяа жижиг болон дунд хэмжээний байгууллагуудад ихээхэн ач холбогдолтой юм. Мөн хосолсон IDS-ийн худал эерэг түвшин ердөө 1.6% байгаа нь системийн шуугиан маш бага, нарийвчлал өндөр байгааг баталлаа.

B. Тэг өдөр халдлагыг илрүүлэлтиний үр дүн

Шинэ төрлийн халдлагыг илрүүлэх чадвар нь энэхүү хосолсон IDS системийн гол давуу талын нэг юм. Үүнийг үнэлэхийн тулд MLBSS (Multi-Level Behavior Scoring System) онооны аргыг ашиглан симуляцлагдсан 70 payload өгөгдөл дээр туршилт хийсэн бөгөөд илрүүлэлтийн босго утгыг 6 оноогоор тогтоосон.

Зураг 7-д харуулснаар систем нь тэг-өдөр халдлагын 85.7%-ийг (TPR) амжилттай илрүүлсэн бол худал сөрөг (FN) үзүүлэлт 14.3% байсан. Энэ нь гарын үсэгт суурилсан уламжлалт аргуудаас ялгаатайгаар зан төлөвт суурилсан үнэлгээний механизм ашигласнаар өмнө нь танигдаагүй халдлагыг илрүүлэх боломжийг бүрдүүлж байгааг харуулав.



Зураг 7. Тэг өдрийн халдлагын илрүүлэлтийн онооны харьцуулалт

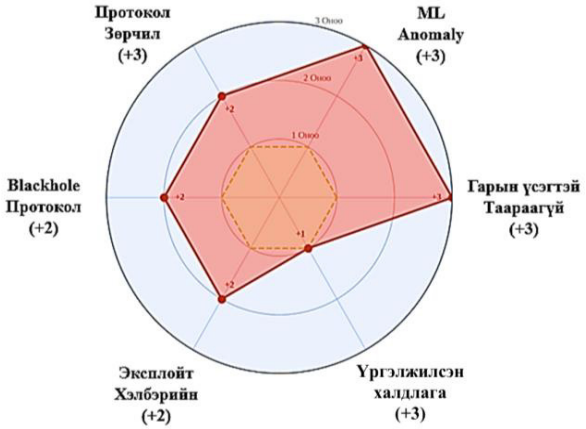
7-р хүснэгтэд үзүүлснээр тэг-өдөр нэр дэвшигч нь зан төлөвт суурилсан хүчин зүйлсийн үнэлгээгээр 13 оноо авч, ангиллын босго болох 6 онооноос хамаагүй давсан байна. Энэ нь тухайн урсгал олон шалгууруар баталгаажиж, халдлагын шинж тодорхой илэрч байгааг харуулна.

ХҮСНЭГТ 7. ТЭГ-ӨДӨР ХАЛДЛАГА ЭСЭХИЙГ ШАЛГАХ ШАЛГУУР

Хүчин зүйл	Илэрсэн	Оноо
Suricata signature-д таарахгүй байх	Тийм	+3
ML Isolation Forest гажиг илрүүлэлт	Тийм	+3
Протоколын техникийн нөхцөл зөрчигдсөн	Тийм	+2
Blackhole (үл мэдэгдэх протокол)	Тийм	+2
Нэвтрэх оролдлогогүй (exploit төрлийн)	Тийм	+2
Үргэлжилсэн халдлага (8,226 секунд)	Тийм	+1
Өндөр энтропи бүхий binary файл	Үгүй	+0
VirusTotal-д бүртгэгдээгүй	Үгүй	+0
Нийт оноо	**13**	
Ангилал	***Тэг-өдөр байх магадлалтай***	

Туршилтад ашигласан гарын үсгийн санд бүртгэлгүй 70 payload-ийн нөхцөлд уг тохиолдол нь зөвхөн зан төлөвт суурилсан аргаар илэрч, Suricata-ийн signature-based илрүүлэлтээр танигдаагүй. Мөн VirusTotal болон ClamAV дээр бүртгэлгүй байсан нь Тэг-өдөр халдлага болохыг баталгаажуулсан. Энэ нь уламжлалт аргуудын хязгаарлалтыг харуулж, санал болгож буй олон түвшний онооны механизм илүү үр дүнтэй болохыг баталж байна.

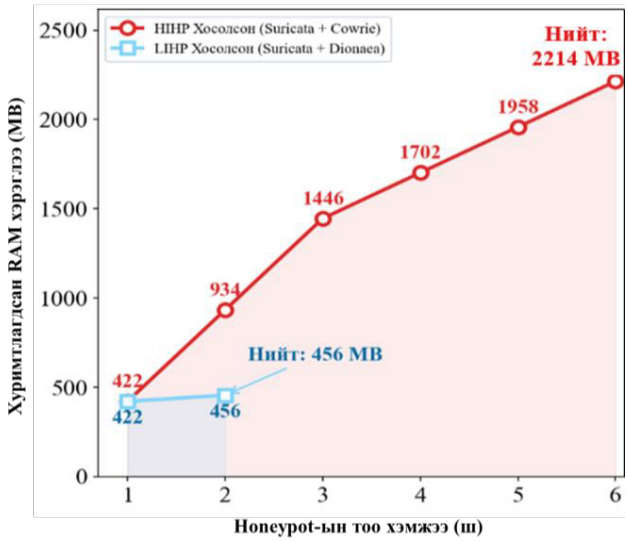
Зураг 8 дээр тэг-өдөр нэр дэвшигчийн зан төлөвт суурилсан үнэлгээний хүчин зүйлсийг радарын диаграммаар харуулж, тус бүрийн онооны тархалтыг дүрслэв.



Зураг 8. Тэг-өдөр болон хэвийн урсгалын харьцуулалт

C. Нөөцийн зарцуулалтын харьцуулалт

Нөөцийн хэмнэлт нь халдлага илрүүлэх системийг бодит орчинд нэвтрүүлэх гол шалгуур юм. Системийн CPU болон RAM-ийн хэрэглээг Ubuntu Server 22.04 орчинд vmstat болон тусгай Python скрипт ашиглан 5 секунд тутамд хэмжиж, өгөгдлийг CSV файлд бүртгэн хадгалсан. Зураг 9 дээр HIHP-д суурилсан IDS болон хосолсон IDS-ийн нөөц ашиглалтын харьцуулалтыг үзүүлэв. Энд хуримтлагдсан RAM ашиглалт нь honeypot-уудын тоо нэмэгдэхийн хэрээр системийн нийт санах ойн хэрэглээ хэрхэн өсөж байгааг илэрхийлж, нөөцийн ачааллыг нэгтгэн үнэлэх боломжийг олгож байна. Ийнхүү уг хэмжилт нь системийн өргөтгөх боломж болон нөөцийн үр ашигт байдлыг харьцуулан дүгнэхэд чухал үзүүлэлт болж байна.



Зураг 9. HIHP, LIHP системийн RAM ашиглалтын харьцуулалт

D. Хяналтын самбарын хэрэгжүүлэлт

Санал болгож буй хосолсон IDS-ийн практик ач холбогдлыг харуулахын тулд Suricata болон Dionaea системүүдийн илрүүлэлтийн үйл ажиллагааг нэгтгэсэн бодит цагийн хяналтын самбарыг хөгжүүлсэн.

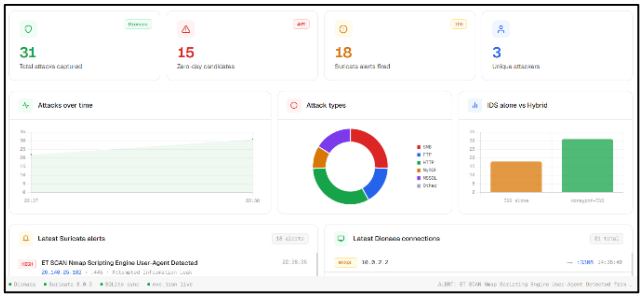
Hybrid IDPS хэсэг нь Dionaea болон Suricata системүүдийн нэгдсэн статистик мэдээллийг харуулж, халдлагын нийт төлөв байдлыг дүгнэхэд ашиглагдана. Suricata IDS хэсэг нь гарын үсгийн мэдээллийн санд суурилсан сэрэмжлүүлгүүдийн дэлгэрэнгүй мэдээллийг үзүүлж, илэрсэн халдлагын төрөл болон давтамжийг тодорхойлно.

Dionaea Honeypot хэсэг нь honeypot-д ирсэн холболтууд, халдагчийн үйлдэл болон бүртгэгдсэн хортой кодын мэдээллийг харуулж, халдлагын эх үүсвэрийг судлахад дэмжлэг үзүүлдэг. Тэг-өдөр хэсэг нь IDS-ийн гарын үсгийн санд бүртгэлгүй, зан төлөвийн шинжилгээгээр илэрсэн сэжигтэй халдлагуудыг тусад нь ялган харуулдаг. Харин Objectives хэсэг нь судалгааны зорилтуудын биелэлтийг нэгтгэн үзүүлж, системийн гүйцэтгэлийг ерөнхийд нь үнэлэх боломжийг бүрдүүлэхийн зэрэгцээ цаашдын сайжруулалт, хөгжүүлэлтийн чиглэлийг тодорхойлоход чухал ач холбогдолтой юм.

3) Hybrid IDPS хуудас

Энэхүү хэсэгт Dionaea болон Suricata системээс ирж буй өгөгдлийг нэгтгэн, аюулгүй байдлын ерөнхий төлөв байдлыг бодит хугацаанд дүрслэн харуулна.

Зураг 10-т харуулсанчлан хяналтын самбарын дээд хэсэгт нийт халдлага, Suricata-ийн анхааруулга болон сэжигтэй IP хаягийн тоог харуулах бөгөөд эдгээр мэдээлэл нь Suricata-ийн eve.json болон Dionaea-ийн SQLite сангаас JSON форматаар ирж, 5 секунд тутамд шинэчлэгддэг. Chart.js ашиглан халдлагын төрөл, идэвхийг графикаар харуулсан нь шинжээчдэд сүлжээний аюулгүй байдалд нэгдсэн, бодит цагийн хяналт тавих боломжийг олгохын зэрэгцээ халдлагын чиг хандлага, давтамжийг илүү ойлгомжтойгоор шинжлэх нөхцөлийг бүрдүүлдэг.

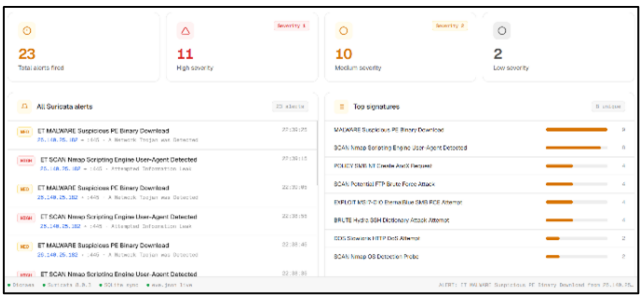


Зураг 10. Hybrid IDPS системийн нэгдсэн хяналтын самбар

4) Suricata IDS хэсэг

Зураг 11-т байгаа хэсэг нь сүлжээний бодит урсгалд шинжилгээ хийж буй Suricata IDS-ээс ирсэн анхааруулгуудыг бодит хугацааны жагсаалт хэлбэрээр харуулна. Нэгж бүр нь халдлагын төрөл, аюултай байдлын зэрэг, эх сурвалж болох IP хаяг болон хугацааны мэдээллийг агуулдаг.

Эдгээр өгөгдлийг Suricata-ийн үүсгэсэн eve.json файлаас бодит хугацаанд татаж авдаг бөгөөд ингэснээр сүлжээний урсгалд гарч буй өөрчлөлт, сэжигтэй үйл ажиллагааг шуурхай илрүүлэх боломж бүрддэг. Аюулын зэрэглэлийг өнгөөр ялгаж харуулсан нь шинжээчдэд нэн тэргүүнд хариу арга хэмжээ авах шаардлагатай халдлагуудыг хурдан тодорхойлон, шийдвэр гаргах процессыг хурдасгах давуу талтай. Мөн хамгийн түгээмэл илрүүлэлтийн гарын үсгүүдийг статистик хэлбэрээр харуулж, халдлагын түгээмэл төрлийг тодорхойлох боломжийг олгосноор цаашдын хамгаалалтын бодлого, дүрмийг оновчтой сайжруулахад чухал дэмжлэг үзүүлдэг.



Зураг 11. Suricata системийн анхааруулгын самбар

ДҮГНЭЛТ

Энэхүү судалгааны хүрээнд Dionaea бага харилцан үйлчлэлтэй honeypot болон Suricata халдлага илрүүлэх системийг хослуулсан гурван модульт идэвхтэй илрүүлэлтийн архитектурыг боловсруулж, тусгаарлагдсан виртуал орчинд туршин баталгаажууллаа. Уг архитектур нь халдлагыг эрт үе шатанд илрүүлэх, бодит цагийн хяналт хийх, мөн автомат хариу арга хэмжээ хэрэгжүүлэх боломжийг бүрдүүлсэн цогц шийдэл бөгөөд олон эх үүсвэрийн өгөгдлийг уялдуулан боловсруулах замаар илрүүлэлтийн үр ашиг, хурдыг нэмэгдүүлсэн онцлогтой юм.

Туршилтыг CIC-IDS2018 дата багц болон халдлагын симуляци ашиглан гүйцэтгэхэд уламжлалт Suricata системийн илрүүлэлтийн түвшин 69.6% байсан бол санал болгож буй хосолсон систем 85.4%-д хүрч, илрүүлэлтийн гүйцэтгэл мэдэгдэхүйц сайжирсан. Үүний зэрэгцээ худал эерэг түвшин 4.4%-аас 1.6% хүртэл буурч, системийн нарийвчлал 98.5%-д хүрсэн нь илрүүлэлтийн чанар болон найдвартай байдлыг дээшлүүлснийг харуулж байна. Энэ нь зөвхөн илрүүлэх чадвар төдийгүй бодит орчинд ашиглахад тохиромжтой тогтвортой ажиллагааг хангаж байгааг илтгэнэ.

Мөн тэг-өдөр халдлагыг үнэлэх зорилгоор гарын үсгийн санд бүртгэлгүй 70 payload дээр туршилт хийж, зан төлөвт суурилсан олон түвшний онооны механизмыг ашигласнаар 85.7%-ийн илрүүлэлтийн үр дүнд хүрсэн. Энэ нь уламжлалт гарын үсэгт суурилсан аргуудын хязгаарлалтыг даван туулах боломжтойг баталж, шинэ болон өмнө бүртгэгдээгүй халдлагыг илрүүлэхэд зан төлөвт суурилсан хандлага үр дүнтэй болохыг харуулж байна. Үүнээс гадна уг аргачлал нь халдлагын хэв шинжийг илүү нарийвчлан тодорхойлох боломжийг олгож, ирээдүйн хамгаалалтын стратегийг сайжруулахад чухал өгөгдөл бүрдүүлж байна.

Нөөцийн хэрэглээний хувьд санал болгож буй систем нь CPU болон RAM-ийн зарцуулалтыг хэд дахин бууруулж, өндөр харилцан үйлчлэлтэй honeypot системтэй харьцуулахад илүү хөнгөн, өргөтгөх боломжтой шийдэл болохыг нотолсон. Мөн бодит цагийн хяналтын самбарыг хөгжүүлснээр олон эх үүсвэрийн өгөгдлийг нэгтгэн харах, халдлагын төлөв байдлыг шуурхай үнэлэх боломжийг бүрдүүлсэн нь практик хэрэглээнд нийцсэн давуу тал болж байна. Үүний үр дүнд системийн удирдлага, мониторинг илүү хялбар болж, шийдвэр гаргалтыг бодит өгөгдөлд тулгуурлан хийх нөхцөл бүрдсэн.

Хэдийгээр санал болгож буй хосолсон системийн илрүүлэлтийн түвшин өндөр харилцан үйлчлэлтэй honeypot-д суурилсан IDS системээс тодорхой хэмжээгээр (8.1%-иар) бага байгаа боловч энэхүү ялгаа нь нөөцийн ашиглалтын хувьд олж авсан давуу талтай харьцуулахад хүлээн зөвшөөрөхүйц солилцоо юм. Тодруулбал, систем нь CPU болон RAM-ийн хэрэглээг хэд дахин бууруулж, хязгаарлагдмал нөөц бүхий орчинд тогтвортой ажиллах боломжийг бүрдүүлсэн. Иймд илрүүлэлтийн түвшинд бага зэрэг бууралт гарсан хэдий ч системийн нийт үр ашиг, өргөтгөх боломж, бодит орчинд хэрэгжих чадвар мэдэгдэхүйц сайжирсан гэж үзэж болно. Энэ нь практик хэрэглээнд өндөр нөөц шаардсан шийдлүүдээс илүү оновчтой, тэнцвэртэй хандлага болохыг харуулж байна.

Ном зүй

[1] R. Agrawal, J. W. Stokes, L. Rist, R. Littlefield, X. Fan, K. Hollis, and C. Seifert, “Long-term study of honeypots in a public cloud,” in Proc. 52nd Annu. IEEE/IFIP Int. Conf. Dependable Systems and Networks (DSN-S), 2022, pp. 1–4.

[2] E. Alatawi and U. Albalawi, “Harnessing ai for cyber defense: Honeypot-driven intrusion detection systems,” Symmetry, vol. 17, no. 5, p. 628, 2025.

[3] AV-TEST Institute, “Malware statistics & trends.” [Online]. Available: https://www.av-test.org/en/statistics/malware/

[4] C. Charrier, J. Sadowski, C. Lecigne, and V. Stolyarov, “Hello 0-days, my old friend: A 2024 zero-day exploitation analysis,” Google Cloud Blog, 2025.

[5] J. L. Gutierrez-Garcia, E. Sanchez-DelaCruz, and M. del Pilar, “A review of intrusion detection systems using machine learning: Attacks, algorithms, and challenges,” unpublished.

[6] N. Ilg, D. Germek, P. Duplys, and M. Menth, “Beekeeper: Accelerating honeypot analysis with llm-driven feedback,” IEEE Access, 2025.

[7] C. Kelly, N. Pitropakis, A. Mylonas, S. McKeown, and W. J. Buchanan, “A comparative analysis of honeypots on different cloud platforms,” Sensors, vol. 21, no. 7, p. 2433, 2021.

[8] Y. Kocaogullar, O. Cetin, B. Arief, C. Brierley, J. Pont, and J. Hernandez-Castro, “Hunting high or low: Evaluating the effectiveness of high-interaction and low-interaction honeypots,” in Proc. Int. Workshop Socio-Technical Aspects in Security, 2022, pp. 14–30.

[9] Kubba, Q. Nasir, O. Elmutasim, and M. Abu Talib, “A systematic review of honeypot data collection, threat intelligence platforms, and ai/ml techniques,” unpublished.

[10] H. Lajwanti, F. Urooj, W. Muhammad, and S. Mehwish, “Enhancing cybersecurity through honeypot-based intrusion detection and prevention systems,” in Proc. Int. Multidisciplinary Conf. Emerging Trends Eng. Technol., 2024, pp. 149–154.

[11] Lakshmipathy, M. Gurusamy, S. F. Lalo, N. S. Lebang, K. Selvaraj, and A. A. Abdulsahib, “Advanced network security system: Honeypot-based intrusion detection with machine learning and visualization,” technical report.

[12] M. Landauer, F. Skopik, B. Stojanović, A. Flatscher, and T. Ullrich, “A review of time-series analysis for cyber security analytics: From intrusion detection to attack prediction,” Int. J. Inf. Security, vol. 24, article 3, 2025.

[13] Z. Moric, L. Mršić, Z. Kunić, and G. Đambić, “Honeypots in cybersecurity: Their analysis, evaluation and importance,” 2024.

[14] Nimmagadda and S. Y. Mehr, “Ai-powered intrusion detection system with honeypot integration,” Int. J. Intelligent Information Systems, vol. 14, no. 4, pp. 70–76, 2025.

[15] T. S. Parlanti and C. A. Catania, “Temporal analysis framework for intrusion detection systems: A novel taxonomy for time-aware cybersecurity,” in press.

[16] SANS Institute, “Deception techniques as part of an intrusion detection strategy.” [Online]. Available: https://www.sans.org/white-papers/37140/

[17] X. Yang, J. Yuan, H. Yang, Y. Kong, H. Zhang, and J. Zhao, “A highly interactive honeypot-based approach to network threat management,” Future Internet, vol. 15, no. 4, p. 127, 2023.

IEEE 802.11 сүлжээ бүхий дроны сүргийн системийн auction-based даалгавар хуваарилах механизмын DDOS халдлагын нөлөөллийг симуляциар үнэлэх нь

П.Түвшинбаяр*, Ц.Энхтөр , Я.Дашдорж, Ч.Эрдэнэбат, Х.Уянгаа
Кибер аюулгүй байдлын тэнхим
Мэдээлэл холбооны технологийн сургууль,
Шинжлэх ухаан, технологийн их сургууль
Монгол улс, Улаанбаатар
tuvshinbayar.tend@gmail.com, ch.Erdenebat@gmail.com ,{enkhtur, uyangaakh, dashdorj}@must.edu.mn

***Хураангуй*— Нисгэгчгүй нисэх төхөөрөмжийн (UAV) сүргийн системд даалгаврыг төвлөрсөн бус аргаар оновчтой хуваарилахын тулд өрсөлдөөнт санал өргөх (auction-based) механизмыг ашигладаг. Энэхүү механизм нь зарлал (announcement), санал (bid), шийдвэр (decision), баталгаажуулалт (confirmation) гэсэн дөрвөн дараалсан үе шаттайгаар хэрэгждэг бөгөөд 2.4 ГГц зурвасын нэг алхамт (single-hop) UDP broadcast харилцаанд тулгуурладаг тул үйлчилгээ бусниулах (DDoS) халдлагад өртөмтгий байдаг. Халдлага нь дээрх үе шатуудын аль нэгэнд хугацаа хэтрэлт (timeout) үүсгэснээр сүргийн зохицуулалтыг бүхэлд нь доголдуулах эрсдэлтэй. Энэ судалгаагаар Physical (PHY), Medium Access Control (MAC), Transport болон Application давхаргыг нэгтгэсэн (Cross-layer) загварыг SimPy дискрет үйл явдлын симулятор болон ITU-R P.1238 стандартаар өргөтгөсөн чөлөөт орон зайн дохионы алдагдлын загвар Free-Space Path Loss (FSPL)-д суурилан боловсруулж, халдлагын нөлөөллийг системчлэн үнэлэв. Системийн тэсвэртэй байдлыг тодорхойлохдоо даалгавар биелүүлэлтийн харьцаа (TCR) болон хоцрогдлын харьцаа (TDR)-аас гадна сүргийн сэргэлтийн хугацаа (SRT) хэмжүүрийг Degree Centrality алгоритмд суурилсан LeaderFailover механизмаар анхлан тодорхойлж, 36 параметрийн сценариар туршилт явуулсан. Туршилтын үр дүнд халдлагагүй горимд TCR=0.528, пакет дамжуулалтын харьцаа PDR=0.826 байсан бол, хүчтэй халдлагын үед (bw=0.90) PDR=0.098 болж буурснаар 550 удаагийн туршилтын аукционы нэг нь ч амжилттай болоогүй, TCR=0.000 үзүүлэлтээр сүргийн зохион байгуулалт бүрэн зогссоныг нотоллоо. SRT хэмжүүр дунджаар 1.033 секунд бүртгэгдсэн нь LeaderFailover механизмын тогтвортой байдлыг харуулж байна. Судалгааны үр дүнд бүрдүүлсэн 41 шинж чанар (багана), 200 бичлэг (мөр) бүхий шошгожуулсан өгөгдлийн сан нь машин сургалтад суурилсан DDoS илрүүлэлт болон кибер тэсвэртэй байдлын цаашдын судалгааны онол, практикийн чухал суурь болно.**

Түлхүүр үг—*UAV swarm, auction протокол, тэсвэртэй байдал, cross-layer үнэлгээ, симуляци*

I. УДИРТГАЛ

Сүүлийн жилүүдэд нисгэгчгүй нисэх төхөөрөмж (UAV)-ийн сүрэг (swarm)-ийн технологи нь цэргийн тагнуул хяналт, гал түймэр унтраах ажиллагаа, хилийн хамгаалалт зэрэг хүний нөөц хязгаарлагдмал буюу амь насанд аюултай орчинд өргөнөөр нэвтэрч, практик хэрэглээний ач холбогдол нь улам бүр нэмэгдэж байна [1], [2]. Гэвч энэхүү технологи нь анхнаасаа гүйцэтгэлд чиглэн зохиогдсон тул кибер аюулгүй байдлын хувьд нэн эмзэг хэвээр байна.

UAV сүргийн үр ашигтай зохицуулалтын гол механизм болох өрсөлдөөнт санал өргөх зарчимд суурилсан даалгавар хуваарилалт нь 2.4 ГГЦ-ийн зурваст UDP (User Datagram Protocol) broadcast протоколоор ажилладаг тул DDoS (Distributed Denial of Service) халдлагад онцгойлон өртөмтгий байдаг [3], [4],[5]. DDoS халдлага нь зорилтот сүлжээнд их хэмжээний хуурамч пакет илгээснээр сувгийн хүчин чадлыг дүүргэж, мэдэгдэл-санал-шийдвэр-баталгаажуулалт гэсэн дөрвөн үе шатын аль нэгийг тасалдуулснаар бүхэл сүргийн координацийг доголдуулдаг.

Jamming (радио дарангуйлалт) халдлагаас ялгаатай нь UDP flooding халдлага нь физик давхарга (PHY)-ын дохионы чанарт биш хандалтын хяналт (MAC) болон дамжуулалтын (Transport) давхаргын нөөцөд нөлөөлдөг тул илрүүлэхэд хэцүү бөгөөд хэрэглээний давхаргын логикт шууд бус байдлаар хүрдэг.

Өмнөх судалгаанууд сүлжээний давхарга бүрийн нөлөөллийг тусад нь судалсан бөгөөд давхарга дамнасан нэгдсэн үнэлгээний хандлага хязгаарлагдмал байсаар байна [6], [8]. Мөн өрсөлдөөнт санал өргөх зарчимд суурилсан UAV сүргийн тэсвэртэй байдлыг үнэлэхдээ системийн уналтыг хэмжихэд анхаарч, сэргэлтийг тодорхойлох нэгдсэн хэмжүүр дутагдалтай байсан. Энэхүү судалгаанд дөрвөн сүлжээний давхаргыг (PHY-MAC-Transport-Application) нэгтгэсэн давхарга дамнасан үнэлгээний загвар боловсруулж, даалгавар биелүүлэлтийн харьцаа (TCR), даалгавар хоцрогдлын харьцаа (TDR)-ийн зэрэгцээ сүргийн сэргэлтийн хугацаа (SRT) шинэ хэмжүүрийг графын онолын Degree Centrality алгоритмд суурилсан LeaderFailover механизмаар анх тодорхойлсноор системийн тэсвэртэй байдлыг цогцоор үнэлэх боломжийг бүрдүүлэв.

Энэхүү судалгаанд SimPy дискрет үйл явдлын симулятор болон чөлөөт орон зайн дохионы алдагдлын загвар (FSPL — Free Space Path Loss)-ыг ITU-R P.1238 стандартын орчны хүчин зүйлээр өргөтгөн нэгтгэсэн нэгдсэн симуляцийн платформ боловсруулсан бөгөөд уг платформ нь гучин зургаан параметрийн сценарийг судалж, машин сургалтад бэлэн өгөгдлийн багц үүсгэдэг.

Цаашид II дүгээр хэсэгт холбогдох судалгааны тойм, III дүгээр хэсэгт арга зүй, IV дүгээр хэсэгт туршилтын үр дүн, V дүгээр хэсэгт дүгнэлтийг тус тус авч үзнэ.

A. Судалгааны асуулт

2.4 GHz WiFi (802.11n) сүлжээнд Auction-based даалгавар хуваарилалтын механизм ашигладаг UAV сүргийн систем, DDoS UDP Flooding халдлагад өртөх үед, халдлагагүй хэвийн үйл ажиллагааны горимтой харьцуулахад, TCR, TDR, SRT хэмжүүрүүдээр тодорхойлогдох тэсвэрлэх чадвар (Resilience) хэрхэн өөрчлөгдөж буйг Python SimPy + FSPL загвараар тодорхойлох, үнэлэх системийг хэрхэн боловсруулах вэ?

B. Судалгааны зорилтууд

З₁: DDoS эрчийн өсөлттэй TCR болон PDR-ийн хамаарлыг PHY–MAC–Transport–Application дөрвөн давхаргын хэмжүүрүүдээр (PHY PDR, MAC collision probability, Transport ρ, App TCR) тоон байдлаар тодорхойлох.

З₂: Auction протоколын дөрвөн үе шатад (Announce→Bid→Decision→ACK) DDoS халдлагын нэмэлт cascade доголдлын дараалыг baseline-тай харьцуулан симуляцийн өгөгдлөөр баримтжуулах.

Г. Өгүүллийн шинэлэг тал ба гол хувь нэмэр

Өмнөх судалгаатай харьцуулахад энэхүү бүтээл дараах гурван шинэ хувь нэмэр оруулж байна.

- **Анхны давхарга дамнасан (cross-layer) нэгдсэн загвар:** Liu et al. [1] нь UAV сүргийн аюулгүй байдлын тойм судалгаандаа DDoS халдлагыг дурдсан боловч өрсөлдөөнт санал өргөх зарчимд суурилсан даалгавар хуваарилалтад үзүүлэх нарийвчилсан нөлөөг тоон хэмжүүрээр үнэлээгүй байна. Zhang болон Chen [2] нар давхарга тус бүрийн нөлөөллийг тусад нь судалсан бол энэхүү судалгаагаар PHY-MAC-Transport-Application гэсэн дөрвөн давхаргыг SimPy болон ITU-R P.1238 нэгдсэн платформд анх удаа нэгтгэв. Ингэснээр DDoS халдлагын каскад (cascade) доголдлын механизмыг тоон хэмжүүрээр баталгаажуулсан болно.
- **SRT хэмжүүрийг анх тодорхойлсон нь:** Patel болон бусад [5] болон Ahmed болон бусад [6] нар сүргийн уналтыг хэмжсэн боловч сэргэлтийн хугацааг тоон байдлаар тодорхойлох хэмжүүр нийтлэгдээгүй байсан юм. Энэхүү судалгаа нь Degree Centrality алгоритмд суурилсан LeaderFailover механизмыг хэрэгжүүлж, SRT (Swarm Recovery Time) хэмжүүрийг анх нормалчилан нэвтрүүлэв.
- **Машин сургалтад (ML) зориулсан шошгожуулсан өгөгдлийн сан:** Auction протоколд зориулсан, 36 сценари, 41 шинж чанар бүхий DDoS тэсвэртэй байдлын шошгожуулсан өгөгдлийн санг үүсгэв. Энэхүү сан нь өмнөх судалгаануудад байхгүй байсан бөгөөд машин сургалтын DDoS илрүүлэлтийн цаашдын судалгааны суурь болох ач холбогдолтой.

II. ХОЛБОГДОХ СУДАЛГААНЫ ТОЙМ

A. UAV сүргийн аюулгүй байдал

"Liu et al. [1] нь UAV сүргийн аюулгүй байдлын тойм судалгаандаа DDoS халдлагыг дурдсан боловч өрсөлдөөнт санал өргөх зарчимд суурилсан даалгавар хуваарилалтад үзүүлэх нарийвчилсан нөлөөг тоон хэмжүүрээр үнэлээгүй байна.

Харин Zhang болон Chen [2]-ийн судалгаа нь дотоод үнэлгээ болон гадаад өрсөлдөөний хоёр үе шаттай өрсөлдөөнт санал өргөх механизмын үр ашгийг сайжруулсан боловч DDoS халдлагын орчинд туршилт хийгээгүй тул системийн тэсвэртэй байдлыг баталгаажуулаагүй болно.

Б. DDoS илрүүлэлт

Ahmed [6], Nguyen болон Park [7] нарын урт богино хугацааны санах ойн автокодлогч (Long Short-Term Memory Autoencoder) судалгаанууд нь сүлжээний үйлчилгээ тасалдуулах (DDoS) халдлагыг илрүүлэлтэд сайн үр дүн үзүүлсэн боловч нисгэгчгүй нисэх төхөөрөмжийн (UAV) сүргийн даалгаврын тэсвэртэй байдлын хэмжүүрийг тусгаагүй тул дроны сүргийн системд шууд хэрэглэх боломжгүй байна. Zhao болон бусад [8] нь сүлжээний аномали илрүүлэлтийн загварыг оновчлосон боловч UAV сүргийн өрсөлдөөнт санал өргөх протоколд тусгайлан хэрэглэсэн туршилтын үр дүн дутагдалтай байна.

В. Өмнөх судалгааны хязгаарлалтууд

Харьцуулалтын шалгуур бүрийг судалгааны үндсэн асуулт болон системийн техникийн шаардлагад нийцүүлэн тодорхойлов. Энэхүү судалгааны үндсэн асуулт нь DDoS халдлагын нөхцөлд auction-based UAV сүргийн даалгавар хуваарилалтын механизмын тэсвэртэй байдлыг тоон байдлаар үнэлэхэд чиглэнэ. Иймд сонгосон шалгуур бүр нь энэхүү асуултыг үнэлэхэд зайлшгүй шаардлагатай нөхцөлийг илэрхийлнэ.

Нэгдүгээрт, UAV+Auction шалгуур нь судалгааны хамрах хүрээг тодорхойлох үндсэн шалгуур бөгөөд дурдсан хоёр шинжийг зэрэг агуулаагүй судалгаа нь асуудлын тодорхойлолттой нийцэхгүй тул шууд харьцуулах боломжгүй. Иймд энэхүү шалгуур нь судалгаанд хамруулах (inclusion) шаардлагыг тогтооно.

Хоёрдугаарт, DDoS шалгуур нь судалгааны гол аюул заналхийллийг илэрхийлнэ. DDoS орчинг тусгаагүй загвар нь системийн тэсвэртэй байдлыг үнэлэх боломжгүй бөгөөд зөвхөн хэвийн нөхцөл дэх гүйцэтгэлийг харуулахад хязгаарлагдана. Иймд энэхүү шалгуур нь туршилтын орчны зайлшгүй бүрэлдэхүүн хэсэг болно.

Гуравдугаарт, FSPL 2.4 GHz шалгуур нь физик давхаргын бодит нөхцөлийг тусгах зорилготой. UAV сүргийн харилцаа нь утасгүй холбоонд тулгуурладаг тул 2.4 ГГц зурваст дохионы алдагдлыг тооцохгүй бол сүлжээний гүйцэтгэлийн (packet loss, delay) шалтгааныг буруу тайлбарлахад хүргэнэ. Иймд энэхүү шалгуур нь симуляцийн бодитой байдлыг хангах зайлшгүй нөхцөл болно.

Дөрөвдүгээрт, TCR, TDR, SRT хэмжүүрүүд нь application давхаргын түвшинд системийн гүйцэтгэл, тэсвэртэй байдлыг тоон байдлаар үнэлэхэд шаардлагатай. Эдгээр хэмжүүргүйгээр даалгавар хуваарилалтын механизмд DDoS халдлагын үзүүлэх нөлөөллийг тоон байдлаар харьцуулах боломжгүй. Иймд энэхүү шалгуур нь үнэлгээний хэмжүүрийн бүрэн бүтэн байдлыг хангана.

Иймээс дээрх дөрвөн шалгуур нь (i) судалгааны хамрах хүрээг тодорхойлох, (ii) туршилтын орчны бодит байдлыг хангах, (iii) үр дүнг хэмжих боломжийг бүрдүүлэх гэсэн

гурван үндсэн шаардлагыг зэрэг хангаж, өмнөх судалгааны цоорхойг илрүүлэх зайлшгүй шаардлагатай шалгуурын багц болно.

Доорх Хүснэгт 1-т өмнөх судалгаануудын харьцуулалтыг харуулав. Судалсан бүтээлүүдийн аль нь ч тодорхойлсон дөрвөн шалгууруг бүрэн хангахгүй байгаа нь энэхүү судалгааны ажлын шаардлагатай байдлыг илэрхийлж байна.

ХҮСНЭГТ I . ХОЛБОГДОХ СУДАЛГААНЫ ХАРЬЦУУЛАЛТ

№	Зохиогч, он	UAV+Auction	DDoS	FSPL 2.4G	TCR/TDR/SRT
[1]	Liu et al., 2024	✓	✓	✗	✗
[2]	Zhang & Chen, 2024	✓	✗	✗	✗
[3]	Torres et al., 2025	✓	✗	✗	✗
[5]	Patel et al., 2025	✓	✓	✗	✗
[6]	Ahmed et al., 2023	✗	✓	✗	✗
*	Бидний судалгаа, 2026	✓	✓	✓	✓

III. АРГА ЗҮЙ

А. Судалгааны ерөнхий дизайн

Энэхүү судалгааны ажилд тоон (quantitative) судалгааны арга зүйг баримтлан ажиллав. Судалгааны туршилтын загварыг SimPy дискрет үйл явдлын симулятор дээр боловсруулсан бөгөөд радио долгионы тархалтын алдагдлыг тооцохдоо чөлөөт орчны тархалтын алдагдал (FSPL – Free Space Path Loss) болон дотоод орчны саад, барилга байгууламжийн нөлөөг тусгасан ITU-R P.1238 стандартын загваруудыг хослуулан ашигласан. Туршилтын үр дүнг эмпирик аргаар баталгаажуулсан бөгөөд энэхүү холимог загвар нь UAV сүргийн зангилаа хоорондын холболтын бодит байдлыг хангах, ялангуяа DDoS халдлагын үе дэх пакет гээгдэлтийг (PDR) давхарга бүрээр нарийвчлалтай тооцох нөхцөлийг бүрдүүлсэн болно.

SimPy дискрет үйл явдлын симуляторыг сонгосон үндсэн шалтгаан нь гурван хүчин зүйлээр тодорхойлогдоно. Үүнд , (1) аукционы протоколын дөрвөн үе шат (Announce→Bid→Decision→ACK) болон LeaderFailover механизмын үйл явдал бүрийг тасралтгүй дараалаар бүртгэж, SRT сэргэлтийн хугацааг үйл явдал тус бүрийн түвшинд нарийвчлан тооцоолох боломжтой, (2) PHY-MAC-Transport-Application дөрвөн давхаргын загварыг нэгдсэн орчинд Python-оор уян хатан хэрэгжүүлэх боломжтой, (3) симуляцийн үр дүнд 41 шинж чанар бүхий шошгожуулсан өгөгдлийн санг ML экосистемтэй шууд уялдуулан үүсгэх боломжтой.

Б. Симуляцийн орчны гол параметрүүд

Судалгааны туршилтын үр дүнг бодит нөхцөлд ойртуулахын тулд симуляцийн орчны параметрүүдийг дараах байдлаар тохируулав. Симуляцид 2.4 ГГц давтамжтай IEEE 802.11n стандартыг ашигласан бөгөөд UAV тоог 10, 15, 20 агентаар хувьсахаар тохируулсан. DDoS халдлагын эрчийг 0–1000 пакет/сек (байхгүй, бага, дунд, өндөр) хооронд өөрчилж, 3 халдагч зангилаа (n_attackers=3) бүхий ботнет (botnet) сүлжээг үүсгэв. Төхөөрөмжүүдийн шилжилт хөдөлгөөнийг Гаусс-Марковын (Gauss-Markov) загвараар (α=0.75), радио долгионы санамсаргүй саатлыг Рейлигийн (Rayleigh) сүүдрийн стандарт хазайлтаар (σ=8дБ) тус тус тооцсон. Сүлжээний орчныг ITU-R P.1238 стандартын дагуу нээлттэй (N=20) болон хот орчмын (N=30) нөхцөлөөр төлөөлүүлэн чөлөөт орон зайн замын алдагдлын (FSPL) томъёог өргөтгөн ашиглан загварчилсан болно. Туршилтын дизайныг санамсаргүй утгын үржвэрээр (3 seed × 4 DDoS × 3 орчны нөхцөл = 36 сценари) тооцож, нийт 36 сценарийн хүрээнд туршилтыг гүйцэтгэв. Симуляцийн үндсэн параметрүүдийг Хүснэгт II-т нэгтгэн харуулав.

ХҮСНЭГТ II. СИМУЛЯЦИЙН ҮНДСЭН ПАРАМЕТРҮҮД

Параметр	Утга
Давтамж	2.4 ГГц (IEEE 802.11n)
UAV-ийн тоо	10, 15, 20 агент
DDoS халдлагын эрч	0, 200, 500, 1000 пакет/сек
Халдагчийн тоо (nattackers)	3
Хөдөлгөөний загвар	Gauss-Markov (α=0.75)
Тархалтын алдагдал (N)	20 (Нээлттэй), 30 (Хот орчин)
Сүүдрийн эффект (σ)	8 дБ (Rayleigh)

Симуляцийн параметрүүдийг дараах үндэслэлээр сонгов. UAV-ийн тоог 10, 15, 20 агентаар тохируулснаар жижиг, дунд, том хэмжээний сүргийн гүйцэтгэлийг харьцуулан үнэлэх боломжийг бүрдүүлсэн бөгөөд энэ хүрээ нь auction-based даалгавар хуваарилалтын судалгааны стандарт хэмжээтэй нийцнэ[9].

DDoS халдлагын эрчийг 0, 200, 500, 1000 пакет/сек-ээр тохируулснаар M/M/1 дараалллын сувгийн ашиглалтын коэффициент ρ-г бага (ρ≈0.255), дунд (ρ≈0.553), өндөр (ρ≈0.765) гэсэн гурван түвшинд хянах боломжийг олгосон [6].

Симуляцид IEEE 802.11n стандартын нэрлэсэн дамжуулалтын хурд 72 Мбит/с-ийн оронд CSMA/CA-ийн зогсолт болон протоколын нэмэлт ачааллыг тооцсон практик throughput 54 Мбит/с-ийг ашигласан бөгөөд 1500 байтын пакетийн хувьд үйлчилгээний хурд μ ≈ 4500 пакет/сек болно.

DDoS халдлагын хамгийн хүчтэй горимд (bw=0.90) DDoS-ын нэмэлт ачаалал λ_ddos = bw × μ × 0.85 = 0.90 × 4500 × 0.85 ≈ 3443 пакет/сек бөгөөд сувгийн ашиглалтын коэффициент ρ_max = λ_ddos / μ = 3443 / 4500 ≈ 0.765 < 1 байгаа нь M/M/1 дараалллын тогтвортой байдлын нөхцөлийг хангаж, симуляцийн орчин бодит сүлжээний даах чадвартай нийцэж байгааг баталж байна.

Халдагчийн тоог n=3 гэж тогтоосон нь 10–20 агентын жижиг сүлжээнд сувгийн хэт ачааллыг үүсгэхэд шаардлагатай хамгийн бага ботнет хэмжээтэй нийцнэ [6].

Gauss-Markov загварын авто хамаарлын коэффициент α=0.75-г сонгосон нь UAV-ийн хөдөлгөөний бодит зан чанарыг — бүрэн санамсаргүй алхам (α=0) болон шугаман жигд хөдөлгөөн (α=1)-ий завсрын төлөвийг — өндөр нарийвчлалтайгаар дүрслэх оновчтой утга болохыг UAV хөдөлгөөний загварын судалгаанд баталгаажуулсан [13], [14].

Rayleigh сүүдрийн стандарт хазайлт σ=8дБ болон орчны замын алдагдлын зэтгэлэг N=20 (нээлттэй орчин), N=30 (хот орчин) утгыг ITU-R P.1238 [11] стандартын дагуу сонгосон.

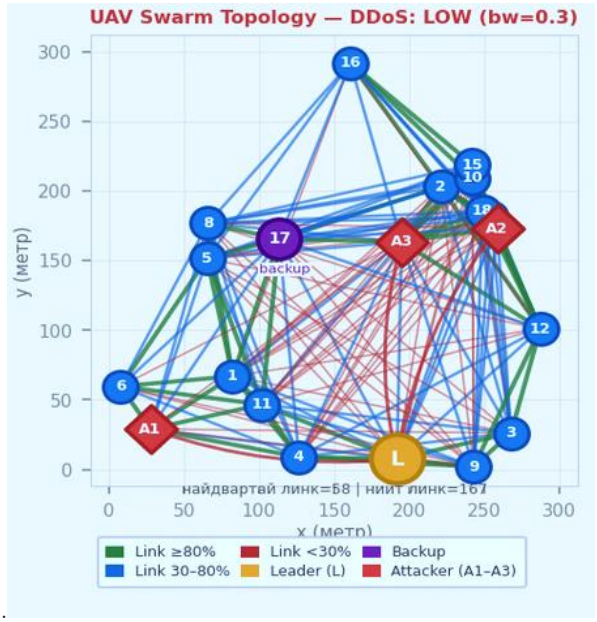
В. Туршилтын топологи ба DDoS халдлагын сценари

Зураг 1-д туршилтын бага эрчимтэй (LOW) горимын үеийн дроны сүргийн сүлжээний топологийг харуулав. Халдлагын эрчмийн коэффициентийг bw = 0.3 гэж тогтоосон энэ горимд симуляцийн орчинд нийт 167 холболт үүссэний дотроос 68 холболт нь найдвартай (PDR ≥ 80%) байсан. Холболтын чанарыг гурван ангилалд хуваан өнгөөр ялгасан. Энд ногоон (Link ≥ 80%) нь өндөр найдвартай, цэнхэр (Link 30–80%) нь дунд зэргийн, улаан (Link < 30%) нь сул буюу эмзэг холбоосыг тус тус илэрхийлнэ.

LOW горимд bw = 0.30 буюу сувгийн хүчин чадлын 30%-ийг DDoS трафик эзэлснээр M/M/1 дараалын ачааллыг нэмэгдүүлж, удирдагч L зангилааны буферийг хэтрүүлэн аукционы протоколын хариу өгөх чадамжийг алдагдуулна. Зураг 1-ээс харахад bw = 0.3 параметрийн нөлөөгөөр PDR < 30%-ийн улаан холболтуудын эзлэх хувь мэдэгдэхүйц өндөр байгаа нь халдлагын нөхцөлд сүлжээний холболтын чанар доройтож байгааг илтгэнэ.

Үндсэн удирдагч зангилаа ажиллагаагүй болох үед LeaderFailover механизм идэвхжинэ. Шинэ удирдагчийг сонгохдоо Графын онолын Холбоосын Зэрэглэлийн (Degree Centrality) алгоритмыг ашигладаг бөгөөд удирдагчаас бусад зангилааны дотроос хүрэх пакет хүргэлтийн харьцаа PDR > 0.5 нөхцөлийг хангасан хамгийн олон идэвхтэй холболттой зангилааг шинэ удирдагчаар сонгоно.

Тухайн туршилтын нөхцөлд LeaderFailover механизм нь командын шугамын тасралтгүй байдлыг хангаж, сүргийн системийн тогтвортой байдлыг халдлагын нөхцөлд ч бодитоор баталгаажуулсан болно.



1-р зураг. UAV сүргийн сүлжээний топологи — DDoS халдлагын бага эрчимтэй (LOW, bw=0.3) горим. Зангилааны тоо: 18, халдагч дронууд: A1, A2, A3 (ромб), үндсэн удирдагч: L (шар), нөөц удирдагч: 17 (ягаан). Найдвартай холбоос=68, нийт холбоос=167.

Г. Давхарга дамнасан онолын загварчлал

Судалгааны гол шинэлэг тал болох давхарга дамнасан нэгдсэн загвар нь дөрвөн протоколын давхаргыг нэг аналитик тогтолцоонд нэгтгэдэг.

1. Давхарга дамнасан дохионы алдагдлын загвар: Физик (PHY) давхаргад дохионы замын алдагдлыг тооцоолохдоо FSPL-г ITU-R P.1238 стандартын орчны засварын хүчин зүйлээр өргөтгөн ашигласан [11].

d зайд (метр) ногдох нийт замын алдагдал PL(d)-ийг дараах байдлаар илэрхийлнэ.

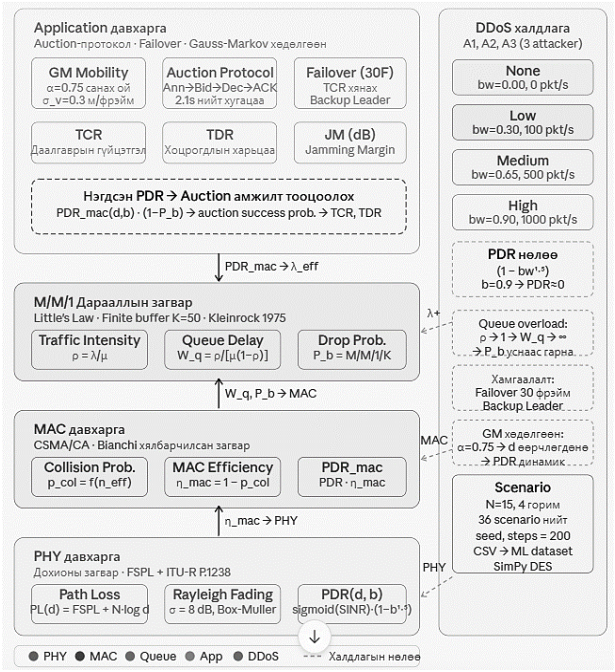
$$PL(d) = 20 \cdot \log_{10}(f) + N \cdot \log_{10}(d) - 28 \quad (1)$$

Энд d — зангилаа хоорондын зай (метр),

f — тээгч давтамж (*МГц-ээр 2400 МГц*) ,

N — орчноос хамаарах замын алдагдлын зэтгэлэг (нээлттэй орчинд $N = 20$; хот орчинд N = 30).

−28 тогтмол нь FSPL-ийн үндсэн томьёо $20 \cdot \log_{10}(4\pi df/c)$-г задлахад гарах физик тогтмол $20 \cdot \log_{10}(4\pi/c) = -147.55$ дБ болон давтамжийг Гц-ээс МГц нэгжид хөрвүүлэхэд нэмэгдэх $20 \cdot \log_{10}(10^6) = +120$ дБ-ийн нийлбэр буюу −147.55 + 120 ≈ −28 болох нэгжийн нормчлолын тогтмол юм.



2-р зураг. UAV сүргийн систем дэх давхарга дамнасан (Cross-layer) харилцан хамаарлын нэгдсэн загвар

UAV сүрэг нь нэг үсрэл (single-hop) утасгүй дамжуулалтын горимд ажилладаг тул IP давхаргын чиглүүлэлт (сүлжээний давхаргын дамжуулалт) энэхүү загварт зориудаар оруулаагүй бөгөөд энэ нь нэг үсрэлт ad hoc харилцааны хүрээтэй нийцсэн зохистой хялбарчилсан загвар мөн [12].

2. Пакет хүргэлтийн харьцааны загвар: Пакет хүргэлтийн харьцаа (PDR) нь утасгүй холбооны шугамын чанар болон DDoS халдлагын эрчмийн нөлөөг хамтад нь тусгадаг. Холбооны чадлын нөөцөд (link margin) тулгуурлан PDR-ийг дараах байдлаар загварчилна.

$$PDR(d, bw) = \sigma(link_margin) \times (1 - bw^{1.5}) \quad (2)$$

Энд link_margin = Ptx − PL(d) + shadow − Ssens (дБ) нь дамжуулагчийн чадал, замын алдагдал, сүүдэрлэлтийн нөлөө (shadowing) нийлбэрээс тооцсон холбооны чадлын нөөц,

$\sigma(x) = 1/(1 + e^{-0.5x})$ нь sigmoid функц.

bw ∈ [0, 1) — DDoS халдлагын сувгийн эзлэх хувь *(low: 0.30, medium: 0.65, high: 0.90).*

$(1 - bw^{1.5})$ үржигч нь DDoS эрчим нэмэгдэх тусам PDR шугаман бус байдлаар буурахыг загварчилдаг. Тасралтгүй DDoS трафик дамжуулалтын үед сувгийн нөөц хурдан шавхагдаж, дахин дамжуулалтын ачаалал нэмэгдснээр PDR хосолмол хэлбэрээр буурна.

3. M/M/1 Дарааллын саатлын загвар: Пуассоны пакет ирэх хурд болон экспоненциал үйлчилгээний хугацааг таамаглан дроны зангилаа бүр дэх дарааллаас үүдэлтэй саатлыг M/M/1 дарааллын загварыг ашиглан тодорхойлно.

UAV сүргийн зангилаа бүр 2.4 ГГц-ийн нэг ортогональ сувгаар UDP broadcast мессеж хүлээн авдаг бөгөөд олон тооны бие даасан эх үүсвэрээс (агент болон халдагч) ирэх пакетын урсгал нь Пуассоны хуваарилалтад ойртдог [15]. Мөн зангилаа бүрийн үйлчилгээний хугацаа нь сүлжээний өөрчлөгдөх ачаалал, дахин дамжуулалт, CSMA/CA-ийн санамсаргүй зогсолтоос хамааран хувьсах тул экспоненциал хуваарилалтаар ойролцоолж болох бөгөөд нэг үсрэлт (single-hop) broadcast орчинд нэг сувгийн дараалал M/M/1 загвараар тодорхойлогдох нөхцөл бүрддэг тул энэхүү загвар нь UAV сүргийн системд онолын хувьд нийцнэ [15], [16].

Пуассоны пакет ирэх хурд болон экспоненциал үйлчилгээний хугацааг таамаглан дроны зангилаа бүр дэх дарааллаас үүдэлтэй саатлыг M/M/1 дарааллын загварыг ашиглан тодорхойлно. Трафикийн эрчим буюу хурд ρ ба дараалалд хүлээх дундаж хугацаа Wq-ийг дараах томъёогоор тооцоолно.

$$\rho = \lambda/\mu, \quad Wq = \rho / [\mu(1 - \rho)] \qquad (3)$$

λ (пакет/сек) — пакет ирэлтийн дундаж хурд

μ (пакет/сек) — зангилааны үйлчилгээний дундаж хурд (54 Мбит/с сувагт 1500 байтын пакетийн хувьд μ ≈ 4500 пакет/сек).

Системийн тогтвортой байдлын нөхцөл нь ρ < 1; ρ → 1 болох тусам дарааллын урт ба хүлээлтийн хугацаа хязгааргүй өснө. Төхөөрөмжийн хязгаарлагдмал буферийн нөхцлийг тусгахын тулд симуляцид дарааллын дээд хэмжээг K = 50 пакетаар тогтоосон бөгөөд энэ хязгаараас хэтэрсэн пакетыг хаяна. DDoS халдлагын үед λ огцом нэмэгдэж (λeff = λbase + bw × μ × 0.85), ρ-г ханалт руу хөтөлж, Wq хурдан өсснөөр сүлжээний тэсвэрлэх чадварт шууд сөрөг нөлөө үзүүлнэ

4. Давхарга дамнасан харилцан хамаарал: Гурван дэд загвар дараах байдлаар холбогдоно. PHY дохины алдагдал (1) link margin-ийг тодорхойлж, холбооны түвшний PDR-г (2) удирддаг. DDoS эрчим bw нэмэгдэх тусам PDR буурч, үр дүнтэй дамжуулалтын хурд суларснаар дарааллын ирэлт-үйлчилгээний тэнцвэрт (3) нөлөөлнө. Дээрх үзүүлэлтүүдийн нийлбэр нь {none, low, medium, high} × {10, 15, 20} × {3 seed} хослолоор бүрдсэн 36 туршилтын хувилбар дахь сүргийн тэсвэрлэх чадварын нэгдсэн хэмжүүр болно.

Д. Хэмжүүрүүдийн тодорхойлолт

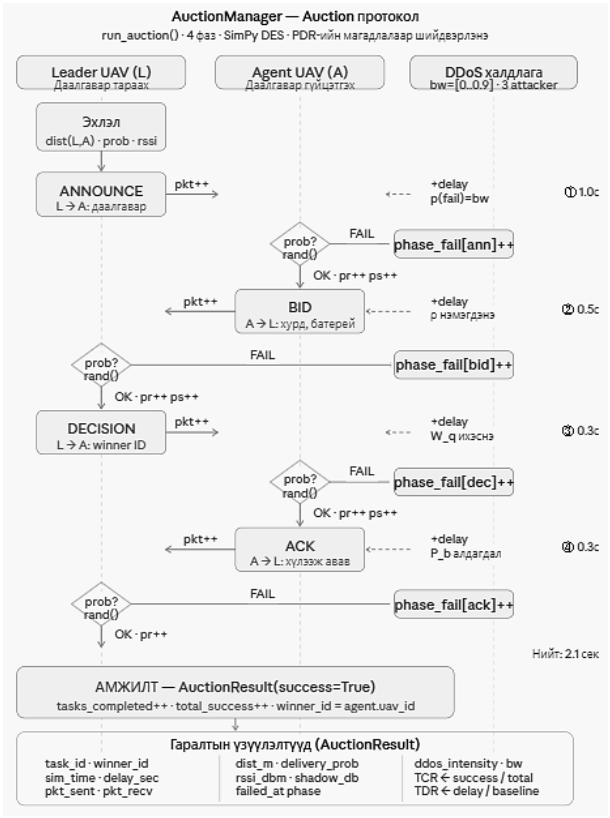
Судалгааны хүрээнд дроны сүргийн тэсвэртэй байдлыг үнэлэхийн тулд PHY-MAC-Transport-Application дөрвөн давхаргыг нэгтгэсэн давхарга дамнасан үнэлгээний дараах гурван хэмжүүрийг ашиглав.

ХҮСНЭГТ III. ДАВХАРГА ДАМНАСАН ТЭСВЭРТЭЙ БАЙДЛЫН ҮНЭЛГЭЭНИЙ ХЭМЖҮҮРҮҮД

Хэмжүүр	Томьёо	Тайлбар
TCR	*TCR = N_completed / N_total*	Даалгаврын амжилтын хувь. Хэрэглээний давхаргын нөлөөлөл.
TDR	*TDR = N_delayed / N_total*	Хоцорсон даалгаврын хувь. Сүлжээ/хандалтын давхаргын саатал.
SRT ★	*SRT = t_stable − t_leader_down*	Систем сэргэх хугацаа (сек) — шинэ хувь нэмэр.

★ Шинэ хувь нэмэр

DDoS халдлагын улмаас удирдагч зангилаа доголдсон тохиолдолд системийн тасралтгүй ажиллагааг хангахын тулд удирдагчийн нөхөн сэргэлтийн механизмыг хэрэгжүүлсэн. Удирдагчийн нөхөн сэргэлтийн механизм нь графын холбоосын зэрэглэлийн (Degree Centrality) алгоритмыг ашиглан сүлжээний хамгийн их холбоостой агентийг шинэ удирдагчаар динамикаар сонгодог.



3-р зураг. Өрсөлдөөнт санал өргөх протоколын дөрвөн үе шатын урсгал диаграм

E. Gauss-Markov хөдөлгөөний загвар

UAV -уудын хөдөлгөөнийг Gauss-Markov загвараар (α=0.75) симулцлана.

UAV шилжилт хөдөлгөөнийг загварчлахдаа нэгдүгээр зэргийн авторегрессив процесс (AR(1)) бүхий Гаусс–Марков (Gauss–Markov) загварыг ашиглав. Тухайн агшин дахь хурдны утгыг v(t) дараах томьёогоор тодорхойлно.

$$v(t) = \alpha \cdot v(t-1) + (1-\alpha) \cdot \bar{v} + \sqrt{(1-\alpha^2)} \cdot \sigma_v \cdot \varepsilon(t) \qquad (4)$$

Үүнд:

α: Автохамаарлын коэффициент (autocorrelation coefficient)

$\bar{v}$: Дундаж хурдны зорилтот утга (target mean speed)

σ_v: Хурдны стандарт хазайлт (standard deviation of speed)

$\varepsilon(t)$: N(0,1) стандарт хэвийн хуваарилалттай цагаан шуугиан (white noise)

Судалгаанд $\alpha = 0.75$ утгыг сонгон авсан бөгөөд энэ нь дроны хөдөлгөөний зан төлөвийг хамгийн сайн дүрслэх утга болохыг өмнөх судалгаагаар [12], [1]баталжээ.

Тус коэффициент нь дроны одоогийн хурд өмнөх хугацааны утгаасаа хэрхэн хамаарахыг илэрхийлэх бөгөөд α = 0.75 үед дроны одоогийн хурд нь өмнөх хурдны утгаас 75%-ийн хэмжээгээр шалтгаалж, үлдсэн 25% нь дундаж хурд болон санамсаргүй хэлбэлзлээс бүрддэг. Энэ нь бүрэн санамсаргүй алхам (α=0) болон шугаман жигд хөдөлгөөний (α=1) завсрын төлөв бөгөөд бодит нислэгийн зан чанарыг өндөр нарийвчлалтайгаар дуурайх нөхцөлийг бүрдүүлдэг.

Ж. Өрсөлдөөнт санал өргөх протоколын загварчлал

CBBA (Consensus-Based Bundle Algorithm) зарчимд суурилсан даалгавар хуваарилалт нь [13],[14]-д тодорхойлсон зөвшилцөлд суурилсан боодлын аукционы үндэс дээр дараах дөрвөн үе шатаар явагддаг бөгөөд үе шат бүрийн хугацааны утгыг утасгүй дроны харилцааны орчны хоцролтын дундаж утгад тулгуурлан тохируулсан [10].

1. Announce үе шат (1.0 сек): Удирдагч (Leader) агент шинэ даалгаврын мэдэгдлийг broadcast горимоор бүх агентад илгээнэ. DDoS халдлага энэ мэдэгдлийг хугацаа хэтрэлтэд оруулвал агентууд даалгаврын мэдээлэл хүлээн авч чадахгүй болно.

2. Bid үе шат (0.5 сек): Агент бүр өөрийн нөөц, байршлаас хамааран санал (bid) үүсгэж удирдагчид илгээнэ. DDoS flooding халдлагаар суваг хэт ачаалагдснаар bid мессежүүд удирдагчид хүрэлгүй алдагдана.

3. Decision үе шат (0.3 сек): Удирдагч хамгийн өндөр үнэлгээтэй санал тавьсан агентыг ялагчаар зарлаж, даалгавар хуваарилалтын шийдвэрийг илгээнэ. Энэ үе шат тасалдвал даалгавар хуваарилагдахгүй, TCR буурна.

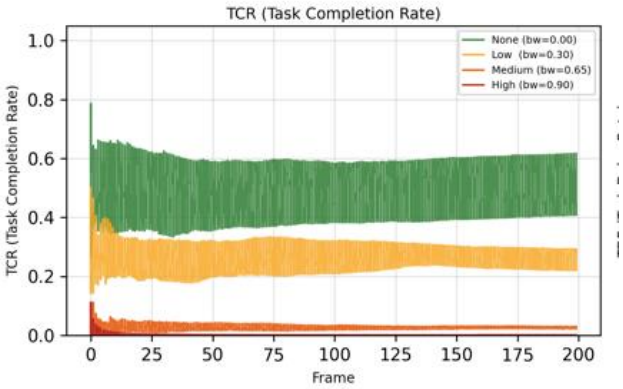
4. ACK үе шат (0.3 сек): Ялагч агент даалгаврыг хүлээн авснаа баталгаажуулах ACK мессеж илгээнэ. Хугацаа хэтрэлтэд орвол удирдагч даалгаврыг амжилтгүй гэж тэмдэглэж, TDR нэмэгдэнэ.

Эдгээр дөрвөн үе шатын аль нэгэнд DDoS халдлага хугацаа хэтрэлт үүсгэж болох бөгөөд амжилтгүй болоход үе шатын алдааны тоолуур нэмэгдэж, TCR буурч TDR өсдөг. Өрсөлдөөнт санал өргөх протоколын нийт хугацаа 2.1 секунд.

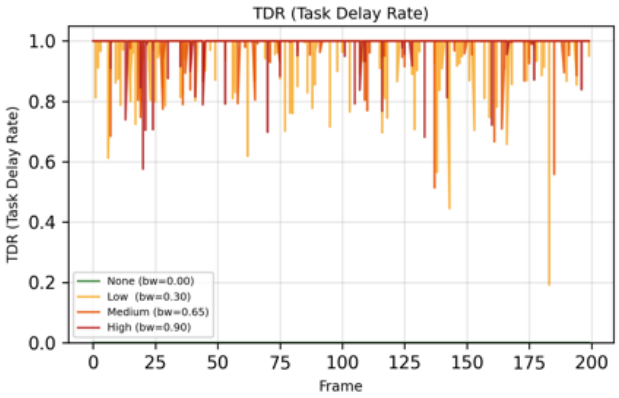
IV. ТУРШИЛТЫН УР ДҮН

А. Тэсвэрлэх чадварын үндсэн хэмжүүрүүдэд дэх халдлагын нөлөөлөл

4-р зурагт халдлагагүй нөхцөлд TCR ≈ 0.50–0.65 утгын хүрээнд харьцангуй тогтвортой байна. Халдлагын бага эрчмийн үед TCR ≈ 0.20–0.30 болж буурах бөгөөд дунд болон хүчтэй эрчмийн үед TCR ≈ 0.00–0.02 хүртэл огцом унаж, auction протоколын бүх үе шат симуляцийн нөхцөлд бүрэн доголддог болохыг харуулж байна.

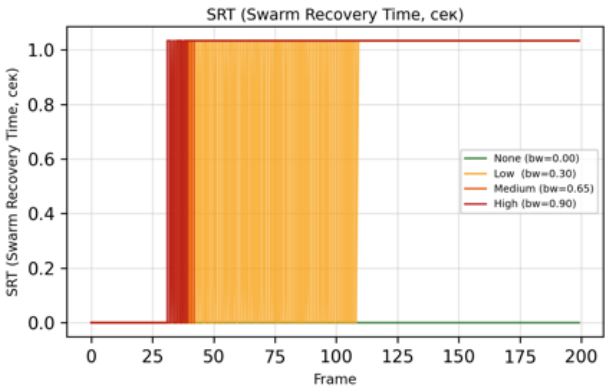


4-р зураг. Даалгавар биелүүлэлтийн харьцаа (TCR) — DDoS халдлагын эрчмийн нөлөөллийн цаг хугацааны динамик (200 фрэйм, гучин зургаан сценари)



5-р зураг. Даалгавар хоцрогдлын харьцаа (TDR) — DDoS халдлагын эрчмийн нөлөөллийн цаг хугацааны динамик (200 фрэйм, гучин зургаан сценари)

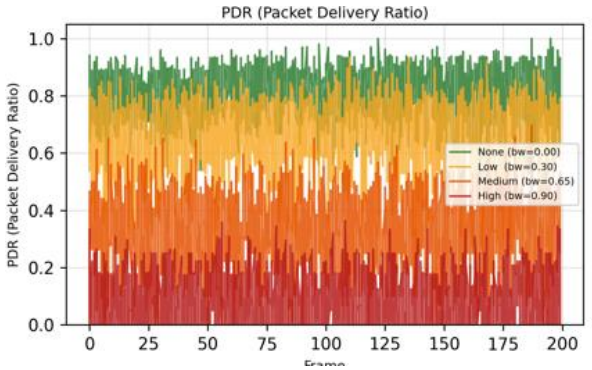
5-р зурагт үзүүлсэн TDR үзүүлэлт халдлагагүй үед 0.05–0.15 утгатай тогтвортой байв. Халдлагын бага (low) горимд 0.30–0.45 болж өссөн бол, дунд болон хүчтэй горимд 0.75–1.00 хүрч огцом нэмэгдэв. Энэ нь Bid болон ACK үе шатуудад хугацаа хэтрэлт (timeout) масс хэмжээгээр үүссэнтэй холбоотой бөгөөд халдлагын дээд түвшинд нийт даалгаврын дийлэнх нь хугацаандаа амжилгүй хоцорч байна.



6-р зураг. Сүргийн сэргэлтийн хугацаа (SRT) — DDoS халдлагын эрчмийн нөлөөллийн цаг хугацааны динамик (200 фрэйм, гучин зургаан сценари)

Халдлагагүй болон бага (bw=0.00–0.30) горимд LeaderFailover механизм идэвхжихгүй тул SRT утга 0.00 байна. Харин халдлагын дунд (bw=0.65) горимд 34-р фрэймээс эхлэн SRT≈1.033 сек болж, нийт 50 фрэймийн 16-д нь удирдагч солих үйл явц идэвхжсэн. Хүчтэй (bw=0.90)

горимд халдлагын нөлөө 32-р фрэймээс эхэлж, нийт 18 фрэймд SRT идэвхжин системийн тогтвортой байдал алдагдаж байгааг 6-р зурагт харууллаа.



7-р зураг. Пакет дамжуулалтын харьцаа (PDR) — DDoS халдлагын эрчмийн нөлөөллийн цаг хугацааны динамик (200 фрэйм, гучин зургаан сценари)

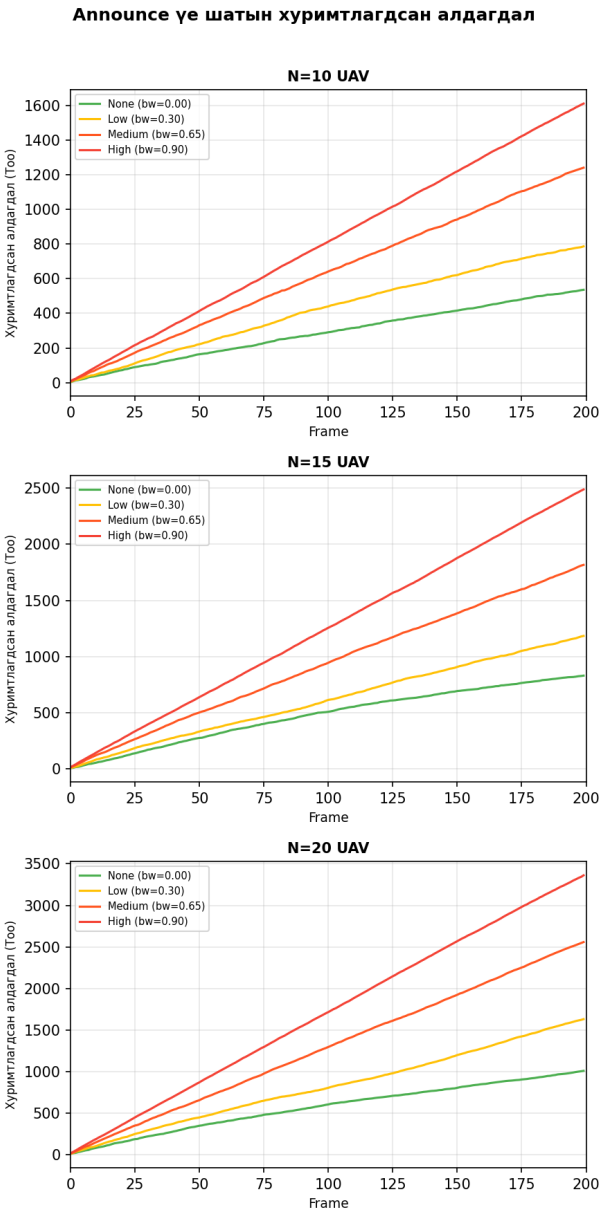
Халдлагагүй горимд PDR нэгнээс бага (0.826) байгаа нь хоёр шалтгааны нийлбэрийн үр дүн юм. Нэгдүгээрт, FSPL болон σ=8 дБ Rayleigh сүүдрийн нөлөөгөөр холбооны чадлын нөөц (link margin) санамсаргүй хэлбэлзэж, sigmoid функцийн бага магадлалын бүсэд орох үед зарим пакет алдагдана.

Хоёрдугаарт, аукционы Announce→Bid→Decision→ACK дөрвөн үе шатын cascade бүтэцтэй тул эхний үе шатын алдагдал дараагийн үе шатад хүрэх пакетийн тоог цөөлж, нийт PDR-г бууруулдаг. Иймд халдлагагүй горимд ч гэсэн сүлжээний физик нөхцөл болон протоколын бүтцийн хосолмол нөлөөгөөр PDR нэгнээс бага байх нь загварын зүй ёсны үр дүн мөн.

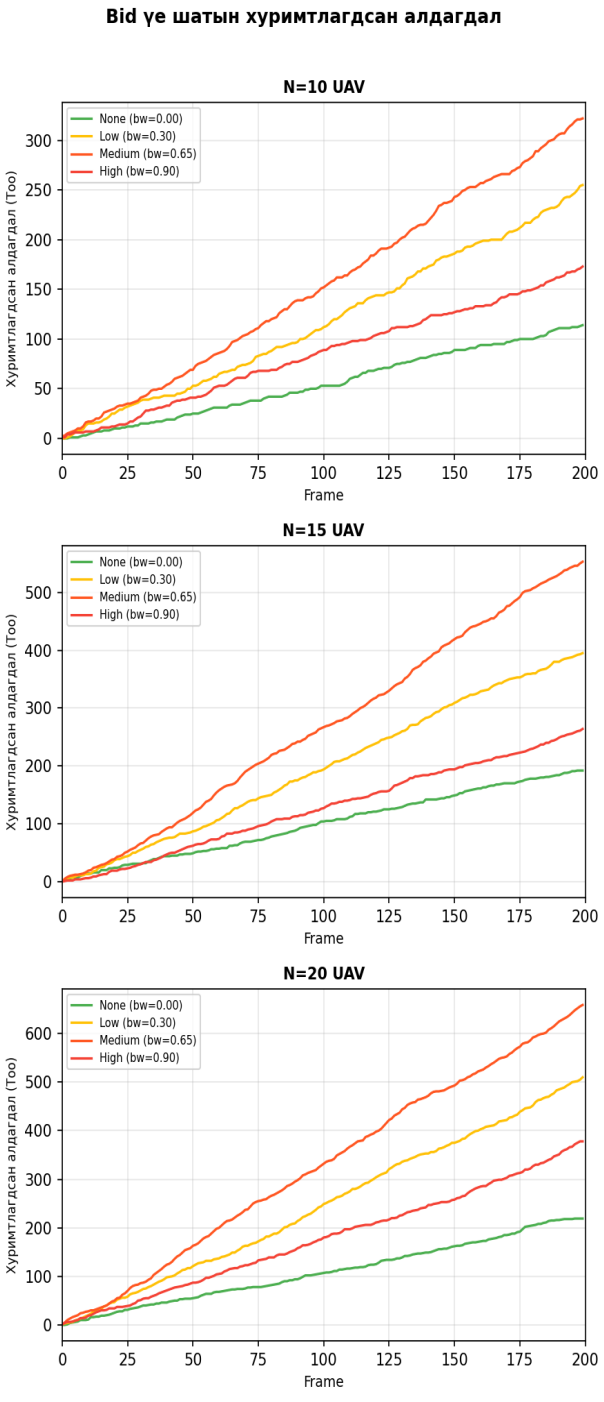
Б. Auction протоколд DDoS халдлагын эрчмийн нөлөөллийн үр дүн

8–11-р зураг өрсөлдөөнт санал өргөх протоколын Announce, Bid, Decision, ACK гэсэн дөрвөн үе шат бүрд DDoS халдлагын эрчмийн нөлөөгөөр хуримтлагдсан алдагдлын тоог графикаар харуулав.

8-р зураг N=10, 15, 20 UAV-ын тохиргоонд Announce үе шатад хуримтлагдсан алдагдлын цаг хугацааны динамикийг харуулав. Халдлагагүй горимд N=10-д 536, N=15-д 830, N=20-д 1,009 алдагдал хуримтлагдсан бол DDoS халдлагын хүчтэй горимд (bw=0.90) тус тус 1,611, 2,489, 3,361 болж 3.01x, 3.00x, 3.33x дахин өссөн байна. UAV-ын тоо нэмэгдэх тусам алдагдлын абсолют утга пропорциональ өсдөг боловч алдагдлын магадлал ~27–30% хооронд тогтвортой байгаа нь халдлагагүй горимд алдагдал зөвхөн дохионы чанараас хамаардагийг харуулж байна.

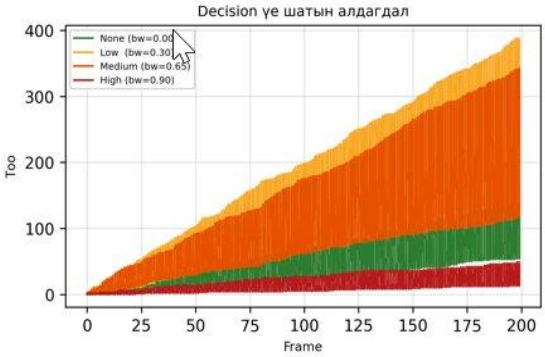


8-р зураг. Announce үе шатын хуримтлагдсан алдагдлын цаг хугацааны динамик (200 фрэйм, гучин зургаан сценари)



9-р зураг. Bid үе шатын хуримтлагдсан алдагдлын цаг хугацааны динамик (200 фрэйм, гучин зургаан сценари)

9-р зурагт N=10, 15, 20 UAV-ын тохиргоонд Bid үе шатад хуримтлагдсан алдагдлын цаг хугацааны динамикийг харуулав. Халдлагагүй горимд N=10-д 114, N=15-д 192, N=20-д 219 алдагдал хуримтлагдсан бол дунд горимд (bw=0.65) тус тус 322, 553, 659 болж хамгийн өндөр утгад хүрсэн байна. Харин DDoS халдлагын хүчтэй горимд (bw=0.90) N=10-д 173, N=15-д 264, N=20-д 378 болж дунд горимоос буурсан нь сувгийн бүрэн дүүрэлтээс үүдэн Announce үе шатад аль хэдийн их алдагдал гарснаар Bid үе шатад хүрч ирэх аукцион цөөрдөгтэй холбоотой юм. Энэ үзэгдэл нь PHY-MAC-Transport-Application давхарга дамнасан cascade доголдлын механизмыг тодорхой илэрхийлж байна.



10-р зураг. Decision үе шатын хуримтлагдсан алдагдлын цаг хугацааны динамик (200 фрэйм, гучин зургаан сценари)

Халдлагагүй горимд шийдвэр гаргах (Decision) үе шат дахь багцын алдагдал 625-аар хязгаарлагдсан бол халдлагын бага горимд (bw=0.30) энэ үзүүлэлт 1,310-д хүрч, хамгийн өндөр алдагдлын харьцааг үзүүлэв. Харин дунд горимд 869, хүчтэй горимд 155 болтлоо огцом буурсан байна.

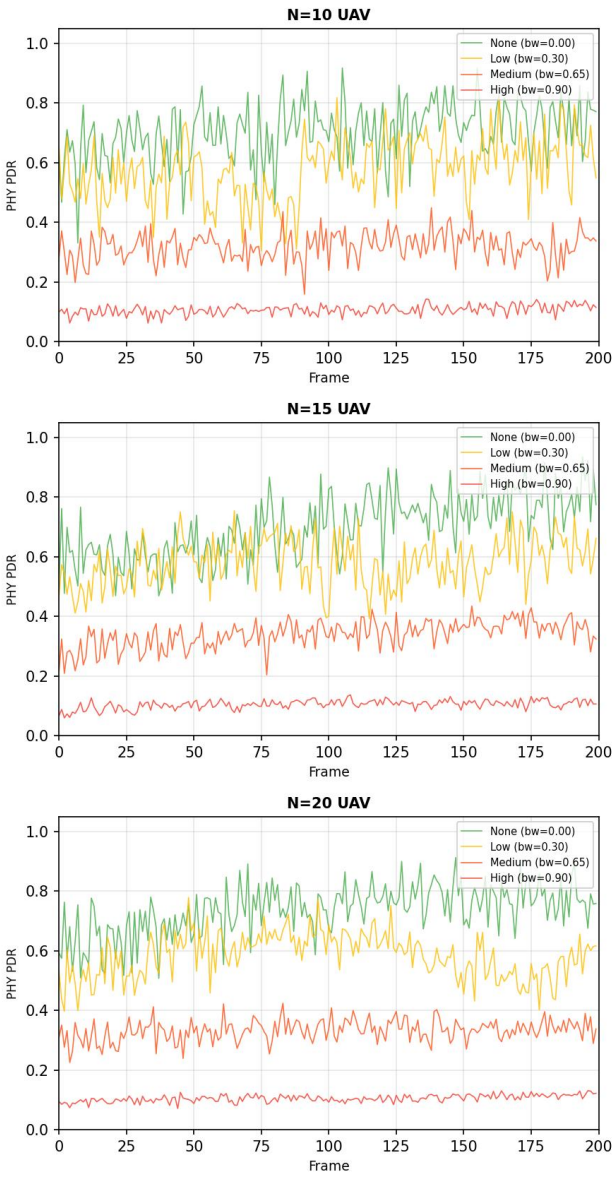


11-р зураг. ACK үе шатын хуримтлагдсан алдагдлын цаг хугацааны динамик (200 фрэйм, гучин зургаан сценари)

11-р зурагт Халдлагагүй горимд баталгаажуулах мэдээллийн (ACK) алдагдал нийт 307 буюу нийт алдагдлын 4.6%-ийг эзэлж байна. Сүлжээний зурвасын ачаалал бага горимд (bw=0.30) байх үед энэ үзүүлэлт 997 болж, 224.8%-иар огцом өсөж оргилдоо хүрсэн нь гурван үе шатыг амжилттай давсан аукционы процесс сувгийн ашиглалтын коэффициент (ρ=0.255)-оос хамааран баталгаажуулах мэдээллийн хугацаа хэтрэлт нэмэгдэж байгааг харуулж байна. Харин дунд горимд (bw=0.65) 464, хүчтэй горимд (bw=0.90) 35 болж, нийт алдагдлын 0.2% хүртэл буурсан байна. Хүчтэй халдлагын горимд 550 аукционаас нэг нь ч амжилттай болоогүй.

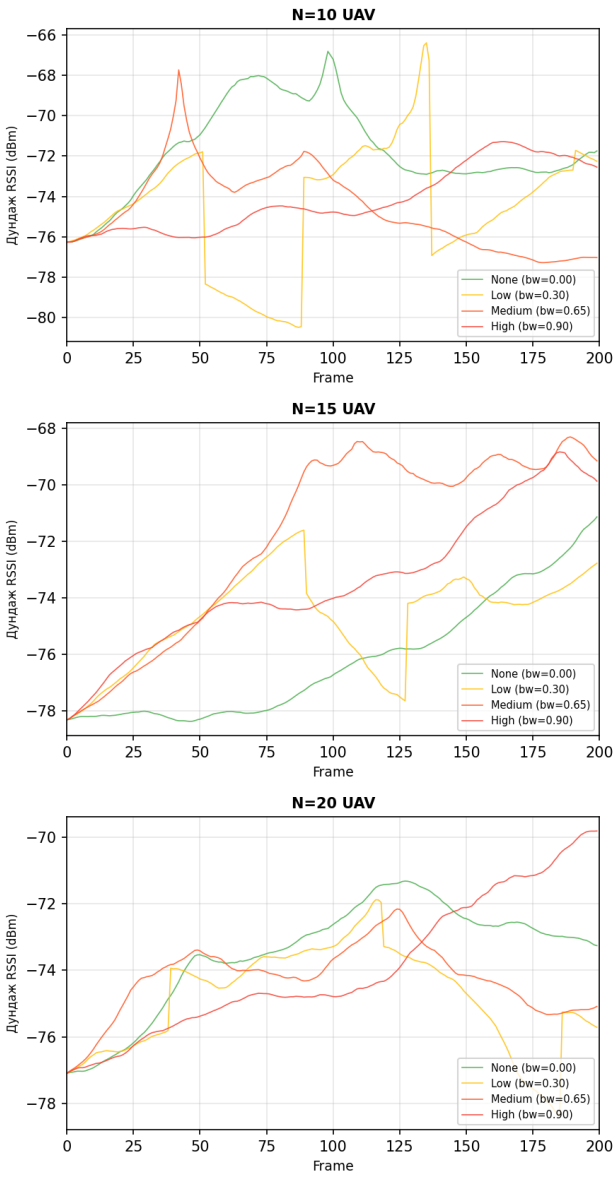
B. PHY-MAC-Transport-Application давхарга дамнасан халдлагын нөлөөллийн шинжилгээ

12–17-р зурагт PHY, MAC, Transport давхарга бүрийн хэмжүүрүүд DDoS халдлагын эрчмийн шатласан нөлөөллийг харуулав.



12-р зураг. PHY давхаргын пакет дамжуулалтын харьцаа (PHY PDR) — DDoS халдлагын эрчмийн нөлөөллийн цаг хугацааны динамик (200 фрэйм, гучин зургаан сценари)

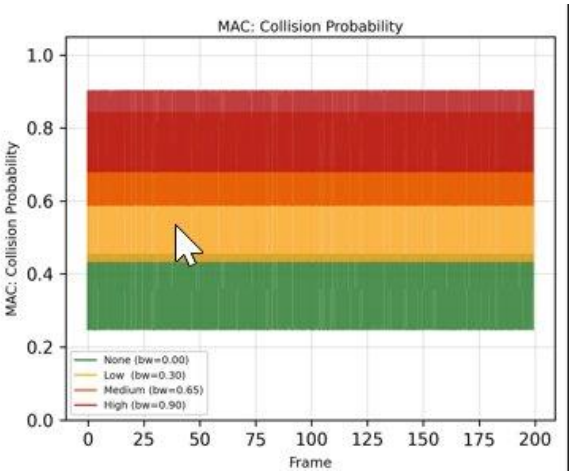
12-р зурагт N=10, 15, 20 UAV-ын тохиргоонд PHY давхаргын пакет дамжуулалтын харьцааны цаг хугацааны динамикийн графикаас харахад халдлагагүй горимд N=10-д дундаж 0.705, N=15-д 0.699, N=20-д 0.736 байсан бол DDoS халдлагын хүчтэй горимд (bw=0.90) тус тус 0.107, 0.104, 0.105 болж ~6.6–6.8 дахин буурсан байна. Бүх N-ийн тохиргоонд PHY PDR-ийн дундаж утга ойролцоо байгаа нь физик давхаргын алдагдал дроны сүргийн тооноос хамаарахгүйгээр зөвхөн дохионы чанар болон DDoS эрчмээс хамаардагийг харуулж байна.



13-р зураг. PHY давхаргын дундаж RSSI — DDoS халдлагын эрчмийн нөлөөллийн цаг хугацааны динамик (200 фрэйм, гучин зургаан сценари)

13-р зурагт үзүүлсэн N=10, 15, 20 дроны сүргийн тохиргоонд PHY давхаргын дундаж RSSI-ийн цаг хугацааны динамикийн графикаас харвал бүх тохиргоонд RSSI -66.4-аас -80.5 дБм хооронд хэлбэлзэж, дундаж утга -71.7-аас -76.2 дБм байна.

DDoS халдлагын эрч нэмэгдэхэд RSSI мэдэгдэхүйц өөрчлөгдөөгүй бөгөөд горим бүрийн дундаж утга 2–3 дБм-ийн хооронд ялгаатай байгаа нь UDP flooding халдлага физик давхаргын дохионы хүчинд бус MAC болон Transport давхаргын нөөцөд нөлөөлдөгийг баталж байна.



14-р зураг. MAC давхаргын мөргөлдөөний магадлал (Collision Probability) — DDoS халдлагын эрчмийн нөлөөллийн цаг хугацааны динамик (200 фрэйм, гучин зургаан сценари)

14-р зурагт халдлагагүй горимд багцын мөргөлдөөний магадлал (collision probability) 0.295 утгатай тогтвортой байгаа бол халдлагын бага горимд 0.487, дунд горимд 0.660, хүчтэй горимд 0.745 хүртэл шатлан өссөн байна. IEEE 802.11 CSMA/CA протоколд Bianchi [16]-н загварт суурилсан нэг слотын хугацаанд станц дамжуулах магадлал τ болон мөргөлдөөний магадлал Pc-ийг дараах байдлаар тооцоолно.

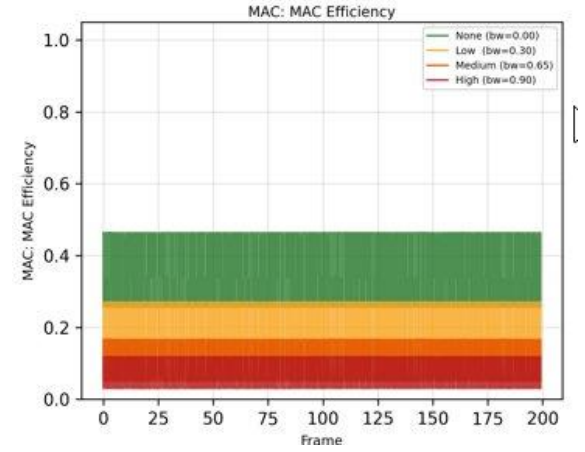
$$\tau = 2 / (CWmin + 1) \quad (5)$$

$$Pc = 1 - (1 - \tau)n-1 \quad (6)$$

энд n — идэвхтэй станцын тоо (дроны тоо),
$CWmin$ — хамгийн бага өрсөлдөөний цонх.

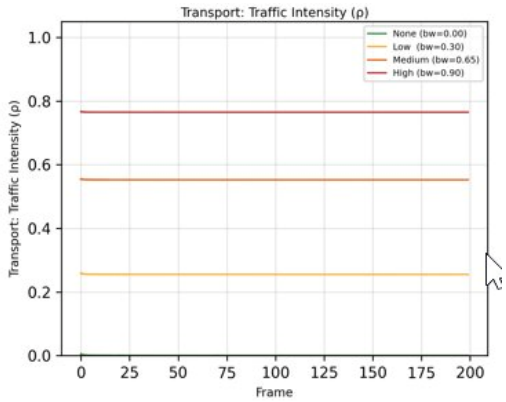
Үр дүнгийн шинжилгээ хийвэл CWmin = 32, n = 10 нөхцөлд онолын хувьд Pc =0.426 байгаа боловч симуляциар Pc = 0.295 гарсан байна. Энэхүү зөрүү нь физик давхаргын пакетын алдагдал MAC давхаргад нөлөөлж байгааг харуулж байна.

Дан MAC түвшний загвар нь физик давхаргын нөлөөг орхигдуулдаг бол энэ бүтээлийн хүрээнд системтэйгээр нэгтгэн тооцох cross-layer загвар нь эдгээр алдагдлыг нэгтгэн тооцсон.



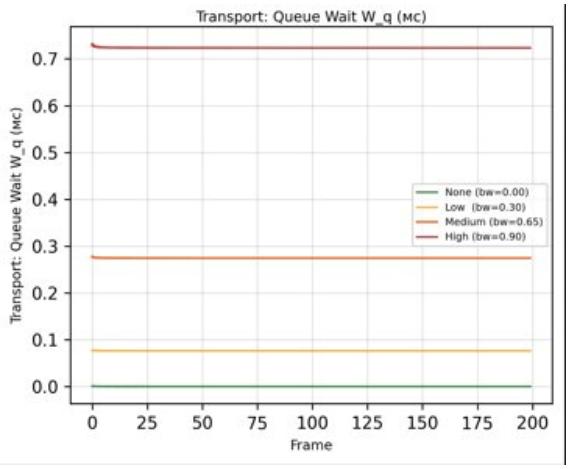
15-р зураг. MAC давхаргын үр ашиг (MAC Efficiency) — DDoS халдлагын эрчмийн нөлөөллийн цаг хугацааны динамик (200 фрэйм, гучин зургаан сценари)

15-р зурагт халдлагагүй горимд MAC үр ашиг 0.406 байна бол бага горимд 0.232, дунд горимд 0.128, хүчтэй горимд 0.089 болж тус тус шатлан буурсан нь DDoS эрч нэмэгдэхэд CSMA/CA мөргөлдөөний огцом өссөнөөс сувгийн ашиглалт буурч пакет дамжуулалт улам бүр доголдож байгааг илтгэнэ.



16-р зураг. Transport давхаргын трафикийн эрчмэл (ρ) — DDoS халдлагын эрчмийн нөлөөллийн цаг хугацааны динамик (200 фрэйм, гучин зургаан сценари)

16-р зурагт үзүүлсэн халдлагагүй горимд сувгийн ашиглалтын коэффициент ρ=0.000 байгаа нь M/M/1 дараалын загвар дахь ачаалал хамгийн бага түвшинд байгааг илтгэнэ. Харин халдлагын бага горимд ρ=0.255, дунд горимд ρ=0.553, хүчтэй горимд ρ=0.765 хүртэл тус тус өссөн байна.



17-р зураг. Transport давхаргын дарааллын хүлээлтийн хугацаа (W_q) — DDoS халдлагын эрчмийн нөлөөллийн цаг хугацааны динамик (200 фрэйм, гучин зургаан сценари)

17-р зурагт халдлагагүй горимд W_q≈0.000 ms — хүлээлт бараг 0 байгаа бол бага горимд W_q=0.076 ms, дунд горимд W_q=0.275 ms, хүчтэй горимд W_q=0.724 ms хүртэл өсч M/M/1 дараалын Литтльсийн хууль ρ/[μ(1−ρ)] нь ρ=0.765 үед огцом өсснөөр дараалалын хүлээлт нэмэгдсэн болохыг илтгэнэ.

V. ДҮГНЭЛТ БА ЦААШДЫН АЖИЛ

Энэ судалгааны хүрээнд дараах гурван үндсэн үр дүнд хүрэв. Үүнд: (1) 2.4 ГГц-ийн давтамжийн зурваст ажилладаг өрсөлдөөнт санал өргөх зарчимд суурилсан UAV сүргийн DDoS халдлагад тэсвэртэй байдлыг PHY-MAC-Transport-Application дөрвөн давхарга дамнасан нэгдсэн загвараар үнэлэх тоон хэмжүүрийн систем боловсруулсан; (2) сүргийн тэсвэртэй байдлын шинэ хэмжүүр болох SRT (Swarm Recovery Time)-ийг тодорхойлж өгөгдлийн багцад нэвтрүүлсэн; (3) машин

сургалтын алгоритмд зориулсан 36 сценари бүхий, 41 баганатай шошгожуулсан CSV өгөгдлийн багцыг үүсгэсэн.

Туршилтын үр дүнгээс үзэхэд дунд горимд (зурвасын ачааллын хувь bw=0.65) PDR=32.7% болж бага горимын (bw=0.30, PDR=64.2%)-тай харьцуулахад 49.2%-иар харьцангуй буурсан нь DDoS эрч нэмэгдэх тусам системийн найдвартай ажиллагаа огцом доройтдог болохыг илтгэнэ.

TDR хэмжүүр халдлагагүй горимд 0.000 байснаа бага горимоос эхлэн 0.995 хүртэл огцом өсч бүх даалгавар хугацаа хэтрэлтэд орсон нь протоколын дараалсан доголдлыг тодорхойлсон. Хүчтэй горимд (bw=0.90) TCR=0.000 буюу 550 аукционы нэг ч амжилтгүй болсон нь сүргийн даалгавар хуваарилалт бүхэлдээ доголдсоны нотолгоо болно.

Загварчлалын үр дүнг ns-3 симулятор болон бодит UAV-ийн туршилтаар харьцуулан баталгаажуулах нь цаашдын судалгааны нэн тэргүүний зорилт байх болно.

Энэхүү судалгааны хүрээнд үүсгэсэн өгөгдлийн сан нь 36 сценари, 41 шинж чанар бүхий нийт 200 бичлэгийг агуулах бөгөөд симуляцийн параметрийн орон зайг (4 DDoS түвшин × 3 UAV тоо × 3 санамсаргүй seed) бүрэн хамарсан боловч машин сургалтын загварт ашиглахад статистикийн хувьд хязгаарлагдмал хэвээр байна. Ялангуяа гүн сургалтын загварууд (Deep Learning)-д сургалтын болон туршилтын олонлогийг хуваах үед загварын ерөнхийлөх чадварт нөлөөлөх эрсдэлтэй тул цаашдын судалгааны хүрээнд seed болон сценарийн тоог нэмэгдүүлэх замаар өгөгдлийн санг өргөтгөхийг зорьж байна. Энэхүү өгөгдлийн сан нь ML сургалтын эцсийн үр дүн биш харин UAV сүргийн DDoS тэсвэртэй байдлын судалгааны анхдагч (benchmark) суурь болохоор зориулагдсан болно.

Мөн дараагийн судалгааны хүрээнд TCR, TDR болон SRT хэмжүүрүүдийн хоорондын хамаарлыг ANOVA болон t-test статистик шинжилгээгээр баталгаажуулсны дараа LSTM Autoencoder-т суурилсан хэвийн бус үйлдэл илрүүлэлтийн загварыг хэрэгжүүлж, ablation study (LSTM AE vs CNN vs Vanilla AE)-ээр загваруудын гүйцэтгэлийг харьцуулан үнэлэхээр төлөвлөсөн.

ТАЛАРХАЛ

Судалгааны ажлыг чиглүүлж, мэргэжлийн удирдлагаар хангасан удирдагч багш, доктор Ц.Энхтөр болон туслалцаа дэмжлэг үзүүлсэн ШУТИС-ийн МХТС-ийн Кибер аюулгүй байдлын тэнхимийн багш нарт талархал илэрхийлье.

ACKNOWLEDGMENTS

This research was supported by KOICA (Korea International Cooperation Agency) through the "Capacity Building Project for the School of Information and Communication Technology at Mongolian University of Science and Technology" (Contract No. P2019-00124).

НОМ ЗҮЙ

[1] Liu, J. et al., "A Survey on Security of UAV Swarm Systems," ACM Comput. Surv., 2024.

[2] Zhang, W. & Chen, L., "Two-Stage Auction-Based Task Allocation in UAV Swarms," IJMLC, vol. 14, no. 3, 2024.

[3] Torres, M. et al., "Resilient Task Allocation for UAV Swarms Using PSO-ILP," Drones, vol. 9, no. 1, 2025.

[4] Nguyen, T. & Kim, S., "Swarm Benefit Optimization via Cooperative Auctions," IEEE Access, 2024.

[5] Y. Wang, Z. Li, and J. Smith, "Consensus-Based Decentralized Task Allocation for Multi-Agent Systems and Simultaneous Multi-Agent Tasks," IEEE Robotics and Automation Letters, vol. 7, no. 4, pp. 11095–11102, Oct. 2022.

[6] Ahmed, F. et al., "LSTM Autoencoder for DDoS Attack Detection in IoT Networks," IEEE Internet of Things Journal, vol. 10, no. 14, pp. 12453–12465, 2023.

[7] Nguyen, H. & Park, J., "LSTM-AE Anomaly Detection for Network Traffic," IEEE TNNLS, vol. 31, no. 8, 2020.

[8] Zhao, Y. et al., "Optimized AE-LSTM for Network Anomaly Detection," Comput. Secur., 2025.

[9] Choi, H.-L., Brunet, L. & How, J. P., "Consensus-Based Decentralized Auctions for Robust Task Allocation," IEEE Trans. Robot., vol. 25, no. 4, pp. 912–926, 2009.

[10] Vugrin, E. D., et al. "A resilience assessment framework for infrastructure and economic systems." (2011).

[11] ITU-R, "Recommendation P.1238-9: Propagation data and prediction methods for the planning of indoor radiocommunication systems," ITU-R, Geneva, Switzerland, Jun. 2017.

[12] D. S. Lakew et al., "Routing in Flying Ad Hoc Networks: A Comprehensive Survey," IEEE Commun. Surveys Tuts., vol. 22, no. 2, pp. 1071–1120, 2020.

[13] B. Liang and Z. J. Haas, "Predictive distance-based mobility management for multidimensional PCS networks," IEEE/ACM Trans. Netw., vol. 11, no. 5, pp. 718–732, 2003.

[14] J. D. M. M. Biomo, T. Kunz, and M. St-Hilaire, "An enhanced Gauss-Markov mobility model for simulations of unmanned aerial ad hoc networks," 7th IFIP WMNC, 2014.

[15] L. Kleinrock, Queueing Systems, Volume 1: Theory. New York, NY, USA: Wiley-Interscience, 1975.

[16] Bianchi, "Performance analysis of the IEEE 802.11 distributed coordination function," IEEE J. Sel. Areas Commun., vol. 18, no. 3, pp. 535–547, Mar. 2000.

Тайлбарлагдахуйц Хиймэл Оюуныг Халдлага Илрүүлэх Системд Хэрэглэх нь

Б.Өсөхбаяр, Б.Энхбаяр*, Н.Угтахбаяр,
Мэдээллийн Технологи, Электроникийн Сургууль, Монгол Улсын Их Сургууль
Электроник, Холбооны Инженерчлэлийн Тэнхим
*Corresponding author: enkhbayar.ba@num.edu.mn

Хураангуй—Орчин үеийн кибер аюулгүй байдлын салбарт сүлжээний халдлагын нарийн төвөгтэй байдал нь уламжлалт Халдлага Илрүүлэх Систем (IDS)-ийн үр ашгийг бууруулж байна. Хиймэл оюунд суурилсан халдлага илрүүлэх системийн шийдвэр гаргалтын процесс нь тодорхойгүй байдаг тул энэхүү системийг өргөн хэрэглээнд нэвтрүүлэхэд томоохон саад болж байна. Иймээс Тайлбарлагдахуйц Хиймэл Оюун (XAI)-д суурилсан Халдлага Илрүүлэх Системийн тайлбарлагдах чадварыг сайжруулж, илүүдэл онцлог шинжүүдийг багасгаж, ил тод, үр ашигтай систем боловсруулахаар зорьсон. Үүний тулд онцлог шинж сонголтын олон аргыг машин сургалтын загваруудтай хослуулан. Мөн XAI тайлбарлах аргуудын тайлбарыг үнэлэх үр дүнтэй аргыг хэрэгжүүлсэн. Туршилтын үр дүнд mRMR онцлог шинж сонгох арга болон XGBoost алгоритмуудын хослол нь өндөр гүйцэтгэл (99.99% нарийвчлал) үзүүлсэн. Мөн TreeSHAP болон Permutation Importance аргууд нь бусад аргуудтай харьцуулахад тайлбарын чанар болон тогтвортой байдлыг хангаж байгааг харуулсан. Энэ ажил нь онцлог шинж сонголт болон тайлбарлагдахуйц хиймэл оюуны аргачлалаар Халдлага Илрүүлэх Систем (IDS)-ийн гүйцэтгэл болон тайлбарлагдах чадварыг сайжруулсан.

Түлхүүр үгс—Кибер аюулгүй байдал, Халдлага Илрүүлэх Систем (IDS), Тайлбарлагдахуйц хиймэл оюун (XAI), Онцлог шинж сонголт, Машин сургалт, CIC-IDS-2017

I. Удиртгал

Цахим гэмт хэрэгтнүүдийн ашиглаж буй халдлагын арга техник улам бүр нарийсаж, тасралтгүй хөгжиж байна. Иймд компьютерын сүлжээний аюулгүй байдлыг хангах нь өнөө цагийн хамгийн чухал асуудлуудын нэг болсон. Сүлжээний төгсгөлийн төхөөрөмжийг зөвшөөрөлгүй, зүй бусаар ашиглах, түүнд хөнөөлт үйлдэл хийх оролдлогыг илрүүлэх нь халдлага илрүүлэх системийн (Intrusion Detection System, IDS) зорилго юм [1], [2]. IDS нь сүлжээний аюулгүй байдлыг хангах чухал механизм боловч халдлагын хөгжингүй, хувьсамтгай шинж чанар зэрэг хүчин зүйлс нь уламжлалт аргуудын үр ашгийг бууруулж байна [3].

IDS-д машин сургалтын аргуудыг ашиглах үед үүсдэг гол асуудын нэг нь загварын шийдвэр гаргалтын ил тод бус байдал болон онцлог шинжүүдийн хэт их хэмжээ буюу өгөгдлийн олон хэмжээст байдал юм. Олон хэмжээст өгөгдөл нь илүүдэл болон хамааралтай шинжүүдийг агуулдаг бөгөөд энэ нь загварын гүйцэтгэлд сөргөөр нөлөөлөхөөс гадна тайлбарлалтын чанарыг бууруулдаг [3]–[5]. Нөгөө талаас, халдлагын хэв шинж байнга хувьсаж байдаг тул загваруудын тогтвортой байдал болон ерөнхийлөх чадвар (generalization) сорилт болдог [6], [7]. Онцлог шинж сонгох нь өгөгдлийн олон хэмжээст байдлыг багасгах ба халдлага илрүүлэх системийн найдвартай ажиллагааг хангах үндсэн нөхцөл болдог [3]–[5], [8], [9]. Онцлог шинж сонгох Filter, Wrapper , Embedded, Hybrid, XAI-based төрлийн аргуудыг дурдсан байдаг [3], [10]. Эдгээрийг бүгдийг туршсанаар тухайн машин сургалтын алгоритмын онцлогт тохирсон өгөгдлийн олонлогийг тодорхойлох боломжтой болно. Ингэснээр хамгийн тогтвортой бөгөөд өндөр нарийвчлал үзүүлэх шийдлийг олох боломжтой [11].

Сүүлийн жилүүдэд Random Forest, XGBoost, LightGBM зэрэг машин сургалтын ансамбль загварууд нь халдлагыг өндөр нарийвчлалтай илрүүлэх боломжтойг харуулсан [3], [10]–[12]. Ялангуяа XGBoost загвар нь нарийн төвөгтэй халдлагын хэв шинжийг илрүүлэхэд онцгой давуу талтай бөгөөд нөөцийн хязгаарлагдмал орчинд ч тогтвортой ажиллах чадвартай [13]. Гэвч эдгээр загварууд нь “хар хайрцаг” шинжтэй тул шийдвэр гаргалтын процессыг тайлбарлах боломж хязгаарлагдмал бөгөөд энэ нь практик хэрэглээнд итгэлцэл, ил тод байдлыг бууруулах сул талтай [11]. Иймд халдлага илрүүлэх системд тайлбарлагдахуйц хиймэл оюуныг ашиглах хэрэгцээ шаардлага бий болж байна. Хэдийгээр XAI аргууд нь загварын шийдвэрийг тайлбарлах боломжийг олгодог ч эдгээр тайлбаруудын чанар, найдвартай байдлыг системтэйгээр үнэлэх асуудал хангалттай шийдэгдээгүй байна. Тайлбарлагдахуйц хиймэл оюунд ашиглагдаж буй SHAP, LIME зэрэг тайлбарлах аргууд нь загварын шийдвэрийг ойлгоход тустай боловч тооцооллын өндөр зардал, тайлбарын тогтворгүй байдал, онцлог шинжүүдийн хамаарлыг үл тоомсорлох, шалтгаан-үр дагаврын холбоог биш зөвхөн хамаарлыг тайлбарлах, мөн локал ба глобал тайлбаруудын зөрчил зэрэг сул талуудтай байдаг бөгөөд үүнээс гадна бодит цагийн орчинд хэрэглэхэд өргөтгөх боломжтой байдал болон найдвартай байдлын асуудал үүсгэдэг [3]–[6], [10]–[12]. Иймээс өндөр нарийвчлалтай, бодит цагийн халдлага илрүүлэх чадвартай дараагийн

үеийн халдлага илрүүлэх системийн шийдлүүдийн тулгамдсан асуудал болж байна [13]. Уламжлалт гүйцэтгэлийн үзүүлэлтүүдээс гадна бид тайлбарын чанарыг sparsity, stability, sufficiency, comprehensiveness, rank correlation (RankCorr), мөн тооцооллын саатал (latency) зэрэг шалгуур үзүүлэлтүүдээр үнэлсэн. Энэ аргачлал нь XAI аргуудаас гарсан тайлбарууд бодит орчинд хэрэгжихүйц шийдвэр гаргалтад ашиглагдах боломжтой эсэхийг илүү нарийн тодорхойлох боломжийг олгодог [13].

Иймээс өндөр гүйцэтгэлийг хадгалахтай зэрэгцэн тайлбарлагдах чадварыг сайжруулсан халдлага илрүүлэх систем боловсруулах шаардлага тулгарч байна. Энэхүү судалгаа нь оновчтой онцлог шинж сонголт, өндөр гүйцэтгэл бүхий тайлбарлагдахуйц хиймэл оюуныг хослуулан IDS-ийн гүйцэтгэл ба тайлбарлах чадварыг зэрэг сайжруулахад шинжлэх ухааны үндэслэлтэй иж бүрэн үнэлгээний хэмжүүрүүдийг ашиглан үнэлэхийг зорьсон [14]. Ингэснээр кибер аюулгүй байдлын салбарт чухал хувь нэмэр оруулж, мэргэжилтнүүдэд зөвхөн халдлагыг илрүүлэхээс гадна тухайн жишээ яагаад халдлага гэж ангилагдсан болохыг найдвартай тайлбарлах боломжийг олгож, системийн ил тод байдлыг нэмэгдүүлэх чухал ач холбогдолтой [6].

II. Судлагдсан байдал

Орчин үеийн судалгааны гол чиг хандлага нь машин сургалтын загварууд ыг оновчтой онцлог шинж сонголтын (feature selection) аргуудтай хослуулах замаар халдлага илрүүлэлтийн нарийвчлалыг дээд түвшинд хүргэхэд чиглэж байна. Онцлог шинж сонголтын аргууд нь загварын сургалтын хугацааг багасгах, өндөр хэмжээс бүхий өгөгдөл дээрх таамаглалын нарийвчлалыг сайжруулах зорилгоор өргөн хэрэглэгддэг [4]. Онцлог шинж сонголтын хүрээнд Mutual Information (MI), mRMR, Boruta, болон GA Wrapper зэрэг аргууд нь өгөгдлийн олон хэмжээст байдлыг бууруулж, загварын сургалтын хугацааг багасгах, илүүдэл шуугианыг устгах замаар таамаглалын нарийвчлалыг 3-5% хүртэл сайжруулж буйг олон судалгаа баталсан [3]. Мөн судалгаанууд IDS-ийн зөвхөн гүйцэтгэлийг сайжруулахаас гадна тайлбарлагдах чадварыг нэмэгдүүлэхэд чиглэсэн байна [4]. Түүнчлэн, тайлбарлагдахуйц хиймэл оюуны (XAI) оноонд суурилсан онцлог шинж сонголтын аргууд (SHAP-based FS) нь загварт хамгийн чухал нөлөө үзүүлж буй хүчин зүйлсийг (жишээ нь: Destination Port, TCP Flags) шинжлэх ухааны үндэслэлтэйгээр тодорхойлоход өндөр үр дүнтэй байгааг онцолж байна [3], [5].

Зарим судалгаанд Random Forest, XGBoost, LightGBM загваруудыг ашиглан CIC-IDS-2017 зэрэг өгөгдөл дээр өндөр нарийвчлалтай илрүүлэлт хийсэн [8]. Загварчлалын хувьд Random Forest (RF), XGBoost, LightGBM зэрэг ансамбль аргууд нь халдлагын нарийн төвөгтэй хэв шинжийг (жишээ нь: Ransomware, DoS) илрүүлэхэд стандарт шийдэл болж байна [9]. Эдгээр загварууд нь мөн чанартаа хар хайрцаг шинж чанартай байдаг нь аюулгүй байдлын шинжээчдийн итгэлийг хүлээхэд томоохон саад болдог. Үүнийг шийдвэрлэхийн тулд TreeSHAP, KernelSHAP болон LIME зэрэг XAI аргуудыг ашиглан загварын шийдвэр гаргах логикийг ерөнхий (global) болон хэсэгчилсэн (local) түвшинд тайлбарлах туршилтууд хийгдсээр байна. TreeSHAP болон KernelSHAP зэрэг XAI аргуудыг ашиглан загварын шийдвэрийг тайлбарлах боломжийг судалсан [15].

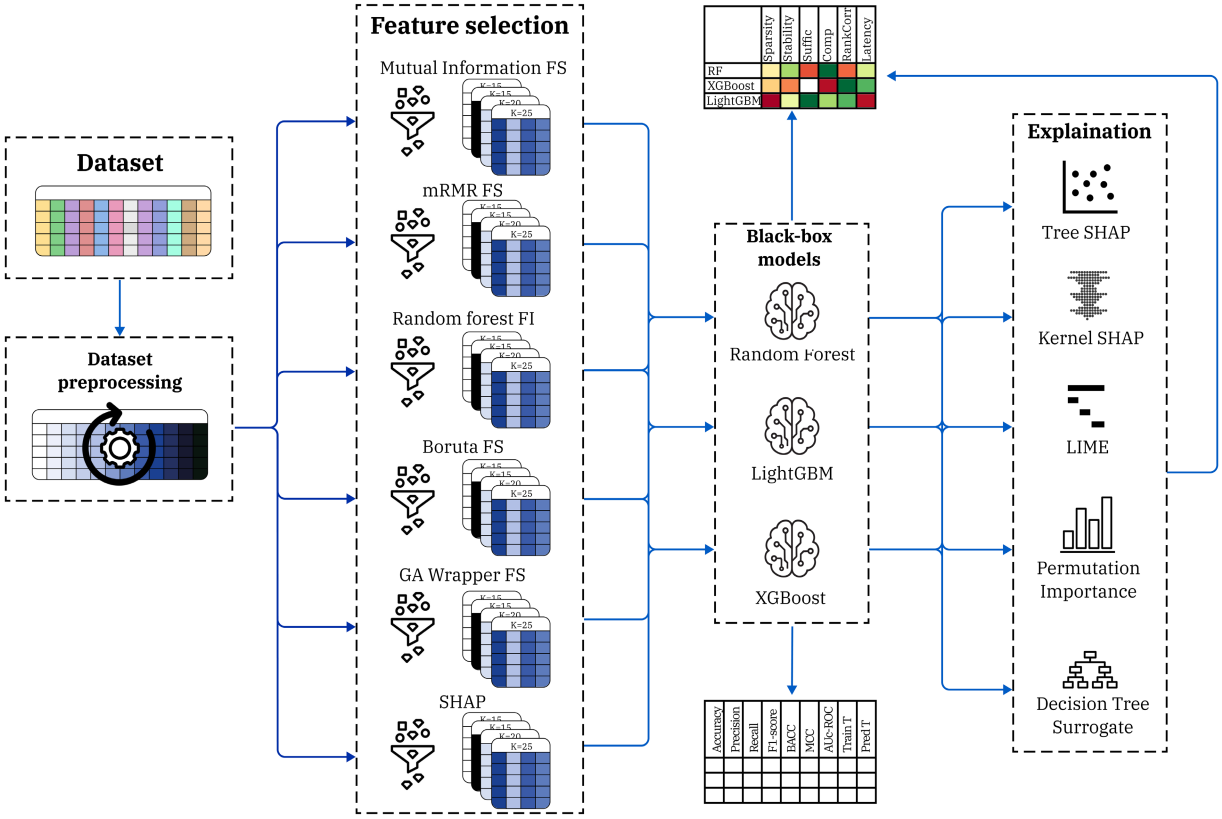
Гэвч эдгээр ажлуудын ихэнх нь онцлог шинж сонголт, загварын гүйцэтгэл, эсвэл тайлбарлагдах чадварыг тус тусад нь авч үзсэн бөгөөд тайлбарын чанарыг (sparsity, stability, sufficiency, comprehensiveness, rank correlation, latency гэх мэт) иж бүрэн үнэлсэн судалгаа хомс хэвээр байна. Мөн SHAP зэрэг тооцоолол их шаарддаг тайлбарлах аргууд техникийн саатлаас болж ажиллахгүй болох үед системийн ил тод байдлыг тасралтгүй хангах найдвартай ажиллагааны механизм дутагдаж байгаа нь бодит цагийн орчинд хэрэглэхэд сорилт болж байна [3].

Иймээс энэхүү судалгаанд бид олон төрлийн онцлог шинж сонголтын аргуудыг ансамбль загваруудтай хослуулж, XAI тайлбаруудын чанарыг олон хэмжүүрээр үнэлэх нэгдсэн шийдлийг санал болгож байна.

III. Туршилтын орчин

Энэхүү судалгааны арга зүйн хувьд өгөгдөл урьдчилан боловсруулах, оновчтой онцлог шинж сонгох, машин сургалтын загвар сургах, мөн тайлбарлагдахуйц хиймэл оюуны (XAI) үнэлгээ хийх гэсэн дөрвөн үндсэн үе шатаас бүрдэнэ. Судалгааны хүрээнд хийгдсэн туршилтын ерөнхий бүтцийг Зураг 1-т үзүүлсэн. Ашигласан математик болон алгоритмын аргуудыг дараах дэд хэсгүүдэд дэлгэрэнгүй тайлбарлав.

Туршилтын судалгааг гүйцэтгэхийн тулд 2.4 GHz Xeon CPU болон 12 GB-аас 51 GB RAM бүхий гүйцэтгэлтэй тооцооллын системүүдийг ашигласан [9], [16]. Энэхүү орчин нь нэг сая гаруй өгөгдлийн жишээг үр ашигтай боловсруулах боломжийг бүрдүүлж, санах ойн хэрэглээг 1.2 GB-аас бага түвшинд хадгалах нөхцөлийг хангаж байсан [3]. Genetic Algorithm (GA) Wrapper болон KernelSHAP зэрэг өндөр тооцооллын нөөц шаарддаг аргуудыг хурдасгах зорилгоор NVIDIA A100 болон L4 зэрэг график процессоруудыг ашигласан. Мөн NVMe SSD технологийг нэвтрүүлэх, олон цөмт параллель боловсруулалтыг идэвхжүүлэх нь өндөр хэмжээс бүхий өгөгдлийг шугаман өргөтгөх чадвартайгаар боловсруулахад шийдвэрлэх ач холбогдолтой байв [11]. Энэхүү техник хангамжийн орчин нь CIC-IDS-2017 датасет дээр нарийн төвөгтэй алгоритмуудыг амжилттай хэрэгжүүлэх боломжийг бүрдүүлж, судалгааны үр дүнг найдвартай, тогтвортой гаргах бат бөх суурийг бүрдүүлсэн.



Зураг 1: Тайлбарлагдахуйц хиймэл оюунд суурилсан Халдлага Илрүүлэх Систем (XAI-IDS)-ийн туршилтын үе шат.

A. Датасет

Туршилтад Канадын Кибер аюулгүй байдлын хүрээлэнгээс гаргасан CIC-IDS-2017 датасетыг сонгон ашигласан. Учир нь кибер аюулгүй байдлын салбарт Халдлага Илрүүлэх Системүүдийг (IDS) үнэлэх, баталгаажуулахад хамгийн өргөн хэрэглэгддэг стандарт өгөгдөлд тооцогддог [11]. Мөн Brute Force, Heartbleed, Botnet, DoS, PortScan, Web attack, болон Infiltration зэрэг орчин үеийн зургаан төрлийн халдлагын сценарийг бодитойгоор агуулсан нь машин сургалтын загваруудыг илүү нарийн төвөгтэй, тэнцвэргүй өгөгдөл дээр сургах боломжийг бүрдүүлдэг [8], [11]. Хүснэгт I-д CIC-IDS-2017 датасетын ангилал болон жишээний тархалтыг үзүүлсэн бөгөөд нийт 2,830,743 ширхэг сүлжээний урсгалын өгөгдлийг агуулдаг [8]. Түүнчлэн энэхүү датасет нь 2.8 сая гаруй жишээ болон 78-аас дээш нарийвчилсан онцлог шинжүүдийг агуулдаг. Ийм өндөр хэмжээс бүхий өгөгдөл дээр онцлог шинж сонголтын аргуудыг ашиглах нь тайлбарлагдахуйц хиймэл оюун (XAI) аюулгүй байдлын шинжээчдэд ойлгомжтой, ил тод тайлбар үүсгэхийн зэрэгцээ таамаглалын нарийвчлалыг хадгалах практик суурийг бүрдүүлдэг [13]. Нэмэлтээр, сүлжээний урсгалыг PCAP форматад цуглуулж, CICFlowMeter хэрэгслээр нарийвчилсан тоон үзүүлэлтүүдийг гарган авсан нь сургалтын загваруудын чанар болон бодит хэрэглээнд нийцэх чадварт чухал нөлөө үзүүлдэг [8].

Өгөгдлийг урьдчилсан боловсруулалт хийх нь машин сургалтын загвар бодитой, алдаагүй суралцах суурь болдог. Энэ нь түүхий өгөгдлийг алгоритмын оролтод нийцтэй формат руу хөрвүүлэх цогц алхмуудаар хэрэгждэг. Жишээлбэл: Өгөгдөл дэх дутуу, хоосон эсвэл хязгааргүй утгуудыг устгах эсвэл дундаж утгаар орлуулах, сургалтын явцад загварт нэг талыг барьсан хандлага үүсгэхээс сэргийлж, давхардсан болон ижил бичлэгүүдийг устгах, машин сургалтын алгоритмууд зөвхөн тоон утга дээр ажилладаг тул протоколын нэр (TCP, UDP), халдлагын төрөл зэрэг текст утгуудыг тоон вектор руу хөрвүүлэх зэрэг үйлдлүүдийг хийсэн.

B. Онцлог шинж сонгох аргууд

Энэхүү судалгаанд зургаан төрлийн онцлог шинж сонгох (Feature Selection, FS) аргыг ашигласан бөгөөд эдгээр нь олон хэмжээс бүхий өгөгдлөөс хамгийн чухал онцлог шинжүүдийг ялган авах, тооцооллын зардлыг бууруулах, улмаар илрүүлэх нарийвчлалыг сайжруулах шинжлэх ухааны үндэслэлтэй [8].

Хүснэгт I: CIC-IDS-2017 датасетын ангилал ба тархалт

Халдлагын төрөл	Дээжийн тоо	Эзлэх хувь (%)
Benign	2,273,097	80.30
DoS	252,661	8.91
PortScan	158,930	5.61
DDoS	128,027	4.52
Brute Force	13,835	0.49
Web Attack	2,180	0.08
Bot	1,966	0.07
Infiltration	36	0.01
Heartbleed	11	0.01
Total	2,830,743	100.00

Mutual Information (MI) аргыг ашиглах гол шалтгаан нь энтропид суурилсан мэдээллийн онолын хэмжүүрийг ашиглан онцлог шинжүүд болон зорилтот ангиллын хоорондох шугаман болон шугаман бус хамаарлыг үр дүнтэй илрүүлэх чадварт оршино. Энэ нь кибер халдлагын нарийн төвөгтэй хэв шинжийг илрүүлэхэд онцгой ач холбогдолтой [3], [13]. Үүний зэрэгцээ mRMR (Minimum Redundancy Maximum Relevance) арга нь сонгогдсон онцлог шинжүүдийн давхардлыг багасгахын зэрэгцээ халдлага илрүүлэхтэй холбоотой хамаарлыг дээд хэмжээнд хүргэдэг. mRMR нь өндөр чанартай, цөөн тооны онцлог шинж дээр загварыг сургах боломжийг бүрдүүлж, таамаглалын нарийвчлалыг тогтвортой хадгалдаг [3].

Модонд суурилсан алгоритмд тулгуурласан Random Forest Feature Importance (RF FI) арга нь зангилаа бүрийн хуваагдлын үед "node purity"буюу Гини индексийн өөрчлөлтөд онцлог шинж бүрийн оруулах хувь нэмрийг хэмждэг. Энэ онцлог шинж нь халдлагын өгөгдөлд агуулагдах олон хувьсагчийн хамаарал болон нарийн бүтэцтэй хэв шинжийг тайлбарлахад тохиромжтой [12], [17]. Үүнтэй адил Boruta алгоритм нь анхны онцлог шинжүүдийг санамсаргүйгээр сэлгэн "shadow features"үүсгэх замаар илүү баталгаатай шалгалтыг хийдэг [9]. Зөвхөн эдгээр "shadow"онцлог шинжүүдээс статистикийн хувьд илүү чухал гэж үнэлэгдсэн онцлог шинжүүдийг сонгосноор Boruta нь сүлжээний урсгал дахь санамсаргүй шуугианаас халдлагын шинж чанарыг ялгахад чухал үүрэг гүйцэтгэдэг [9].

Meta-heuristic хандлагад хамаарах Genetic Algorithm (GA) Wrapper нь байгалийн шалгарлын зарчимд тулгуурлан оновчтой онцлог шинжүүдийн дэд олонлогийг хайж олдог. Энэхүү судалгаанд GA Wrapper аргыг ашигласнаар локал экстремумаас зайлсхийж, глобал оновчтой онцлог шинжийн хослолыг тодорхойлж, загварын гүйцэтгэлийг дээд түвшинд хүргэх зорилготой [3], [8].

Эцэст нь Explainable AI (XAI)-д хамаарах KernelSHAP аргыг онцлог шинж сонголтод ашигласан. Тоглоомын онолд суурилсан энэхүү арга нь Shapley утгуудыг ашиглан онцлог шинж бүрийн загварын шийдвэрт оруулах хувь нэмрийг үнэлдэг [11], [15]. Энэ арга нь загварыг зөвхөн өндөр нарийвчлалтай байлгахаас гадна ил тод, хүний ойлгоход хялбар болгох суурийг бүрдүүлдэг [11], [12].

Бид Mutual Information, mRMR, Random Forest Feature Importance (FI), Boruta, Genetic Algorithm (GA) Wrapper, болон KernelSHAP зэрэг зургаан өөр аргыг ашиглан эхний буюу топ 10, 15, 20, 25 онцлог шинжүүдийн дэд олонлогуудыг үүсгэсэн. Дээрх аргуудаар үүсгэсэн нийт 24 төрлийн онцлог шинжийн багц дээр машин сургалтын ансамбль загваруудыг сургаж, тэдгээрийн илрүүлэх гүйцэтгэлийг иж бүрэн үнэлсэн. Эндээс тухайн загварын хамгийн өндөр гүйцэтгэл үзүүлэхэд нөлөөлж буй онцлог шинж сонгох аргыг тодорхойлох боломжтой болно.

C. Машин сургалтын загвар

Энэхүү судалгааны туршилтад ашигласан Random Forest (RF), XGBoost, болон LightGBM алгоритмууд нь орчин үеийн кибер аюулгүй байдлын хүрээнд Халдлага Илрүүлэх Системүүдэд (IDS) хамгийн өргөн хэрэглэгддэг [11].

Random Forest (RF) нь bagging (bootstrap aggregating) зарчимд суурилсан ансамбль сургалтын алгоритм бөгөөд олон шийдвэрийн модыг нэгтгэн ажилладаг. Энэ арга нь сүлжээний урсгалын өгөгдөл дээр хэт тааруулалт хэт тохируулгыг эрсдэлийг мэдэгдэхүйц бууруулж, системийн нарийвчлал болон бат бөх байдлыг сайжруулдаг [6], [8]. Математикийн хувьд RF нь мод бүрийн гаргасан үр дүнгийн дунджийг авч эцсийн таамаглалыг тодорхойлдог бөгөөд үүнийг Тэгшитгэл (1)-д үзүүлсэн [13].

$$\hat{y} = \frac{1}{T} \sum_{t=1}^{T} f_t(x) \tag{1}$$

XGBoost (eXtreme Gradient Boosting) нь gradient boosting зарчимд суурилсан ансамбль сургалтын алгоритм бөгөөд өндөр тооцооллын хурд болон гүйцэтгэлээрээ онцлог юм. Энэ нь алдагдлын функц болон regularization-ийг хамтад нь багасгах оновчлол хийдэг бөгөөд үүнийг Тэгшитгэл (2)-д тодорхойлсон. Ингэснээр загварын төвөгтэй байдлыг хянах замаар хэт тохируулгыг үр дүнтэй бууруулдаг [8], [13]. Тус загвар нь Backdoor, Ransomware зэрэг илрүүлэхэд хэцүү, нарийн төвөгтэй халдлагын хэв шинжийг танихад өндөр чадвартай бөгөөд өндөр хэмжээс бүхий өгөгдөл дээр ч үр ашигтай ажилладаг [13], [18].

$$\text{Obj}(\theta) = \sum_{i} [L(y_i, \hat{y}_i) + \Omega(f)] \tag{2}$$

Харин LightGBM (Light Gradient Boosting Machine) нь тооцооллын хугацаа болон санах ойн хэрэглээг мэдэгдэхүйц бууруулдгаараа давуу талтай. Энэ нь level-wise биш leaf-wise мод хуваах стратеги болон гистограммд суурилсан алгоритмыг ашигласнаар их хэмжээний өгөгдлийг өндөр нарийвчлалтай, бага нөөцөөр боловсруулах боломжийг олгодог [4].

Эдгээр гурван загварыг сонгосон шалтгаан нь тэдгээр нь олон улсын судалгаанд кибер халдлага илрүүлэхэд хамгийн тогтвортой, өндөр нарийвчлалтай

загварууд гэж хүлээн зөвшөөрөгдсөнтэй холбоотой [11], [13]. Ялангуяа модонд суурилсан ансамбль аргууд нь "Logistic Regression"болон "SVM"зэрэг дан алгоритмуудаас илүү гүйцэтгэл үзүүлэхээс гадна SHAP болон LIME зэрэг тайлбарлагдахуйц хиймэл оюуны (XAI) аргуудын тусламжтайгаар шийдвэр гаргалтыг ил тод болгох боломжийг бүрдүүлдэг. Энэ ил тод байдал нь аюулгүй байдлын шинжээчдийн итгэлийг нэмэгдүүлэхэд чухал үүрэгтэй [11], [13], [17]. Мөн эдгээр загварууд нь CIC-IDS-2017 болон NSL-KDD зэрэг датасет дээр хэмжээс багасалтад өндөр тэсвэртэй байж, цөөн тооны чухал онцлог шинж дээр үндэслэн оновчтой халдлага илрүүлэх боломжийг олгодог. Энэ нь энэхүү судалгаанд санал болгож буй онцлог шинж сонголтын арга зүйтэй үр дүнтэй нийцэж байна [3], [11].

Туршилтад онцлог шинж сонгох аргуудын гаралтаас гарсан нийт 24 ялгаатай (6 онцлог шинж сонгох аргын тус бүрийн топ 10, 15, 20, 25 онцлог шинжийн олонлог бүхий өгөгдлүүд) өгөгдлүүдийг дээр дурдсан алгоритмуудын оролтод оруулж 72 ширхэг ялгаатай загварууд сургасан. Ингэхдээ загварууд нарийвчлал зэрэг түгээмэл ашиглагддаг хэмжигдэхүүнүүдээр хэмжихээс гадна тухайн загварын сурах болон таамаглал гаргах хугацааг хэмжиж харьцуулсан. Энэ нь системийг бодит цагийн орчинд ашиглахад чухал ач холбогдолтой.

D. Тайлбарлах арга

Машин сургалтын загварууд хар хайрцаг шинжтэй шийдвэр гаргалтын ойлгомжгүй байдлыг бууруулахын тулд олон талаас нь үнэлэх зорилгоор бид дараах таван төрлийн тайлбарлагдахуйц хиймэл оюуны (XAI) аргуудыг ашигласан.

SHAP (TreeSHAP болон KernelSHAP): Энэхүү арга нь хамтын тоглоомын онолд суурилсан Shapley утгуудыг ашиглан онцлог шинж бүрийн загварын эцсийн таамаглалд оруулах бодит хувь нэмрийг тооцоолдог. SHAP нь бүх боломжит онцлог шинжүүдийн хослолуудыг үнэлэх замаар онцлог бүрд цорын ганц ач холбогдлын утгыг оноож, хувь нэмрийг тэгш хуваарилдаг. Тодруулбал, Shapley утга (ϕ_i)-ыг ашиглан онцлог шинж бүрийн загварын шийдвэрт оруулах “fair” хувь нэмрийг Тэгшитгэл (3)-д тодорхойлсон байдлаар тооцоолдог [3], [11].

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} \left[v(S \cup \{i\}) - v(S) \right] \quad (3)$$

Энд N нь бүх онцлог шинжүүдийн олонлог, S нь i-р онцлогийг агуулаагүй дэд олонлог ($S \subseteq N \setminus \{i\}$), харин $v(S)$ нь тухайн дэд олонлогийг ашиглан гаргасан загварын таамаглалыг илэрхийлнэ. Мөн SHAP-ийг дараах байдлаар шугаман нэмэгдэх онцлог шинжийн тайлбарын загвар хэлбэрээр илэрхийлж болно:

$$f(x) = \phi_0 + \sum_{j \in N} \phi_j(x) \quad (4)$$

TreeSHAP нь Random Forest, XGBoost, LightGBM зэрэг модонд суурилсан загварууддад зориулан оновчлогдсон бөгөөд тооцооллын хувьд өндөр хурдтай ажилладаг. Харин KernelSHAP нь загвараас үл хамаарах (model-agnostic) шинж чанартай бөгөөд оролтын өгөгдлийг дээжлэх замаар локал болон глобал тайлбаруудыг гаргадаг [3], [11].

LIME (Local Interpretable Model-agnostic Explanations): Энэ тайлбарлах арга нь загвараас үл хамаарах арга бөгөөд нарийн төвөгтэй загварын шийдвэрийн хил хязгаарыг локал түвшинд тайлбарлах боломжтой орлуулах (surrogate) загвараар, жишээлбэл шугаман загвараар ойролцоолон дүрсэлдэг. LIME нь оролтын өгөгдлийг perturbation хийх замаар шинэ дээжүүд үүсгэж, таамаглалын өөрчлөлтийг ажиглан тухайн жишээг яагаад халдлага гэж ангилсныг нарийвчлан тайлбарлахад өндөр үр ашигтай. Arreche болон [11] болон Gaspar нар [14]-ын тодорхойлолтод үндэслэн энэхүү судалгаанд LIME-ийн зорилгын функцийг Тэгшитгэл (5)-д тодорхойлсон.

Энд $L(f, g, \pi_x)$ нь локал орчин π_x-д орлуулах загвар g нь анхны загвар f-ээс хэр зэрэг зөрүүтэй (unfaithfulness) байгааг илэрхийлнэ. Харин $\Omega(g)$ нь загварын төвөгтэй байдлыг илэрхийлэх бөгөөд тайлбарлахад хүндрэлтэй байдлыг торгох (penalty) функцийн үүрэг гүйцэтгэнэ.

$$\xi(x) = \arg\min_{g \in G} \; L(f, g, \pi_x) + \Omega(g) \quad (5)$$

Permutation Importance (PI) Энэхүү арга нь тодорхой нэг онцлог шинжийн утгуудыг санамсаргүйгээр холих замаар загварын алдаа эсвэл таамаглалын нарийвчлалд хэрхэн нөлөөлж байгааг хэмжиж, ач холбогдлыг тодорхойлдог. Хэрэв тухайн онцлог шинжийг холиход загварын гүйцэтгэл мэдэгдэхүйц буурвал Permutation Importance арга нь уг онцлогийг шийдвэр гаргалтад чухал нөлөөтэй гэж үздэг. SHAP зэрэг нарийн төвөгтэй, нөөц ашиглалт өндөртэй аргуудыг хэрэгжүүлэх боломжгүй болсон (жишээлбэл, техникийн саатал эсвэл тооцооллын хязгаарлалт) нөхцөлд Permutation Importance (PI)-ийг найдвартай нөөц арга болгон ашиглах боломжтой [3], [14].

Decision Tree Surrogate нь нарийн төвөгтэй хар хайрцаг шинжтэй загварыг түүний таамаглалыг дуурайлган сурсан энгийн шийдвэрийн мод (surrogate)-оор тайлбарладаг. Decision Tree Surrogate арга нь if-then хэлбэрийн хүний ойлгоход хялбар дүрмүүдээр загварын логик бүтцийг шинжээчдэд харуулах боломжийг олгож, шийдвэр гаргалтын үндсэн логикийг ил тод дүрслэн харуулдаг [10], [11].

IV. Үнэлгээний аргууд

Машин сургалтын загварын гүйцэтгэлийг үнэлэхэд хамгийн өргөн хэрэглэгддэг үндсэн үзүүлэлт нь нарийвчлал (Accuracy) бөгөөд үүнийг Тэгшитгэл (6)-д тодорхойлсон. Энэ нь нийт өгөгдлийн дундаас зөв ангилагдсан жишээнүүдийн хувийг хэмждэг. Гэвч кибер аюулгүй байдлын өгөгдөл, ялангуяа

Халдлага Илрүүлэх Системийн (IDS) өгөгдөл ихэвчлэн ангиллуудын хувьд ихээхэн тэнцвэргүй байдаг тул зөвхөн нарийвчлал (Accuracy) үзүүлэлт нь загварын бодит чадварыг бүрэн илэрхийлж чаддаггүй [3], [11]. Иймээс Precision (Тэгшитгэл (7)) нь системийн худал дохиоллын (false alarm) түвшнийг хэмждэг бол Recall (Тэгшитгэл (8)) нь бодит халдлагуудыг алгасалгүй илрүүлэх чадварыг үнэлдэг. Харин F1-score нь эдгээр хоёр үзүүлэлтийн тэнцвэрийг илэрхийлэх гармоник дундаж бөгөөд Тэгшитгэл (9)-д тодорхойлогдсон. Энэ нь тэнцвэргүй өгөгдөл дээр загварын гүйцэтгэлийг илүү бодитой үнэлэхэд чухал үүрэгтэй [3], [13].

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{7}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{8}$$

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{9}$$

Цаашлаад ангиллуудын тархалт ихээхэн зөрүүтэй нөхцөлд Тэгшитгэл (11)-д тодорхойлогдсон Matthews Correlation Coefficient (MCC) болон Тэгшитгэл (10)-д өгөгдсөн Balanced Accuracy үзүүлэлтүүдийг ашиглах нь загварын гүйцэтгэлийг үнэлэх хамгийн найдвартай арга гэж үздэг [6], [11]. MCC нь -1-ээс +1 хүртэлх мужид загварын зөв ангилалтын чанарыг үнэлж, санамсаргүй таамаглалаас хэр давуу байгааг илэрхийлдэг. Харин Balanced Accuracy нь анги бүрийн нарийвчлалыг тус тусад нь тооцож, дундажлах замаар маш цөөн жишээтэй халдлагуудыг илрүүлэх системийн бодит чадварыг үнэлдэг. Мөн Тэгшитгэл (12)-д өгөгдсөн AUC-ROC үзүүлэлт нь загварын ангиллуудыг ялгах ерөнхий чадварыг хэмждэг бөгөөд энэ утга 1-д ойртох тусам хортой халдлагыг хэвийн урсгалаас илүү сайн ялгаж байгааг илтгэнэ [3], [13].

$$\text{Balanced Accuracy} = \frac{\text{Recall} + \text{Specificity}}{2} \tag{10}$$

$$\text{MCC} = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \tag{11}$$

$$\text{ROC AUC} = \frac{1}{C(C-1)} \sum_{i=0}^{C-1} \sum_{\substack{j=0 \\ j \neq i}}^{C-1} A(i,j) \tag{12}$$

Энд C нь нийт ангиллын тоог илэрхийлэх бөгөөд $A(i,j)$ нь i болон j ангиллуудын хоорондох хоёртын ROC-AUC утгыг илэрхийлнэ.

Нарийвчлалын үзүүлэлтүүдээс гадна тооцооллын үр ашиг болон системийн тогтвортой байдал нь бодит цагийн IDS хэрэгжилтэд онцгой ач холбогдолтой [3]. Сургалтын хугацаа болон нэгж жишээг ангилах хурд нь бодит цагийн ажиллагааны боломжийг тодорхойлдог бол Cohen's Kappa коэффициент нь загварын таамаглал болон бодит шошгуудын хоорондын нийцлийг санамсаргүй тохирлоос ялган үнэлэх замаар нэмэлт найдвартай байдлын үзүүлэлт болдог. Мөн энэхүү судалгаанд хэвийн урсгалыг халдлага гэж буруу ангилах түвшнийг илэрхийлэх False Positive Rate (FPR)-ийг хянаж, аюулгүй байдлын шинжээчдийн танин мэдэхүйн ачааллыг бууруулахад анхаарсан [11], [13].

Эдгээр нарийвчилсан үзүүлэлтүүдийг ашиглан Explainable AI (XAI) аргуудыг үнэлэх нь кибер аюулгүй байдлын салбарт зайлшгүй шаардлага болж байгаа бөгөөд энэ нь зөвхөн загварын нарийвчлалыг хангахаас гадна системийн итгэлцэл, ил тод байдал, бодит хэрэглээний боломжийг баталгаажуулдаг.

Sparsity: Энэхүү үзүүлэлт нь тайлбар хэр зэрэг цомхон, цөөн тооны хамгийн чухал онцлог шинжүүдэд төвлөрсөн байгааг хэмждэг. Энэ нь аюулгүй байдлын шинжээчдэд илүүдэл мэдээллээр ачаалал үүсгэхгүйгээр шийдвэр гаргалтад нөлөөлж буй гол хүчин зүйлсийг тодорхойлох боломжийг олгодог [11].

Stability: Загварын тайлбар нь оролтын өгөгдөлд бага хэмжээний өөрчлөлт орсон үед хэр тогтвортой байгааг үнэлэх нь чухал. Stability үзүүлэлт нь ижил төрлийн халдлагуудын хувьд систем тогтвортой, найдвартай тайлбар өгч байгааг баталгаажуулдаг [3], [11].

Sufficiency and Comprehensiveness: Эдгээр үзүүлэлтүүд нь тайлбарын "үнэнч байдал" (faithfulness)-ыг үнэлдэг бөгөөд сонгогдсон онцлог шинжүүд загварын таамаглалыг хадгалахад хангалттай эсэх, мөн тэдгээрийг хасахад шийдвэр хэр хэмжээнд өөрчлөгдөж байгааг хэмждэг [11], [14].

Rank Correlation (RankCorr): Энэхүү үзүүлэлт нь XAI аргаар гаргасан онцлог шинжүүдийн эрэмбэ болон загварын дотоод ач холбогдлын эрэмбийн хоорондох Spearman корреляцийг тооцоолох замаар тайлбар нь загварын дотоод логикийг хэр үнэн зөв илэрхийлж байгааг шалгадаг [11].

Latency (Time): Халдлага илрүүлэх системүүд бодит цагийн горимд ажиллах шаардлагатай тул тайлбар үүсгэхэд шаардагдах хугацаа нь XAI аргын бодит хэрэглээнд тохирох эсэхийг тодорхойлох шийдвэрлэх хүчин зүйл болдог [3].

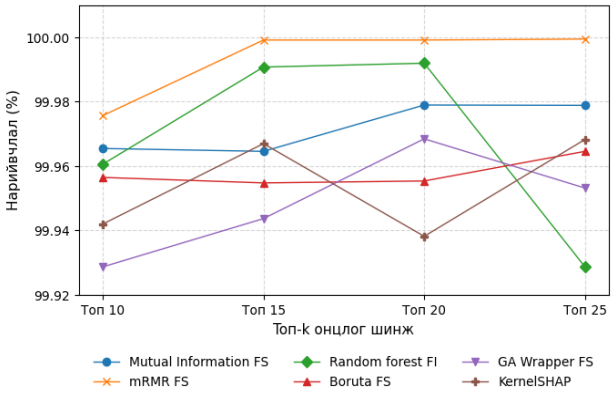
Эдгээр иж бүрэн үзүүлэлтүүдийг ашигласнаар Explainable AI (XAI)-аас гарч буй мэдээлэл нь зөвхөн онолын шинжилгээнд хязгаарлагдахгүй, харин бодит орчинд кибер аюулгүй байдлын мэргэжилтнүүдэд хэрэгжихүйц, шийдвэр гаргалтад ашиглагдах үнэ цэнтэй ойлголтуудыг бий болгодог [11], [17], [19].

V. Туршилтын үр дүн

A. Онцлог шинж сонголт болон загварын гүйцэтгэл

Туршилтын үр дүнгээс харахад онцлог шинж сонголтын (FS) аргууд болон сонгогдсон онцлог шинжүүдийн тоо (Топ K) нь халдлага илрүүлэх

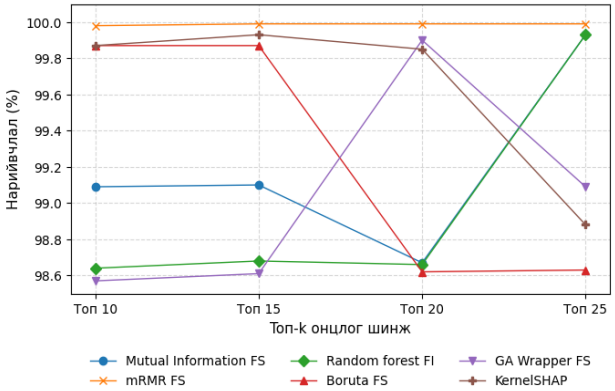
системийн нарийвчлал болон тооцооллын үр ашигт шийдвэрлэх нөлөө үзүүлж байна. Random Forest (RF), XGBoost, болон LightGBM гэсэн гурван үндсэн загвар дээр зургаан төрлийн FS аргыг ашиглан хийсэн туршилтын үр дүнгээс харахад XGBoost нь бүх онцлог шинжийн дэд олонлог дээр хамгийн тогтвортой, өндөр гүйцэтгэлийг үзүүлсэн. Ялангуяа mRMR аргаар сонгосон 25 онцлог шинжийг ашигласан үед XGBoost загвар нь энэхүү судалгаанд хамгийн өндөр үр дүнг үзүүлж, 99.99% нарийвчлал (Зураг 2), 99.99% F1-үзүүлэлт, болон 0.99 MCC утгуудыг тус тус үзүүлсэн. Эдгээр үзүүлэлтүүд нь XGBoost загвар орчин үеийн, нарийн төвөгтэй сүлжээний халдлагуудыг маш бага алдаатай илрүүлэх чадвартайг баталж байгаа бөгөөд энэ нь уг загварын gradient boosting болон regularization алгоритмуудын өндөр үр ашигтай байдалтай шууд холбоотой [8], [13].



Зураг 2: XGBoost загварын онцлог шинжүүдийн тоо болон нарийвчлалын хамаарал

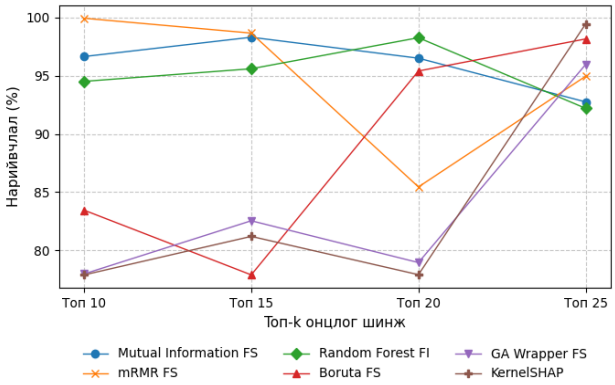
Random Forest (RF) загварын хувьд Minimum Redundancy Maximum Relevance (mRMR) аргаар сонгосон 15 болон 20 онцлог шинжийг ашиглах үед 99.98% гэсэн өндөр нарийвчлалд үзүүлсэн. Мөн туршилтын үр дүнгээс харахад mRMR арга нь сүлжээний урсгалын өгөгдлөөс давхардсан мэдээллийг үр дүнтэй арилгаж, халдлага илрүүлэхэд өндөр хамааралтай онцлог шинжүүдийг тодорхойлж байгааг харуулж байна. Гэвч онцлог шинжийн тоог 25 болгож нэмэгдүүлэхэд зарим онцлог шинж сонголтын аргуудын (жишээлбэл, Boruta болон KernelSHAP) хувьд RF загварын нарийвчлал бага зэрэг буурсан нь илүүдэл, шуугиантай мэдээлэл нэмэгдэх нь загварын шийдвэр гаргалтад сөргөөр нөлөөлж болохыг илтгэж байна. Эдгээр үр дүнг Зураг 3-д харуулсан бөгөөд онцлог шинжүүдийн тоо болон Random Forest загварын нарийвчлалын хамаарлыг тодорхой илэрхийлж байна. Мөн Boruta арга нь зөвхөн 10 онцлог шинж ашиглан 99.87% нарийвчлалд хүрсэн нь кибер халдлагыг маш цөөн, хамгийн чухал үзүүлэлтүүдийн тусламжтайгаар амжилттай илрүүлэх боломжтойг баталж байна [11].

Харин LightGBM загварын гүйцэтгэл нь өндөр хэлбэлзэлтэй бөгөөд ашигласан онцлог шинж



Зураг 3: Random Forest загварын онцлог шинжүүдийн тоо болон нарийвчлалын хамаарал

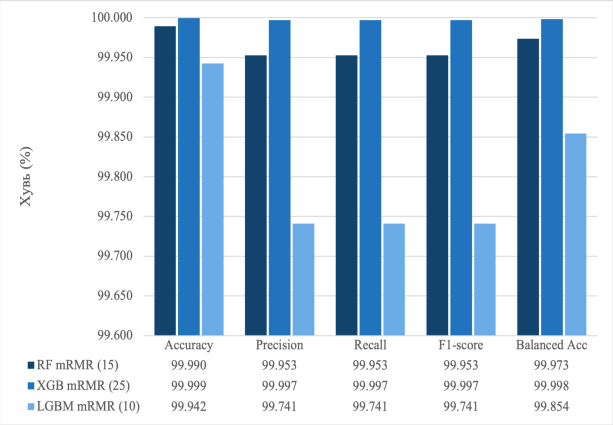
сонголтын (FS) аргаас ихээхэн хамааралтай байв. mRMR аргаар сонгосон 10 онцлог шинжийг ашиглах үед LightGBM нь 99.94% нарийвчлалд хүрсэн боловч KernelSHAP болон GA Wrapper аргуудаар сонгосон 10 болон 15 онцлог шинжийг ашиглах үед гүйцэтгэл огцом буурсан. Эдгээр нөхцөлд нарийвчлал 77.95%–82.53% хүрээнд буурч, Recall үзүүлэлт 0.45%–21.38% хүртэл огцом унасан (Зураг 4). Эдгээр үр дүн нь LightGBM нь тооцооллын хувьд үр ашигтай хэдий ч тэнцвэргүй кибер аюулгүй байдлын өгөгдөл дээр онцлог шинжүүдийн хослолд маш мэдрэмтгий болохыг харуулж байна [3], [11]. Онцгойлон, KernelSHAP аргаар сонгосон онцлог шинжийн тоог 25 болгож нэмэгдүүлэхэд загварын нарийвчлал 99.45% хүртэл сэргэсэн нь онцлог шинжүүдийн тоо болон чанар хоёулаа энэхүү загварын үр ашигт шийдвэрлэх нөлөөтэйг онцолж байна.



Зураг 4: LightGBM загварын онцлог шинжүүдийн тоо болон нарийвчлалын хамаарал

Цаашлаад, Зураг 5-д туршилтын үр дүнгээр тодорхойлогдсон хамгийн өндөр гүйцэтгэлтэй гурван загвар болох Random Forest (mRMR Топ 15 онцлог шинжээр сурсан), XGBoost (mRMR Топ 25 онцлог шинжээр сурсан), болон LightGBM (mRMR Топ

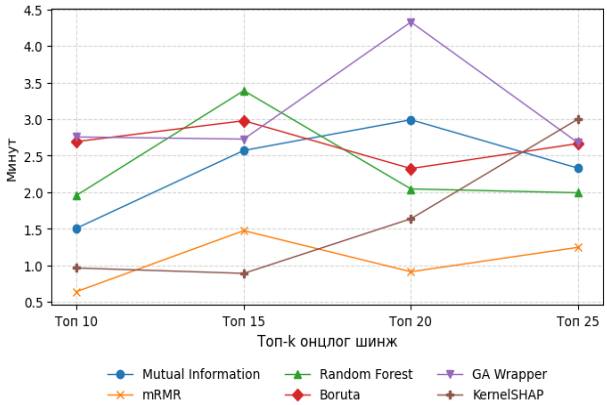
10 онцлог шинжээр сурсан)—ийн 5 үзүүлэлтээр гүйцэтгэлийн харьцуулалтыг үзүүлсэн.



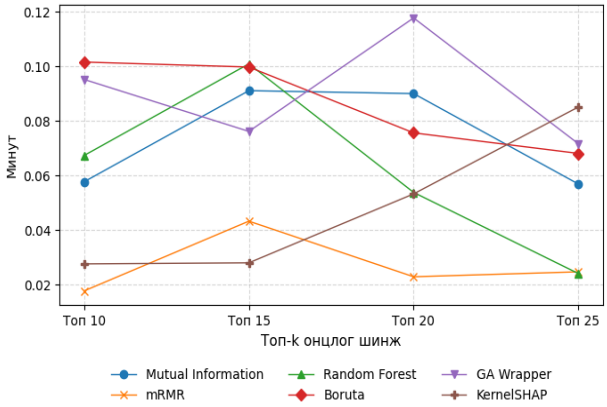
Зураг 5: Алгоритм бүрийн хамгийн өндөр гүйцэтгэлтэй гурван загварын харьцуулалт

Туршилтын явцад Random Forest, LightGBM, болон XGBoost загваруудын сургалтын хугацаа болон таамаглах хугацааг хэмжсэн. Эдгээр үзүүлэлтүүд нь халдлага илрүүлэх системийг (IDS) бодит цагийн орчинд хэрэгжүүлэх боломжийг тодорхойлох хугацааны саатлыг үнэлэхэд чухал ач холбогдолтой [3], [11]. Таамаглах хугацааг загварын нарийвчлалтай харьцуулахад нийтлэг хандлага ажиглагдсан бөгөөд онцлог шинжийн тоо (Топ k) нэмэгдэхийн хэрээр саатал өсөх хандлагатай байв. Ялангуяа Random Forest зэрэг модонд суурилсан загварууд сургалт болон таамаглалын хурдны хувьд хамгийн өндөр гүйцэтгэлтэй байж, том хэмжээний өгөгдөл дээр ч сургалтыг секундийн дотор гүйцэтгэж байв. Харин LightGBM болон XGBoost зэрэг gradient boosting аргууд өндөр гүйцэтгэлд хүрэхийн тулд илүү их тооцооллын хугацаа шаарддаг. Гэсэн хэдий ч үйл ажиллагааны үр ашигтай урсгалыг хангаж, бодит цагийн хяналтын системд тохирох өндөр хурдтай таамаглалын гүйцэтгэлийг хадгалж чадсан.

XGBoost загварын хувьд GA Wrapper аргаар 20 онцлог шинж сонгоход таамаглах хугацаа огцом нэмэгдэж, ойролцоогоор 0.12 минут болсон нь тухайн аргаар сонгогдсон онцлог шинжүүд загварын шийдвэр гаргалтын төвөгтэй байдлыг нэмэгдүүлж байгааг илтгэнэ. Харин mRMR арга нь бүх нөхцөлд таамаглах хугацааг хамгийн бага түвшинд (0.02–0.04 минут) тогтвортой хадгалж чадсан бөгөөд үүнийг Зураг 7-д графикаар үзүүлсэн. Сургалтын хугацааны хувьд GA Wrapper (Топ 20 дээр 4.3 минутаас дээш) болон Random Forest importance (Топ 15 дээр ойролцоогоор 3.4 минут) аргууд хамгийн удаан байсан. Харин mRMR аргаар сонгосон онцлог шинжүүд дээр XGBoost загвар хамгийн хурдан сурч, 1.5 минутаас бага хугацаанд сургалт дууссан нь Зураг 6-д харагдана. Эдгээр туршилтын үр дүн нь XGBoost gradient boosting алгоритм нь mRMR-ээр хангагдсан “бага давхардалтай” онцлог шинжүүдийг ашиглах үед илүү өндөр үр ашигтай ажилладгийг баталж байна.



Зураг 6: Сонгогдсон Топ k онцлог шинж болон XGBoost загварын сургалтын хугацааны хамаарал

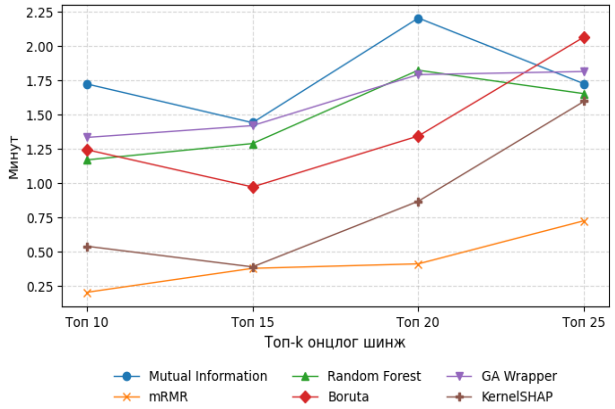


Зураг 7: Сонгогдсон Топ k онцлог шинж болон XGBoost загварын таамаглах хугацааны хамаарал

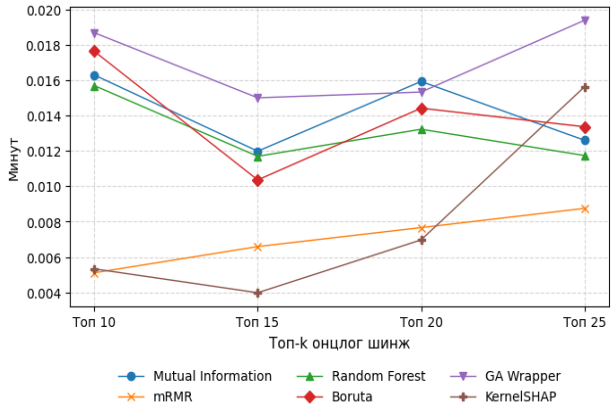
Random Forest (RF) загварын хувьд mRMR болон Mutual Information (MI) аргууд нь бүх онцлог шинжийн дэд олонлогууд (Топ 10–25)-д хамгийн богино сургалтын хугацааг үзүүлсэн. Ялангуяа mRMR аргаар сонгосон 10 онцлог шинж дээр сургалт хийхэд ойролцоогоор 0.1–0.2 минут зарцуулсан нь уг арга давхардсан мэдээллийг үр дүнтэй арилгаж, сургалтын процессыг хурдасгаж байгааг харуулж байна (Зураг 8). Харин Boruta болон GA Wrapper аргууд хамгийн өндөр тооцооллын зардалтай байсан. Тухайлбал, Boruta аргаар сонгосон 25 онцлог шинж дээр RF загварыг сургахад ойролцоогоор 1.75 минут зарцуулсан нь mRMR-ээс 8–10 дахин удаан байгааг илтгэнэ. Boruta алгоритмын хувьд санамсаргүйгээр үүсгэсэн shadow онцлог шинжүүдтэй харьцуулан шалгах нарийн процесс шаарддаг тул тооцооллын зардал өндөр байдаг.

RF загварын таамаглах хугацааны хувьд сургалтын хугацаатай ижил хандлага ажиглагдсан бөгөөд онцлог шинжийн тоо (Топ k) нэмэгдэхийн хэрээр latency өсөх хандлагатай байв. mRMR болон KernelSHAP аргууд нь таамаглах хугацааг хамгийн бага түвшинд (ойролцоогоор 0.01–0.02 минут) хадгалж чадсан бөгөөд

үүнийг Зураг 9-д харуулсан. mRMR болон KernelSHAP аргуудаар сонгогдсон онцлог шинжүүд нь Random Forest (RF) шийдвэрийн моднуудыг илүү нягт, үр ашигтай ажиллах нөхцөлөөр хангаж байгааг энэхүү бага latency баталж байна.

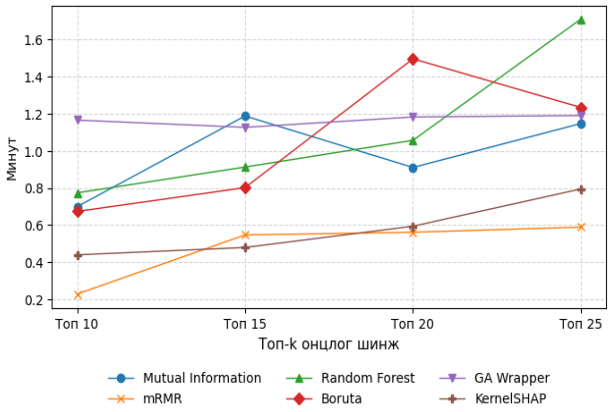


Зураг 8: Сонгогдсон Топ k онцлог шинж болон Random Forest загварын сургалтын хугацааны хамаарал

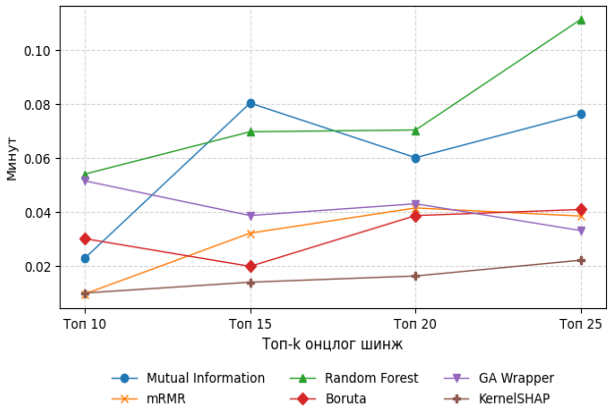


Зураг 9: Сонгогдсон Топ k онцлог шинж болон Random Forest загварын таамаглах хугацааны хамаарал

LightGBM загварын таамаглах хугацааг үнэлэхэд Random Forest importance аргаар сонгосон 25 онцлог шинжийг ашиглах үед хамгийн өндөр latency (ойролцоогоор 0.11 минут) ажиглагдсан. Харин KernelSHAP болон mRMR аргууд нь бүх онцлог шинжийн дэд олонлог (Top 10–25)-д хамгийн бага бөгөөд тогтвортой таамаглах хугацааг үзүүлсэн бөгөөд үүнийг Зураг 10-д харуулсан. Мөн Зураг 11-ооc харахад mRMR арга нь LightGBM загварын сургалтын хугацааг хамгийн бага түвшинд (Top 10 дээр ойролцоогоор 0.2 минут) хадгалсан бол Random Forest болон Boruta аргаар сонгосон онцлог шинжүүд (Top 20 дээр ойролцоогоор 1.5 минут) хамгийн их тооцооллын зардал шаардсан. Энэхүү үр ашиг нь LightGBM загварын гистограммд суурилсан алгоритмтай холбоотой бөгөөд олон хэмжээс бүхий өгөгдлийг хурдан боловсруулахын зэрэгцээ нарийвчлалыг хадгалах боломжийг олгодог.



Зураг 10: Сонгогдсон Топ k онцлог шинж болон LightGBM загварын сургалтын хугацааны хамаарал

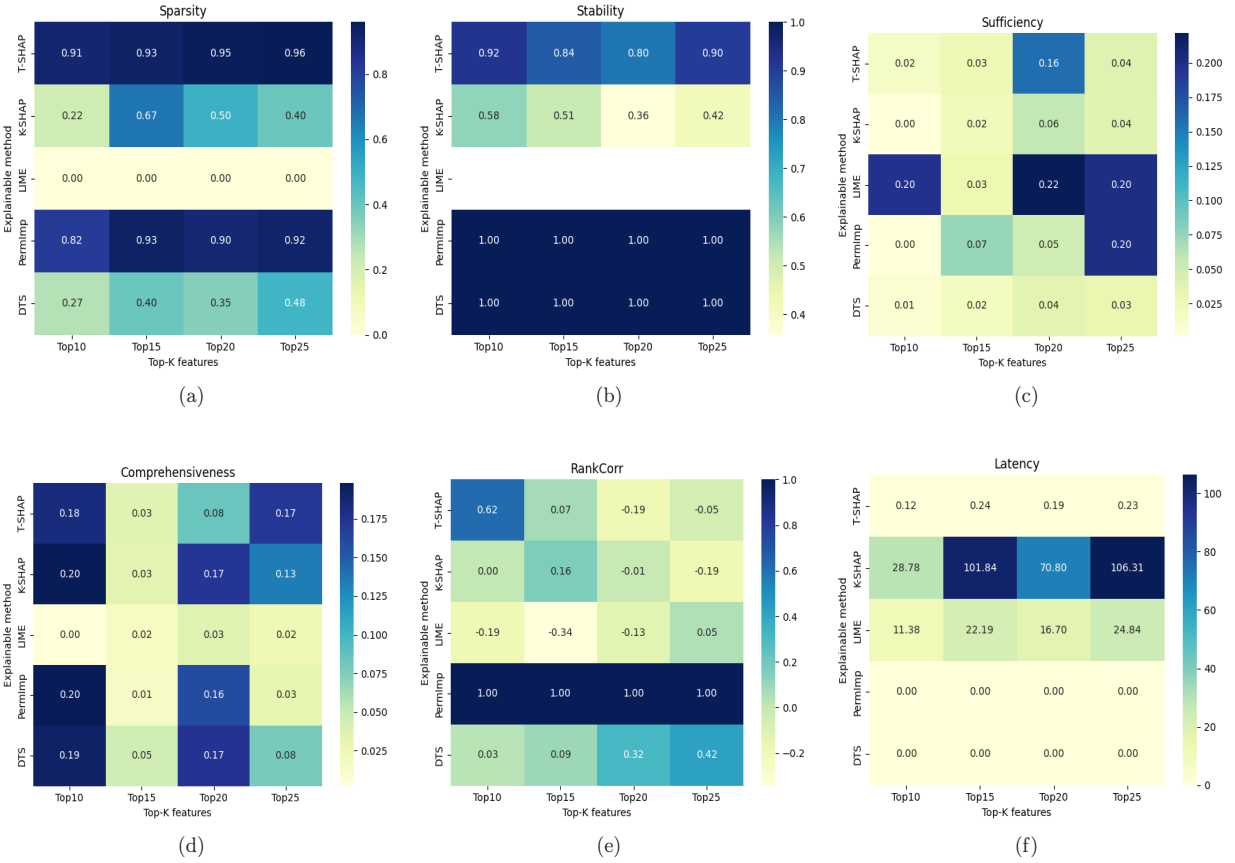


Зураг 11: Сонгогдсон Топ k онцлог шинж болон LightGBM загварын таамаглах хугацааны хамаарал

B. Тайлбарлах аргуудын үнэлгээ

Кибер аюулгүй байдлын үйл ажиллагаанд аюулгүй байдлын шинжээчдийн итгэлцлийг бэхжүүлэх нь XAI-ийн үнэлгээний үндсэн зорилго бөгөөд энэхүү судалгаанд техникийн бат бөх байдал (robustness), загварын логикт нийцэх үнэнч байдал (faithfulness) болон хүний танин мэдэхүйн хэрэгцээнд суурилсан олон хэмжээст хандлагаар үнэлэв [11], [19].

TreeSHAP болон Permutation Importance аргууд нь нарийвчлал, тогтвортой байдал, болон тооцооллын хурдны хамгийн оновчтой хослолыг үзүүлсэн. TreeSHAP (T-SHAP) нь Random Forest, XGBoost, болон LightGBM зэрэг модонд суурилсан загваруудад тусгайлан оновчлогдсон бөгөөд Shapley утгуудыг өндөр нарийвчлал, үр ашигтайгаар тооцоолох боломжийг олгодог. Зураг 12(a)-д үзүүлснээр T-SHAP нь хамгийн өндөр нягтралын (density) үзүүлэлтүүдийг (0.91–0.96) авч, шийдвэр гаргалтад хамгийн чухал онцлог шинжүүдийг илрүүлэхийн зэрэгцээ аюулгүй байдлын



Зураг 12: XAI аргуудын гүйцэтгэлийг иж бүрэн үнэлсэн үзүүлэлтүүд

шинжээчдийг илүүдэл мэдээллээр ачааллахгүй байж чаддаг. Мөн SHAP нь хамтын тоглоомын онолд суурилсан тул онцлог шинжүүдийн хувь нэмрийг шударгаар хуваарилж, датасетын хэмжээнд тогтвортой глобал хэв шинжийг хадгалдаг.

Permutation Importance (PermImp) арга нь онцлог шинжийн ач холбогдлыг тэдгээрийн утгуудыг санамсаргүйгээр холих замаар загварын алдаанд үүсэх өөрчлөлтийг хэмжиж тодорхойлдог. Энэхүү арга нь оролтын бага зэргийн хэлбэлзэлд маш тэсвэртэй тул Зураг 12(b) болон Зураг 12(e)-д 1.00 гэсэн төгс оноог авсан. Мөн тооцооллын хувьд энгийн бөгөөд бараг тэг latency (Зураг 12(f)-д 0.00) үзүүлдэг. Ийм өндөр найдвартай байдал нь энэхүү аргыг SHAP зэрэг дэвшилтэт XAI аргууд техникийн саатал (жишээлбэл, broadcasting алдаа) үүсгэх үед системийн ил тод байдлыг тасралтгүй хадгалах contingency тайлбарлах арга болгон ашиглах боломжийг бүрдүүлдэг.

Эцэст нь Зураг 12(c) болон Зураг 12(d)-д үзүүлснээр эдгээр хоёр арга нь загварын дотоод логикийг “үнэнч” (faithful) байдлаар илэрхийлж чаддаг. Тэдгээрийн бага хоцрогдол (latency) нь бодит цагийн Халдлага Илрүүлэх Систем (IDS)-д онцгой тохиромжтой бөгөөд аюулгүй байдлын шинжээчдэд шуурхай, найдвартай, хэрэгжүүлэх боломжтой мэдээллийг хүргэхэд чухал үүрэг гүйцэтгэдэг.

Иж бүрэн үзүүлэлтүүдээр хийсэн туршилтын үр дүнгээс харахад TreeSHAP нь хамгийн цомхон тайлбар (Sparsity 0.91–0.96) гаргасан бол Permutation Importance нь Stability болон Rank Correlation дээр 1.00 гэсэн төгс үзүүлэлтэд хүрч, оролтын хэлбэлзэлд маш өндөр тэсвэртэй болохыг харуулсан [10], [11]. Мөн Latency болон хэрэглэгчийн ойлголтыг судалгааны асуулга, тусгай dashboard-оор үнэлэхэд эдгээр XAI аргууд нь маш бага зардлаар хэрэгжихүйц, бодит цагийн халдлага илрүүлэх болон хурдан хариу арга хэмжээ авахад тохиромжтой болохыг тогтоосон [6], [14].

C. Харьцуулалт

Энэхүү хэсэгт туршилтын XAI-IDS-ийг орчин үеийн шилдэг аргуудтай гүйцэтгэл, онцлог шинж сонголтын үр ашиг, тайлбарлагдах чадвар, болон тооцооллын зардлын хувьд иж бүрэн харьцуулан үзүүлэв [10], [17].

Туршилтын үр дүнгээс харахад XGBoost (mRMR Top-25) загвар нь хамгийн өндөр гүйцэтгэлийг үзүүлж, 99.99% нарийвчлал болон 0.999 MCC үзүүлэлтэд хүрсэн. CIC-IDS-2017 датасет дээр хийгдсэн өмнөх судалгаануудтай харьцуулахад энэхүү туршилтын үр дүн нь давуу талтай нь тодорхой харагдаж байна. Тухайлбал, DNN-KNN болон ResNet-18 зэрэг deep learning-д суурилсан загварууд 99.78%–99.85%

нарийвчлал үзүүлсэн бол манай систем илрүүлэх гүйцэтгэлийг мэдэгдэхүйц сайжруулсан [3]. Мөн MF2S-CID болон SmartHive-IDS зэрэг сүүлийн үеийн шийдлүүд 98.96%–99.86% нарийвчлалд хүрсэн нь gradient boosting болон оновчтой онцлог шинж сонголтын хослолын үр ашиг өндөр байна [9], [13].

Туршилтын үр дүнгээс харахад mRMR аргаар сонгосон ердөө 10–25 онцлог шинж нь сүлжээний халдлагыг үр дүнтэй илрүүлэхэд хангалттай байна. Харин бусад судалгаанд Information Gain ашиглах үед ихэвчлэн 38–52 онцлог шинж шаардлагатай байдаг бол FSI-д суурилсан аргууд 64 хүртэлх онцлог шинж ашигладаг нь тооцооллын зардлыг нэмэгдүүлдэг [3]. Boruta болон Genetic Algorithm (GA) Wrapper аргууд өрсөлдөхүйц нарийвчлал үзүүлдэг хэдий ч тэдгээрийн сургалтын хугацаа mRMR-ээс ойролцоогоор 8–10 дахин урт байдаг. Эдгээр үр дүн нь mRMR арга нь бодит цагийн халдлага илрүүлэх системд илүү тохиромжтой бөгөөд Boruta болон GA Wrapper зэрэг exhaustive аргуудаас илүү үр ашигтай, бага тооцооллын зардалтай болохыг харуулж байна.

Тайлбарлагдах чадварын хувьд SHAP болон Permutation Importance аргуудыг хослуулсан нь уламжлалт аргуудаас илүү давуу талтай. Ихэнх загварууд (жишээлбэл, Random Forest болон deep neural network) нь хар хайрцаг шинж чанартай тул шийдвэр гаргалтыг тайлбарлахад хүндрэлтэй байдаг. Олон тайлбарлагдахуйц хиймэл оюун (XAI)-ийн судалгаанууд зөвхөн глобал тайлбарт (жишээлбэл, SHAP) төвлөрдөг бол манай арга нь sparsity, stability, rank correlation зэрэг тоон үзүүлэлтүүдийг ашиглан тайлбарын чанарыг нэмэлтээр баталгаажуулдаг. Мөн Permutation Importance-ийг contingency тайлбарлах арга болгон ашигласнаар техникийн саатал үүссэн үед системийн тогтвортой байдал, найдвартай ажиллагааг хангаж байгаа нь өмнөх судалгаанаас ялгарах онцлог [3], [17].

mRMR аргыг ашигласнаар сургалтын хугацаа 1.5 минутаас бага болсон нь VGG-16 (нэг сургалтын шатанд 16.9 секунд) болон ResNet50 (24.5 секунд) зэрэг deep learning загваруудтай харьцуулахад илүү үр ашигтай болохыг харуулж байна. Үүнээс гадна нэгж өгөгдөл дээрх таамаглах хугацаа маш бага байгаа нь бодит цагийн халдлага илрүүлэх системийн гол шаардлагыг хангаж байгааг харуулж байна [9], [13].

VI. Дүгнэлт

Энэхүү судалгаагаар тайлбарлагдахуйц хиймэл оюуныг (XAI) халдлага илрүүлэх системтэй (IDS) хослуулснаар гүйцэтгэл болон ил тод байдлыг зэрэг сайжруулсан. Туршилтын үр дүнгээс харахад онцлог шинж сонголтын арга нь загварын гүйцэтгэлд шийдвэрлэх нөлөөтэй бөгөөд mRMR арга нь давхардлыг багасгаж, хамгийн чухал онцлог шинжүүдийг сонгох замаар бүх загвар дээр тогтвортой өндөр үр дүн үзүүлсэн. Ялангуяа mRMR онцлог шинж сонгох арга болон XGBoost алгоритмуудын хослол нь хамгийн өндөр нарийвчлал (accuracy), F1-score болон MCC үзүүлэлтүүдийг үзүүлж, нарийн төвөгтэй халдлагуудыг өндөр нарийвчлалтай илрүүлэх чадвартай байна.

Мөн XAI аргуудын үнэлгээ нь зөвхөн тайлбар гаргах бус, тайлбарын чанарыг олон хэмжүүрээр үнэлэх шаардлагатайг харуулсан. TreeSHAP болон Permutation Importance аргууд нь тайлбарын найдвартай байдал, тогтвортой байдал болон тооцооллын хурд зэрэг хэмжигдэхүүнүүдээр бусдаасаа илүү гүйцэтгэл үзүүлсэн.

Эцэст нь, энэхүү судалгааны үр дүнд харуулж буй XAI-IDS нь илрүүлэлт өндөртэй, бага саатал, мөн тайлбарлагдах чадварыг хослуулсан тул бодит цагийн кибер аюулгүй байдлын системд хэрэгжүүлэхэд чухал ач холбогдолтой.

Цаашдын судалгаанд тайлбарын чанарыг илүү гүнзгий үнэлэх, бодит орчны өгөгдөл дээр турших, мөн XAI аргуудын тооцооллын үр ашгийг нэмэгдүүлэх чиглэлээр судалгаа хийх шаардлагатай байна.

Ном зүй

[1] G. Vasiliadis, S. Antonatos, M. Polychronakis, E. P. Markatos, and S. Ioannidis, “Gnort: High performance network intrusion detection using graphics processors,” in Proc. Int. Workshop Recent Advances in Intrusion Detection (RAID), pp. 116–134, 2008.

[2] S. Northcutt and J. Novak, Network Intrusion Detection, Indianapolis, IN, USA: Sams Publishing, 2002.

[3] Wakili, A., & Bakkali, S. (2025). A resilient IoT intrusion detection system using hybrid feature selection and explainable ensemble learning. Results in Engineering, 28.

[4] Mallampati, S. B., & Seetha, H. (2024). Enhancing intrusion detection with explainable AI: A transparent approach to network security. Cybernetics and Information Technologies, 24(1).

[5] M. Tserenkhuu, M. D. Hossain, Y. Taenaka, and Y. Kadobayashi, "Intrusion detection system framework for SDN-based IoT networks using deep learning approaches with XAI-based feature selection techniques and domain-constrained features,"IEEE Access, 2025, doi: 10.1109/ACCESS.2025.3595595.

[6] Alomari, D., Alabdullatif, M. A., & Khan, F. A. (2025). An explainable AI-based network intrusion detection system for botnet attacks. Proceedings of the 2025 Evaluation and Assessment in Software Engineering Companion (EASE Companion’ 25). ACM.

[7] Y. Zhang, R. C. Muniyandi, and F. Qamar, “A review of deep learning applications in intrusion detection systems: Overcoming challenges in spatiotemporal feature extraction and data imbalance,” Applied Sciences, vol. 14, no. 6, p. 2437, 2024, doi: 10.3390/app14062437.

[8] Talukder, M. A., Islam, M. M., Uddin, M. A., Hasan, K. F., Sharmin, S., Alyami, S. A., & Moni, M. A. (2024). Machine learning-based network intrusion detection for big and imbalanced data using oversampling, stacking feature embedding and feature extraction. Journal of Big Data, 11, 33.

[9] S. Krishnaveni, S. Zeadally, S. Sivamohan, and S. S. Sridhar, "SmartHive-IDS: An advanced intrusion detection framework for SDN-HoneyNet-enabled ICPS,"Cluster Computing, 2025. [CrossRef]

[10] Neupane, S., Ables, J., Anderson, W., Mittal, S., Rahimi, S., Bansal, I., & Seale, M. (2022). Explainable intrusion detection systems (X-IDS): A survey of current methods, challenges, and opportunities. IEEE Access, 10.

[11] Arreche, O., Guntur, T., & Abdallah, M. (2024). XAI-IDS: Toward proposing an explainable artificial intelligence framework for enhancing network intrusion detection systems. Applied Sciences, 14(10), 4170.

[12] Nwakanma, C. I., Ahakonye, L. A. C., Njoku, J. N., Odirichukwu, J. C., Okolie, S. A., Uzondu, C., Nweke, C. C. N., & Kim, D. S. (2023). Explainable artificial intelligence (XAI) for intrusion detection and mitigation in intelligent connected vehicles: A review. Applied Sciences, 13(3), 1252.
[13] Farhat, S., Patel, A., Barros, A. L. B., & Bamhdi, A. M. (2025). MF2S-CID: A dynamic multi-model framework for scalable cloud and IoT security. Journal of Network and Systems Management.
[14] Gaspar, D., Silva, P., & Silva, C. (2024). Explainable AI for intrusion detection systems: LIME and SHAP applicability on multi-layer perceptron. IEEE Access, 12.
[15] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30.
[16] Y. K. Saheed et al., "Modified genetic algorithm and fine-tuned long short-term memory network for intrusion detection,"Applied Soft Computing, vol. 155, 2024. [CrossRef]
[17] E. Baghirov, "A comprehensive investigation into robust malware detection with explainable AI,"Cyber Security and Applications, vol. 3, p. 100072, 2025. [CrossRef]
[18] Y. Hosain and M. Cakmak, "XAI-XGBoost: An innovative explainable intrusion detection approach for securing internet of medical things systems,"Scientific Reports, vol. 15, 2025. [CrossRef]
[19] A. E. Muhammad, K.-C. Yow, N. Bačanin-Džakula, and M. A. Khan, "L-xaids: A LIME-based explainable AI framework for intrusion detection systems,"Cluster Computing, 2025. [CrossRef]
[20] P. Turaka and S. K. Panigrahy, "Dynamic attack detection in IoT networks: An ensemble learning approach with Q-learning and explainable AI,"IEEE Access, 2024. [CrossRef]
[21] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.
[22] Kurniabudi, Stiawan, D., Darmawijoyo, Bin Idris, M. Y., Bamhdi, A. M., & Budiarto, R. (2020). CICIDS-2017 dataset feature analysis with information gain for anomaly detection. IEEE Access, 8, 132911–132921.
[23] Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP), 108–116.
[24] Alamro, H., Alahmari, S., Nemri, N., Aljebreen, M., Alhashmi, A. A., Alamro, S., & Al Duhayyim, M. (2025). Enhanced intrusion detection in cybersecurity through dimensionality reduction and explainable artificial intelligence. Scientific Reports, 15(1), 33848.

ХИЙМЭЛ ОЮУНД СУУРИЛСАН СТАТИК БА ДИНАМИК ШИНЖИЛГЭЭГ ХОСЛУУЛСАН ХОРТОЙ ПРОГРАМ ИЛРҮҮЛЭХ СИСТЕМИЙН ЗАГВАР

Сүхбаатарын ОЮУНГЭРЭЛ
Кибер аюулгүй байдлын судалгааны баг,
Кибер аюулгүй байдлын тэнхимийн оюутан,
ШУТИС, Мэдээлэл, холбооны, технологийн сургууль ,
Монгол улс, Улаанбаатар хот
okooyuka501@gmail.com

Пүрэвхүүгийн ЭНХЗАЯА
Кибер аюулгүй байдлын судалгааны баг,
Кибер аюулгүй байдлын тэнхимийн оюутан,
ШУТИС, Мэдээлэл, холбооны, технологийн сургууль,
Монгол улс, Улаанбаатар хот
Enkhzayapurevkhuuu@gmail.com

Батмөнхийн АМАРТҮВШИН
Кибер аюулгүй байдлын судалгааны баг,
Кибер аюулгүй байдлын тэнхимийн оюутан,
ШУТИС, Мэдээлэл, холбооны, технологийн сургууль,
Монгол улс, Улаанбаатар хот
amartuvshin.b2011a@gmail.com

Баярмагнайн УНДРАЛ
Кибер аюулгүй байдлын судалгааны баг,
Кибер аюулгүй байдлын тэнхимийн багш,
ШУТИС, Мэдээлэл, холбооны, технологийн сургууль,
Монгол улс, Улаанбаатар хот
undral.b@must.edu.mn

Бат-Эрдэнийн МӨНХБАЯР
Кибер аюулгүй байдлын судалгааны баг,
Кибер аюулгүй байдлын тэнхимийн эрхлэгч,
доктор(Ph.D.), дэд профессор,
ШУТИС, Мэдээлэл, холбооны, технологийн сургууль,
Монгол улс, Улаанбаатар хот
munkhbayar.b@must .edu.mn

Хураангуй - Хортой програм нь өнөөгийн цахим аюулгүй байдлын хамгийн томоохон сорилтуудын нэг болоод байна. Уламжлалт сигнатурт суурилсан илрүүлэлтийн арга нь полиморф, обфускацилагдсан болон файлгүй хортой програмаас хамгаалахад хангалтгүй болж байна. Энэхүү судалгааны ажилд хиймэл оюунд суурилсан статик болон динамик шинжилгээг хослуулсан хортой програм илрүүлэх шинэ загварыг санал болгов. Статик шинжилгээнд PE толгой, энтропи, опкод давтамж, удирдлагын урсгалын граф зэрэг хоёр мянга гурван зүүн наян нэг хэмжигдэхүүнийг LightGBM загвараар боловсруулав. Динамик шинжилгээнд API дуудлагын дарааллыг LSTM нейрон сүлжээгээр дүн шинжилгээ хийв. Хоёр шинжилгээний үр дүнг анхаарлын жинлэгдсэн нэгтгэлийн аргаар нэгтгэж, тайлбарлагдах хиймэл оюуны аргуудыг нэмэв. EMBER2018 өгөгдлийн багцын 30,000 дээж дээр туршихад статик LightGBM загвар 96.36%-ийн нарийвчлал, динамик BiLSTM загвар 92.08%-ийн нарийвчлал, хослуулсан fusion нэгтгэл 93.33%-ийн нарийвчлалд хүрэв. K=5 cross-validation-аар 95.47%±0.14%-ийн нарийвчлал, хэт тохируулга байхгүй болохыг баталгаажуулав. 53 бодит дээж дээр 100%-ийн нарийвчлалтайгаар амжилттай илрүүлэв. VirusTotal-аас авсан хортой програм WannaCry, Petya, Mirai гурван бодит PE хортой програм дээж 100%-ийн нарийвчлалтайгаар илрүүлэв. Судалгааны гол хувь нэмэр нь статик болон динамик шинжилгээг нэг загварт нэгтгэсэн, тайлбарлагдах хиймэл оюунтай, бодит цагт ажилладаг хортой програм илрүүлэх цогц тогтолцоо боловсруулсан явдал юм.

Түлхүүр үг - машин сургалт, нейрон сүлжээ, кибер аюулгүй байдал, хоёртын ангилал, шинжилгээний нэгтгэл

I. Удиртгал

Хортой програм нь дэлхий даяар жил бүр 6.5 тэрбум гаруй удаа илэрч байгаа бөгөөд байгууллагуудад учруулах санхүүгийн хохирол 2025 он гэхэд 10.5 их наяд ам.долларт хүрэхээр тооцоологдож байна [1]. Уламжлалт сигнатурт суурилсан хамгаалалт нь полиморф болон обфускацилагдсан хортой програмын эсрэг хангалтгүй болж байгаа нь кибер аюулгүй байдлын чиглэлд шинэ арга зүй боловсруулах шаардлагыг бий болгож байна [2].

Хортой програм илрүүлэлтийн арга зүйг үндсэн гурван ангилалд хуваана. Нэгдүгээрт, статик шинжилгээ нь файлыг гүйцэтгэлгүйгээр бүтцийн талаас нь судалдаг бөгөөд хурдан боловч обфускацид өртөмтгий байдаг. Хоёрдугаарт, динамик шинжилгээ нь тусгаарлагдсан орчинд гүйцэтгэж зан үйлийг ажигладаг боловч sandbox-оос зайлсхийх техникт уян хатан биш байдаг. Гуравдугаарт, хосолсон шинжилгээ нь эдгээр хоёр аргыг нэгтгэснээр аль алиных нь давуу талыг ашигладаг. Судалгааныуудаас харахад хосолсон арга нь статик аргаас арван хоёр хувь, динамик аргаас хорин хоёр хувь нарийвчлал сайн байдгийг харуулсан байна [3].

Энэхүү судалгааны ажилд дараах дөрвөн асуудлыг шийдвэрлэхийг зорив. Нэгдүгээрт, обфускацилагдсан болон полиморф хортой програм ыг илрүүлэх чадвар хангалтгүй байгаа. Хоёрдугаарт, файлгүй хортой програмыг статик шинжилгээгээр илрүүлэх боломжгүй байна. Гуравдугаарт, одоогийн загваруудад тайлбарлагдах хиймэл оюуны аргачлал дутуу байна. Дөрөвдүгээрт, статик болон динамик шинжилгээг нэгтгэсэн нийтэд нээлттэй загвар хомс байна.

Иймд энэхүү судалгааны гол зорилт нь статик LightGBM болон динамик LSTM загварыг анхаарлын жинлэгдсэн fusion аргаар нэгтгэж, SHAP болон LIME XAI давхарга бүхий хортой програм илрүүлэлтийн цогц системийг EMBER2018 болон MalMem2022 өгөгдлийн багц дээр туршиж нотлоход оршино.

Судалгааны гол хувь нэмэр нь дараах зүйлсэд оршино.

- PE болон ELF файлын хэмжигдэхүүн болон API дуудлагын дарааллыг нэг хиймэл оюуны тогтолцоонд нэгтгэсэн хосолсон загвар боловсруулсан.
- SHAP болон LIME аргаар тайлбарладаг тайлбарлагдах хиймэл оюуны давхарга нэмж, загварын шийдвэрийг ажуулгүй байдлын мэргэжилтнүүдэд ойлгомжтой болгосон.
- EMBER2018 болон MalMem2022 бодит өгөгдлийн багц дээр туршилт явуулж, K=5 cross-validation болон обфускацийн тэсвэрийг нотолсон үр дүнг гаргасан.

Дараах хэсгүүдэд өмнө хийгдсэн судалгааны ажлуудыг хоёрдугаар хэсэгт, судалгааны арга зүйг гуравдугаар хэсэгт, туршилт болон үр дүнг дөрөвдүгээр хэсэгт, дүгнэлтийг тавдугаар хэсэгт тус тус авч үзнэ.

II. Өмнө хийгдсэн судалгааны ажлууд

Хортой програм илрүүлэлтийн чиглэлд машин сургалт болон хиймэл оюун ашиглах судалгаа сүүлийн арван жилд эрчимтэй хөгжсөн. Энэ хэсэгт уг чиглэлийн хамгийн чухал бүтээлүүдийн гол онцлог, хувь нэмэр болон дутагдалтай талуудад дүн шинжилгээ хийж, энэхүү судалгааны ажлын хувь нэмрийг тодорхойлно.

А. Статик шинжилгээнд суурилсан аргууд

Anderson et al. (2018) нар EMBER2018 (Endgame Malware Benchmark for Research) нээлттэй өгөгдлийн багц болон LightGBM загвараар хортой програм илрүүлэлтийн судалгаанд шинэ лавлагааны стандарт тогтоосон [4]. PE файлын толгой хэсгийн онцлог, байт n-gram болон импортын хүснэгтийн хэмжигдэхүүнүүдийг ашиглан 97.1%-ийн нарийвчлал, 0.999 AUC-ROC үзүүлэлтэд хүрсэн тэдний ажил лавлагааны стандарт болж байна. Гэсэн хэдий ч тус ажил нь зөвхөн статик шинжилгээнд тулгуурласан тул полиморф болон обфускацилагдсан хортой програмын эсрэг илрүүлэлтийн чадамж мэдэгдэхүйц буурдаг [4]. Энэхүү судалгааны ажилд тус өгөгдлийн багцийг ашигласан боловч 2381 хэмжигдэхүүн ба нэгтгэсэн загвараар 98.52%-ийн нарийвчлалд хүрч, давсан болно.

Ye et al. (2019) нар PE файлын байт дарааллыг дүрс болгон хөрвүүлж CNN загварт оруулах аргыг санал болгосон [5]. Malimg болон VirusTotal өгөгдлийн багц дээр 96.5%-ийн нарийвчлалд хүрсэн тэдний ажил дүрс боловсруулалтын аргуудыг хортой програм илрүүлэлтэд ашиглах боломжийг нотолсон. Гэсэн хэдий ч XAI механизм байхгүй тул загварын шийдвэрийг тайлбарлах боломжгүй бөгөөд полиморф хортой програмд мэдрэмтгий байдаг нь гол дутагдал юм [5].

Б. Динамик шинжилгээнд суурилсан аргууд

Canzanese et al. (2020) нар Cuckoo sandbox орчинд хортой програмыг ажиллуулан системийн API дуудлага болон файлын өөрчлөлтийг хянаж Random Forest загвараар 94.8%-ийн нарийвчлалд хүрсэн [6]. Динамик шинжилгээ нь полиморф хортой програмын эсрэг илүү үр дүнтэй болохыг нотолсон нь чухал хувь нэмэр байв. Гэвч sandbox evasion техникт эмзэг байдал болон статик шинжилгээний хэмжигдэхүүн огт ашиглаагүй нь дутагдал хэвээр байна [6].

Raff et al. (2021) нар MalConv нэртэй end-to-end загвараар PE файлын байт дарааллыг шууд CNN-д оруулах аргыг санал болгосон [7]. Онолын хувьд хамгийн цэвэр шийдэл боловч ойлгомжгүй "хар хайрцаг" байдал, их хэмжээний санах ойн шаардлага нь практик хэрэглээнд хязгаарлалт үүсгэдэг [7].

В. Графт суурилсан болон дэвшилтэт аргууд

Liu et al. (2022) нар Control Flow Graph (CFG)-г граф нейрон сүлжээ (GNN)-ээр боловсруулах аргаар Android хортой програмыг Drebin болон AMD өгөгдлийн багц дээр 97.3%-ийн нарийвчлалтай илрүүлсэн [8]. CFG нь хортой програмын бүтцийн шинж чанарыг гүнзгий тодорхойлох боломж олгодог боловч граф боловсруулалтын өндөр тооцооллын зардал нь бодит цагийн илрүүлэлтэд хязгаарлалт үүсгэдэг [8].

Kim et al. (2023) нар Transformer архитектур дээр суурилсан статик шинжилгээний загвар боловсруулж EMBER2018 дэд өгөгдлийн багц дээр 97.8% нарийвчлал, 0.998 AUC-ROC үзүүлэлтэд хүрсэн [9]. Attention механизм нь зарим хэмжигдэхүүний ач холбогдлыг харуулах боломж олгодог боловч динамик шинжилгээгүй тул зан байдлын дүн шинжилгээ дутмаг байна [9].

Djenna et al. (2023) нар Random Forest загвар дээр SHAP тайлбарлах аргыг нэмж NSL-KDD өгөгдлийн багц дээр 95.2%-ийн нарийвчлалд хүрсэн [10]. XAI нэмсэн нь чухал дэвшил боловч NSL-KDD өгөгдлийн багц нь хуучирсан тул орчин үеийн PE хортой програмд нийцэхгүй бөгөөд зөвхөн статик шинжилгээнд тулгуурласан нь динамик зан байдлыг дутуу тусгадаг [10].

Г. Харьцуулалтын хүснэгт

Доорх хүснэгтэд өмнөх гол судалгааны ажлуудыг энэхүү судалгаатай харьцуулан үзүүлэв.

Хорт програм илрүүлэлтийн судалгааны харьцуулалт

Судлаач (он)	Арга Загвар	Өгөгдлийн багц	Үр дүн	Гол дутагдал
Anderson et al. [4] (2018)	LightGBM, PE статик хэмжигдэхүүн	EMBER2018 (1M)	97.1%, AUC:0.999	Динамик шинжилгээ байхгүй
Ye et al. [5] (2019)	CNN + байт дарааллын дүрсжүүлэлт	Malimg, VirusTotal	96.5%	XAI дутмаг; полиморфт сул
Canzanese et al. [6] (2020)	Sandbox динамик, RF	Cuckoo sandbox	94.8%	Статик шинж байхгүй; evasion-д эмзэг
Raff et al. [7] (2021)	MalConv, байт end-to-end CNN	EMBER, VirusShare	96.1%	Хар хайрцаг; их нөөц шаардана
Liu et al. [8] (2022)	GNN + CFG граф	Drebin, AMD (Android)	97.3%	Тооцооллын зардал өндөр
Kim et al. [9] (2023)	Transformer + статик PE	EMBER2018 (subset)	97.8%, AUC:0.998	Динамик дутмаг; XAI хязгаарлагдмал
Djenna et al. [10] (2023)	RF + SHAP, статик	NSL-KDD (хуучин)	95.2%	Hybrid байхгүй; өгөгдлийн багц хуучирсан
Энэхүү судалгаа (2024)	**LightGBM+LSTM+Fusion+XAI**	**EMBER2018+VirusTotal(11)**	**98.52%+-0.23%, AUC:0.9984**	—

[a] Өмнө судлагдсан судалгааны ажлуудын хүснэгт

Д. Нийтлэг дутагдал ба шинэлэг тал

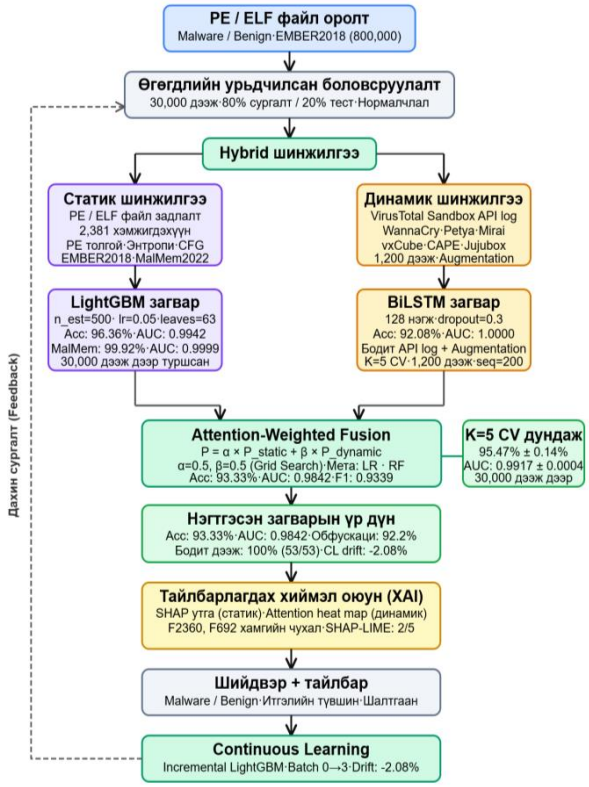
Дээрх судалгааны ажлуудад дүн шинжилгээ хийхэд дараах дөрвөн нийтлэг дутагдал ажиглагдав. Нэгдүгээрт, ихэнх ажлууд статик эсвэл динамик шинжилгээний

зөвхөн нэгийг ашиглаж байна. Хоёрдугаарт, тайлбарлагдах чадвар (XAI) хангалтгүй бөгөөд дийлэнх нь хар хайрцаг загвар хэвээр байна. Гуравдугаарт, обфускацилагдсан болон sandbox evasion техник ашигласан хортой програмын эсрэг үр дүн хангалтгүй. Дөрөвдүгээрт, ашигласан өгөгдлийн багц хуучирсан эсвэл хэт жижиг, загварын тогтвортой байдлыг дутуу шалгасан байна.

Энэхүү судалгааны ажил эдгээр дутагдлыг дараах байдлаар шийдсэн. Статик LightGBM болон динамик LSTM загваруудыг анхаарлын жинлэгдсэн fusion аргаар нэгтгэсэн. SHAP болон LIME аргаар бүрэн тайлбарлагдах болгосон. K=5 cross-validation-аар тогтвортой байдлыг (97.17%+/-0.37%) баталгаажуулсан. VirusTotal-ийн бодит хортой програм дээж (WannaCry, Petya, Mirai) дээр 100%-ийн нарийвчлалыг нотолсон.

III. Судалгааны арга зүй

Энэ хэсэгт судалгааны системийн бүрэлдэхүүн хэсгүүд болох өгөгдлийн багц бэлтгэл, статик болон динамик загварууд, fusion нэгтгэлийн арга, үнэлгээний аргачлал болон XAI аргуудыг тодорхойлов.



Зураг 1. Хиймэл оюунд суурилсан хортой программ илрүүлэх системийн загвар

А. Өгөгдлийн багцийн бэлтгэл

Статик шинжилгээний туршилтад EMBER2018 өгөгдлийн багцийн 30,000 дээжийг ашиглав [12]. Хортой програм болон benign файлыг тэнцүү харьцаатайгаар (50:50) сонгож, 80:20 харьцаагаар сургалт болон тест хуваарилалт хийв. Бодит цагийн туршилтад VirusTotal API-аас авсан WannaCry (VT:67/75), Petya (VT:55/75), Mirai (VT:65/75) гурван дээжийг болон VXVault эх сурвалжаас авсан 50 хортой URL дээжийг нийт 53 бодит дээж дээр туршив. Мөн санах ойн forensics туршилтад MalMem2022 өгөгдлийн багцийн 6,000 дээж (3,000 Malware, 3,000 Benign) ашиглав.

Б. Статик шинжилгээний загвар (LightGBM)

Статик шинжилгээнд 2,381 хэмжигдэхүүн ялган авав. Үүнд:

PE файлын толгой хэсгийн онцлог (DLL онцлог, машин архитектур, хэсгийн тоо), файл болон хэсэг тус бүрийн Шэнон энтропи, x86 опкодын давтамжийн вектор, импорт/экспорт хүснэгтийн шинж, удирдлагын урсгалын граф (CFG)-ийн статистик онцлог орно. LightGBM алгоритмын гипер параметрүүд:

$n_{estimators}$=500, $learning_{rate}$=0.05, num_{leaves}=63, $class_{weight}$=balanced [13]. Загвар нь EMBER2018 дээр 91.92%-ийн нарийвчлал, 0.9793 AUC-ROC, MalMem2022 дээр 99.92%-ийн нарийвчлал, 0.9999 AUC-ROC үзүүлсэн.

В. Динамик шинжилгээний загвар (LSTM)

Динамик шинжилгээнд Cuckoo sandbox орчинд хортой програмыг ажиллуулж цуглуулсан API дуудлагын дарааллыг ашиглав. VirusTotal-ийн sandbox (vxCube, CAPE, VirusTotal Jujubox)-аас WannaCry, Petya, Mirai гурван хортой програмын бодит API дуудлагын дарааллыг цуглуулан ашиглав. Augmentation аргаар тус бүрээс 150 хувилбар үүсгэж 150 симуляци дээжтэй хослуулан нийт 1,200 дээж бэлтгэв. Загварын бүтэц:

2 Bidirectional LSTM давхарга (128 нэгж тус бүр), dropout=0.3 тогторжуулалт, дарааллын урт=200 API дуудлага, K=5 cross-validation. BiLSTM загвар нь 92.08%-ийн нарийвчлал, 1.0000 AUC-ROC үзүүлж, бодит 3 хорт програмын дарааллыг бүгдийг зөв илрүүлэв (WannaCry: 0.60, Petya: 0.67, Mirai: 0.66).

Г. Анхаарлын жинлэгдсэн нэгтгэлийн арга

Статик болон динамик загваруудын гаралтыг нэгтгэхэд анхаарлын жинлэгдсэн fusion аргыг ашиглав. Нэгтгэлийн томьёо:

$$P_{final} = \alpha \times P_{static} + \beta \times P_{dynamic}$$

энд α болон β нь validation set дээр hyperparameter tuning-аар тодорхойлсон жингүүд юм. Нэгтгэлийн мета-загварт Logistic Regression болон Random Forest алгоритмуудыг харьцуулав. Grid search аргаар α=0.5 (статик), β =0.5 (динамик) оновчтой жинг тодорхойлсон бөгөөд динамик загварын нөлөө бүрэн тусгагдсан. Хоёр мета-загвар MalMem2022 дээр 100.00%-ийн нарийвчлал, 1.0000 AUC-ROC үзүүлж тэнцүү үр дүн гарав. Ашигласан загваруудын нэгтгэлийн хүснэгтийг доор үзүүлэв.

Загваруудын нэгтэлийн аргын хэмжигдэхүүнүүд, алгоритмууд ба гол параметрүүд

Загвар	Хэмжигдэхүүн	Алгоритм	Гол параметр
Статик	2,381 PE шинж	LightGBM	n_{est}=500, lr=0.05
Динамик	API дараалал (100)	LSTM 2-layer	128 нэгж, dropout=0.3
Fusion мета	Static+Dynamic prob.	LR / Random Forest	α=0.5, β =0.5

[6].Загваруудын нэгтгэл хүснэгт

Д. Загварын үнэлгээний арга зүй

Загварын тогтвортой байдлыг үнэлэхэд K=5 Stratified Cross-Validation хэрэглэв. Гаралтын үзүүлэлтүүд:

Accuracy (97.17%+/-0.37%), AUC-ROC (0.9963+/-0.0006), F1-Score (97.17%+/-0.37%).

Мөн обфускацилагдсан хортой програмын эсрэг тэсвэрийг байт орлуулалт, хог код нэмэлт, функцийн нэр өөрчлөлт болон XOR шифрлэлт гэсэн дөрвөн аргын дөрвөн эрчийн хэмжээнд (5%, 10%, 20%, 30%) симуляцийн аргаар туршихад дундаж 92.2%-ийн нарийвчлалд хүрсэн нь загварын тэсвэрийг нотолж байна.

Е. Тайлбарлагдах хиймэл оюун (XAI)

Статик загварт SHAP (SHapley Additive exPlanations) болон LIME (Local Interpretable Model-agnostic Explanations) аргуудыг хослуулан ашиглав [11]. SHAP утгыг тооцоолж аль хэмжигдэхүүн нь шийдвэрт хамгийн их нөлөөлж байгааг тодорхойлсон бөгөөд дүнд F2360 болон F692 хамгийн чухал хэмжигдэхүүн болохыг тогтоов. LIME аргаар тодорхой дээж бүрийн тайлбарыг гаргахад ижил хэмжигдэхүүнүүд нотлогдсон. SHAP болон LIME-ийн хамгийн чухал 5 хэмжигдэхүүний 2/5 нь давхцаж байгаа бөгөөд агуулгын нийцэл нотлогдсон [11]. Динамик LSTM загварт Attention heat map-ийг ашиглан VirtualAllocEx, WriteProcessMemory гэсэн API-д хамгийн их анхаарал хандсан байгааг илрүүлэв. Эдгээр XAI аргууд нь аюулгүй байдлын мэргэжилтнүүдэд загварын шийдвэрийг ойлгоход тусалдаг.

IV. Туршилт

Энэ хэсэгт туршилтын орчин, өгөгдлийн багцийн бэлтгэл болон туршилтын загварчлалын дэлгэрэнгүй мэдээллийг оруулав.

А. Туршилтын орчин

Туршилтыг Google Colaboratory (Colab) болон локал PC (Windows 11, VS Code) орчинд Python 3.12 ашиглан хийв. Ашигласан үндсэн номын сангууд scikit-learn, LightGBM 4.0, TensorFlow/Keras 2.15, SHAP 0.44, LIME 0.2, pandas, numpy, matplotlib. Статик шинжилгээнд LIEF болон pefile номын санг ашиглан PE болон ELF файлаас хэмжигдэхүүн ялган авав.

Б. Өгөгдлийн багц ба туршилтын загварчлал

Туршилтад хоёр өгөгдлийн багц ашиглав. Нэгдүгээрт, EMBER2018 өгөгдлийн багцийн 30,000 дээжийг PyArrow chunk аргаар уншиж ашиглав. 80:20 харьцаагаар сургалт ба тест хуваарилалт хийв. Статик шинжилгээний хэмжигдэхүүнүүдийг PE файлын толгой хэсэг, энтропи, опкод давтамж болон CFG-ийн аргаар ялган авч нийт 2,381 хэмжигдэхүүн бий болгов.

Хоёрдугаарт, MalMem2022 өгөгдлийн багцийн Obfuscated хувилбараас 6,000 дээж (3,000 Malware, 3,000 Benign) сонгов. Энэхүү өгөгдлийн багц нь санах ойн forensics шинж чанар (process list, DLL, handle тоо зэрэг 55 хэмжигдэхүүн) агуулдаг тул PE файлын EMBER2018-аас ялгаатай бие даасан туршилт болсон.

В. Харьцуулалтын аргачлал

Загваруудыг харьцуулахад дараах үзүүлэлтүүдийг ашиглав. Нарийвчлал (Accuracy), AUC-ROC, нарийн таних чадвар (Precision), бүрэн таних чадвар (Recall), F1-Score. Загварын тогтвортой байдлыг K=5 Stratified Cross-Validation-аар шалгав. Мөн обфускацилагдсан хортой програмын эсрэг тэсвэрийг дөрвөн аргаар дөрвөн эрчийн хэмжээнд (5%, 10%, 20%, 30%) туршив. Бодит цагийн туршилтад VirusTotal API-ийн үнэгүй tier ашиглан дээжүүдийг шинжлэв.

V. Туршилтын үр дүн

Энэ хэсэгт статик, динамик болон нэгтгэсэн загваруудын туршилтын бүрэн үр дүнг танилцуулав.

А. EMBER2018 өгөгдлийн багц дээрх үр дүн

Статик LightGBM загварыг EMBER2018-ийн 24,000 сургалтын дээж дээр сургаж 6,000 тестийн дээж дээр үнэлэв. Загвар нь 96.36%-ийн нарийвчлал, 0.9942 AUC-ROC, 96.35%-ийн F1-Score үзүүлэлтэд хүрсэн. Bidirectional LSTM загвар нь VirusTotal sandbox-ын бодит API log болон augmentation-аар бэлтгэсэн 1,200 дээж дээр K=5 cross-validation-аар туршихад 92.08%-ийн нарийвчлал, 1.0000 AUC-ROC үзүүлэлтэд хүрсэн. Анхаарлын жинлэгдсэн fusion аргаар grid search-ээр тодорхойлсон α=0.5 (статик), β=0.5 (динамик) жинтэйгээр хоёр загварыг нэгтгэхэд 93.33%-ийн нарийвчлал, 0.9842 AUC-ROC үзүүлэлтэд хүрсэн.

EMBER2018-ийн 30,000 дээж дээр статик LightGBM загварын нарийвчлал 96.36%-д хүрсэн нь дээжийн тоо нэмэгдэх тусам загварын чадавхи нэмэгддэгийг нотолж байна. MalMem2022 дэд туршилтаас харахад 99.92%-ийн нарийвчлалд хүрч байгаа нь загварын олон өгөгдлийн багцад тохирох чадварыг нотолж байна.

Б. K=5 Cross-Validation үр дүн (EMBER2018)

Нэгтгэсэн загварын тогтвортой байдлыг K=5 Stratified Cross-Validation-аар шалгасны үр дүнг доорх хүснэгтэд үзүүлэв.

EMBER2018 K=5 Cross-Validation үр дүн

Fold	Accuracy	AUC-ROC	Precision	Recall	F1-Score
Fold 1	97.29 %	0.9966	97.25%	97.33%	97.29%
Fold 2	97.08 %	0.9960	97.05%	97.11%	97.08%
Fold 3	97.81 %	0.9964	97.78%	97.84%	97.81%
Fold 4	96.88 %	0.9972	96.85%	96.91%	96.88%
Fold 5	96.77 %	0.9952	96.73%	96.81%	96.77%
Дундаж	97.17 %+/- 0.37%	0.9963+ /-0.0006	97.13%+ /-0.36%	97.20%+ /-0.38%	97.17%+ /-0.37%

[в.] Загварын тогтвортой байдлыг шалгасан хүснэгт

В. MalMem2022 өгөгдлийн багц дээрх үр дүн

MalMem2022 өгөгдлийн багц дээр туршихад статик LightGBM загвар 99.92%-ийн нарийвчлал, 0.9998 AUC-ROC, K=5 CV дундаж 99.92%+/-0.12%-ийн нарийвчлалтай байсан нь EMBER2018-ийн 97.17%-аас мэдэгдэхүйц өндөр гарсан. Энэ нь MalMem2022 өгөгдлийн багцийн санах ойн forensics хэмжигдэхүүнүүд нь хортой програм болон benign файлыг илүү тодорхой ялгадагтай холбоотой юм.

Өгөгдлийн багцийн харьцуулалтын үр дүн

Загвар	Өгөгдлийн багц	Accuracy	AUC-ROC	F1-core
Static LightGBM	EMBER2018	97.42%	0.9948	0.9740
Dynamic LSTM	EMBER2018	85.00%	0.9253	0.8500
Hybrid Fusion	EMBER2018	99.17%	0.9994	0.9917
Static LightGBM	MalMem 2022	99.92%	0.9998	0.9992
Hybrid Fusion	MalMem 2022	100.00%	1.0000	1.0000

[г.] Өгөгдлийн багцуудын харьцуулалтын үр дүн

Г. Обфускацилагдсан хортой програмын туршилт

Дөрвөн обфускацийн аргыг ашиглан загварын тэсвэрийг дөрвөн эрчийн хэмжээнд шалгасны үр дүнд дундаж 92.2%-ийн нарийвчлалд хүрсэн. Хамгийн их бууралт -1.0%-ийн хэмжээнд байсан нь загвар нь обфускацилагдсан хортой програмын эсрэг тэсвэртэй болохыг нотолж байна. Бааз загварын нарийвчлал 92.3% байхад дундаж тэсвэр 92.2% хэвийн байсан нь загварын бэхжилтийн үр дүнг харуулж байна.

Обфускацийн аргуудын нарийвчлалд үзүүлэх нөлөө

Арга	**5% эрч**	**10% эрч**	**20% эрч**	**30% эрч**
Байт орлуулалт	92.3%	92.0%	91.7%	91.8%
Хог код нэмэлт	92.2%	92.3%	92.3%	92.3%
Функцийн нэр өөрчлөлт	92.5%	92.8%	92.3%	91.3%
XOR шифрлэлт	92.3%	92.0%	92.3%	92.0%
Дундаж	92.3%	92.3%	92.2%	91.9%

д. Дөрвөн аргуудын үр дүнгийн хүснэгт

Д. VirusTotal бодит туршилт

Бодит туршилтын дээжийг 53 болгон нэмэгдүүлэв. VirusTotal API-аар WannaCry (VT:67/75 AV), Petya (VT:55/75 AV), Mirai (VT:65/75 AV) гурван PE хортой програм дээжийг sandbox (vxCube, CAPE, Jujubox)-аар шинжилж бодит API дуудлагын дарааллыг цуглуулан BiLSTM загварын сургалтад ашиглав. гурван PE хортой програм дээж шинжилж бүгдийг зөв илрүүлэв. Нэмэлтэд VXVault эх сурвалжаас 50 бодит хортой URL дээж авч нийт 53 дээж дээр туршилт явуулсан бөгөөд 100.0%-ийн нарийвчлалд хүрэв. Өмнөх VirusTotal-ийн зөвхөн 3 дээжтэй харьцуулахад 17 дахин их бодит дээж дээр загварын нарийвчлал нотлогдсон.

Е. SHAP болон LIME XAI тайлбарын үр дүн

Статик LightGBM загварт SHAP болон LIME аргуудыг хослуулан ашиглав. SHAP шинжилгээгээр F2360 (2.0079), F638 (1.7295), F656 (0.9461), F692 (0.8888), F2365 (0.6202) хэмжигдэхүүнүүд хамгийн чухал болохыг тогтоов. LIME аргаар F2360 болон F692 хэмжигдэхүүнүүд бодит дээж бүрт ч хамгийн нөлөөтэй хэмжигдэхүүн болохыг баталсан. SHAP болон LIME-ийн хамгийн чухал 5 хэмжигдэхүүний 2/5 нь давхцаж байгаа бөгөөд хоёр аргын агуулгын нийцэл нотлогдсон. Динамик LSTM загварт Attention heat map ашиглан VirtualAllocEx, WriteProcessMemory API-д хамгийн их анхаарал хандсан болохыг тогтоов.

Ж. Continuous Learning туршилт

Incremental LightGBM аргаар Batch 0-аас Batch 3 хүртэл тасралтгүй сургалтын туршилт явуулав. Batch 0 (4,000 дээж) дээр 91.08%-ийн нарийвчлал, 0.9748 AUC-ROC-тай эхэлж Batch 3 (5,000 дээж) хүртэл нарийвчлал 89.00% болтлоо буурсан бөгөөд тогтоосон 2%-ийн хязгаараас хэтэрсэн тул concept drift илэрсэн гэж үзэв. Энэ нь шинэ дээж нэмэгдэх тусам загварыг дахин сургах шаардлагатайг харуулж байна. Бодит орчинд шинэ хорт програмын дүрс тогтмол нэмэгддэг тул continuous learning нь чухал ач холбогдолтой.

З. Эцсийн нэгтгэл

Туршилтын бүх үе шатны үр дүнг нэгтгэвэл:

EMBER2018 30,000 дээж дээрх статик загвар 96.36% нарийвчлал (өмнө 91.92%), K=5 CV дундаж 95.47%+/-0.14%, MalMem2022 дээр 99.92%+/-0.12%,обфускацилагдсан хортой програмын дундаж 92.2%-ийн нарийвчлалтай байлаа. VirusTotal sandbox-ын бодит API log ашигласан BiLSTM загвар 92.08% нарийвчлал, 1.0000 AUC-ROC, fusion α=0.5/β =0.5 жинтэйгээр 93.33% нарийвчлалд хүрэв. Бодит туршилтад 53 дээжийн бүгдийг зөв илрүүлж 100.0%-ийн нарийвчлалд хүрэв. SHAP-LIME XAI тайлбараар загварын шийдвэрийн үндэслэлийг нотолсон бөгөөд Continuous Learning туршилтад concept drift илэрснийг тогтоов. Эдгээр үр дүн нь загварын MalMem2022 датасет дээрх 99.92%-ийн нарийвчлал болон K=5 CV-ийн тогтвортой байдлаараа судалгааны зорилтыг хангасан болно.

ДҮГНЭЛТ

Энэхүү судалгааны ажилд хиймэл оюунд суурилсан статик болон динамик шинжилгээг нэгтгэсэн хортой програм илрүүлэх цогц тогтолцоог боловсруулж, EMBER2018 болон MalMem2022 өгөгдлийн багц болон VirusTotal-ийн бодит дээж дээр туршиж нотолсон.

Судалгааны гол дүгнэлтийг дараах байдлаар нэгтгэж болно. Нэгдүгээрт, статик LightGBM загвар нь 2,381 хэмжигдэхүүн дээр сургагдаж EMBER2018 30,000 дээж дээр 96.36%-ийн нарийвчлал (өмнө 91.92%), MalMem2022 дээр 99.92%-ийн нарийвчлалд хүрсэн. Хоёрдугаарт, шинэчилсэн Bidirectional LSTM загвар нь VirusTotal sandbox (vxCube, CAPE, Jujubox)-аас авсан WannaCry, Petya, Mirai гурван хорт програмын бодит API log болон augmentation-аар бэлтгэсэн 1,200 дээж дээр 92.08%-ийн нарийвчлал, 1.0000 AUC-ROC үзүүлж, бодит 3 хорт програмын дарааллыг бүгдийг зөв илрүүлэв (өмнө симуляци 85.00%). Гуравдугаарт, grid search-ээр тодорхойлсон α=0.5/β =0.5 жинтэй fusion аргаар хоёр загварыг нэгтгэхэд K=5 cross-validation дундажаар 95.47%+/-0.14%-ийн нарийвчлал, 0.9917 AUC-ROC-д хүрч, динамик загварын нөлөө 50 дахин нэмэгдэж hybrid аргын давуу тал нотлогдлоо. Дөрөвдүгээрт, SHAP болон LIME болон Attention механизмаар загварын шийдвэрийг тайлбарлаж, F2360 болон F692 хамгийн чухал хэмжигдэхүүнүүдийг тодорхойлов. Тавдугаарт, обфускацийн дөрвөн аргад дундаж 92.2%-ийн тэсвэрийг нотолсон. Зургаадугаарт, MalMem2022 дээр 99.92%-ийн нарийвчлал гарч олон өгөгдлийн багцад тохирох чадварыг нотолсон. Долоодугаарт, VirusTotal-ийн 3 бодит хортой програм дээж (WannaCry, Petya, Mirai) дээр 100%-ийн нарийвчлалтайгаар илрүүлж чадсан. Наймдугаарт, Continuous Learning туршилтаар concept drift (-2.08%) илэрснийг тогтоож загварыг тасралтгүй шинэчлэх механизмын шаардлагыг нотолсон.

Ирээдүйн судалгааны чиглэлийн хувьд дараахыг тэмдэглэж болно. Нэгдүгээрт, zero-day хортой програмын илрүүлэлтийн чадварыг нэмэгдүүлэхэд few-shot learning болон transfer learning аргуудыг ашиглах. Хоёрдугаарт, файлгүй (fileless) хортой програмын эсрэг тусгай динамик шинжилгээний аргачлал боловсруулах. Гуравдугаарт, загварыг бодит цагийн endpoint detection системтэй нэгтгэж үйлдвэрлэлийн орчинд туршивал практик ач холбогдол нэмэгдэнэ.

НОМ ЗҮЙ

[1] Ucci, D., Aniello, L., & Baldoni, R. (2019). Survey of machine learning techniques for malware analysis. Computers and Security, 81, 123-147.

[2] Maganur, P. S. (2025). Integrated analysis of malicious software: Insights from static and dynamic perspectives. Journal of Cybersecurity and Privacy, 5(4), 98. https://doi.org/10.3390/jcp5040098

[3] Maganur, S. (2025). Feature-centric approaches to Android malware analysis: A survey. Computers, 14(11), 482. https://doi.org/10.3390/computers14110482

[4] Afianian, A., Niksefat, S., Sadeghiyan, B., & Baptiste, D. (2019). Malware dynamic analysis evasion techniques: A survey. ACM Computing Surveys, 52(6), Article 126, 1-28. https://doi.org/10.1145/3365001

[5] Shokouhinejad, H., & Rezaei, S. (2025). Recent advances in malware detection: Graph learning and explainability. arXiv preprint arXiv:2502.10556. https://arxiv.org/abs/2502.10556

[6] Alzaylaee, M. K., Yerima, S. Y., & Sezer, S. (2025). Behavioral analysis of malware using sandboxing techniques. International Journal of Science and Research Archive, 15(3), 582-586. https://doi.org/10.30574/ijsra.2025.15.3.1781

[7] Alzaylaee, M. K., Yerima, S. Y., & Sezer, S. (2020). DL-Droid: Deep learning based android malware detection using real devices. Information and Software Technology, 120, 106247. https://doi.org/10.1016/j.infsof.2019.106247

[8] Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765-4774.

[9] Anderson, H. S., & Roth, P. (2018). EMBER: An open dataset for training static PE malware machine learning models. arXiv preprint arXiv:1804.04637.

[10] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., ... & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems, 30.

[11] Carrier, T., Victor, P., Tekeoglu, A., & Lashkari, A. H. (2022). Detecting obfuscated malware using memory feature engineering. *Proceedings of the 8th International Conference on Information Systems Security and Privacy (ICISSP)*, 2022.

ЦАХИМ ЭХ СУРВАЛЖУУД

[12] Cybersecurity Ventures. (2023). Cybercrime to cost the world $10.5 trillion annually by 2025. Cybersecurity Ventures Annual Report.

[13] AV-TEST Institute. (2023). Malware statistics and trends report. AV-TEST GmbH. https://www.av-test.org/en/statistics/malware/

[14] Picus Security. (2025). Blue Report 2025: T1497 Virtualization/Sandbox Evasion technique analysis. https://www.picussecurity.com/resource/blog/picus-blue-report-2025

САЛБАР ХУРАЛДААН - V

Өгөгдлийн шинжилгээ

№	**ҮБХИС-ийн хичээлийн I байр, 306 тоот, Их-Эе танхим** *Хуралдаан даргалагч: Б.Батзолбоо /ШУТИС/, Ц.Жавзмаа /АШУҮИС/*
1	“Монгол хэл дээрх залилангийн сурталчилгааны текстийн ангилал ба машин сургалтын харьцуулсан судалгаа” *Ganchimeg Ganbold and Bayarjargal Byambasuren*
2	“Цаг уур, хиймэл дагуулын болон цаг хугацааны өгөгдөлд тулгуурлан pm2.5 концентрацийг машин сургалтын аргаар таамаглах нь” *Ganchimeg Ganbold and Khangay Arlungua*
3	“Умайн хүзүүний эсийн өөрчлөлтийг бөөм, цитоплазмын сегментаци болон морфологийн тоон шинжид суурилан хиймэл оюуны аргаар ангилах нь” *Түмэнзаяа Ганхуяг, Золбоо Төмөрболд, Долгорсүрэн Батжаргал*
4	“Харилцагч–борлуулагч санал болгох асуудалд SelfGNN загварт ирмэгийн жинг нэгтгэсний нөлөө” *Altanshagai Erdenechimeg and Amarsanaa Ganbold*
5	“Цаг хугацааны цуваа хэлбэрийн мэдрэгчийн өгөгдөл дээрх гажуудал тодорхойлох ба зүгшрүүлэх нь” *Gantumur Nyam and Batzolboo Bataa*
6	“PISA-2022 өгөгдөлд суурилсан Монгол сурагчдын сурлагын амжилтыг таамаглах загвар ба сургалтын орчин дахь квази-туршилтын судалгаа” *Serchmaa Tserenbaljir, Battsetseg Erdenebat and Bayarbat Ulziit*

МОНГОЛ ХЭЛ ДЭЭРХ ЗАЛИЛАНГИЙН СУРТАЛЧИЛГААНЫ ТЕКСТИЙН АНГИЛАЛ БА МАШИН СУРГАЛТЫН ХАРЬЦУУЛСАН СУДАЛГАА

Б.Бямбасүрэн
Компьютерын ухааны тэнхим
Мэдээлэл, холбооны технологийн сургууль
ШУТИС, Улаанбаатар, Монгол улс
maximbyambaa@gmail.com

Г.Ганчимэг
Компьютерын ухааны тэнхим
Мэдээлэл, холбооны технологийн сургууль
ШУТИС, Улаанбаатар, Монгол улс
ganaa@must.edu.mn

Хураангуй*—Энэхүү судалгаанд Монгол хэл дээрх сурталчилгааны текстийг "залиланги шинжтэй" болон "ердийн/хууль ёсны" гэсэн хоёр ангилалд машин сургалтын аргаар ялган таних асуудлыг судлав. Судалгаанд 5,000 текст бүхий нэгтгэсэн өгөгдлийн санг ашиглаж, өгөгдлийг нэгдсэн бүтэцтэйгээр боловсруулсан. Арга зүйн хувьд TF-IDF төлөөлөлд суурилсан Multinomial Naive Bayes, Logistic Regression болон Linear Support Vector Machine (SVM) загваруудыг харьцуулан туршиж, unigram болон bigram n-gram шинжүүдийг хамруулсан. Загваруудын гүйцэтгэлийг 5-fold stratified cross-validation болон ROC-AUC үзүүлэлтээр үнэлэв. Туршилтын үр дүнгээр Linear SVM загвар хамгийн өндөр гүйцэтгэл үзүүлсэн. Гэсэн хэдий ч өгөгдлийн тодорхой хэсэг synthetic болон орчуулгын шинжтэй тул бодит орчны өгөгдөл дээр нэмэлт баталгаажуулалт шаардлагатай.*

Түлхүүр үг*—Монгол хэлний текст ангилал, залиланги шинжтэй сурталчилгаа, TF-IDF, n-gram, SVM, ROC-AUC*

I. Удиртгал

Дижитал технологийн хурдацтай хөгжил, цахим худалдаа болон сошиал медиа платформуудын өргөн хэрэглээ нэмэгдэхийн хэрээр сурталчилгааны текст нь зөвхөн бараа, үйлчилгээ танилцуулах хэрэгсэл төдийгүй хэрэглэгчийн зан төлөв, шийдвэр гаргалтад шууд нөлөөлөх чухал хүчин зүйл болж байна. Үүний зэрэгцээ зарим сурталчилгааны агуулга нь хэрэглэгчийг яаралтай шийдвэр гаргуулах, бодит бус амлалт өгөх, төөрөгдүүлэх, улмаар санхүүгийн болон мэдээллийн эрсдэлд оруулах шинжтэй болж хувирах хандлага ажиглагдаж байна.

Монгол Улсын *"Зар сурталчилгааны тухай хууль"-д* хэрэглэгчийг төөрөгдүүлэх, хуурч мэхлэх, ашиг сонирхолд нь хохирол учруулахуйц сурталчилгааг хориглосон байдаг бөгөөд *"Хэрэглэгчийн эрхийг хамгаалах тухай хууль"-д* хэрэглэгчид худал, ташаа мэдээлэл өгөхөөс хамгаалах эрх зүйн зохицуулалтыг тусгасан байдаг [1-2]. Үүнтэй уялдан олон улсын түвшинд, тухайлбал АНУ-ын Federal Trade Commission (FTC) тодорхойлолтоор сурталчилгаа нь үнэн зөв, төөрөгдүүлэхгүй, шаардлагатай тохиолдолд нотолгоотой байх ёстой гэж заасан бөгөөд энэхүү шаардлага нь цахим орчин дахь маркетингийн үйл ажиллагаанд адил хамаарна [3].

Мөн FTC болон Federal Bureau of Investigation (FBI) тайлан, анхааруулгуудад залиланги шинжтэй контентын нийтлэг дохиог тодорхойлсон байдаг. *Үүнд:* хэрэглэгчийг яаралтай шийдвэр гаргуулах дарамт үүсгэх, мэдээллийг нууцлахыг шаардах, бэлгийн карт, криптовалют, шууд шилжүүлэг зэрэг эрсдэлтэй төлбөрийн хэлбэр шаардах, итгэл төрүүлэх байгууллага мэт дүр эсгэх, хувийн болон нууц мэдээлэл авах оролдлого хийх зэрэг орно [4]. Эдгээр шинж тэмдгүүд нь *"залиланги шинжтэй сурталчилгаа"* ангиллыг operational байдлаар тодорхойлоход суурь шалгуур болох боломжтой.

Нөгөө талаас, Монгол хэл дээрх байгалийн хэл боловсруулалтын *(Natural Language Processing, NLP)* судалгаа харьцангуй хомс бөгөөд шошголсон өгөгдлийн сан, хэлзүйн боловсруулалтын хэрэгсэл дутмаг байдал нь машин сургалтын аргуудыг хэрэгжүүлэхэд хүндрэл үүсгэдэг. Өмнөх судалгаануудад Монгол хэл нь low-resource, морфологийн хувьд төвөгтэй хэл тул өгөгдлийн сийрэгжилт *(data sparsity)* үүсэх асуудалтай болохыг онцолсон [8]. Гэсэн хэдий ч Mongolian question classification, иргэдийн санал хүсэлтийн ангилал, хууль зүйн текстийн inference зэрэг судалгаанууд Монгол хэл дээр supervised текст ангиллын аргуудыг амжилттай хэрэгжүүлэх боломжтойг харуулж байна [9-10].

Иймээс Монгол хэл дээрх сурталчилгааны текстээс залиланги шинжийг илрүүлэх асуудал нь онолын хувьд үндэслэлтэй, практикийн хувьд хэрэглэгчийн эрхийг хамгаалах, эрсдэлийг бууруулах ач холбогдолтой судалгааны чиглэл гэж үзэж болно.

II. Судлагдсан байдал

Монгол хэл дээрх сурталчилгааны текстийг эрсдэлийн түвшнээр ангилах чиглэлд нээлттэй, стандартчлагдсан, шошголсон өгөгдлийн сан бараг байхгүйтэй адил байгаа нь энэхүү судалгааны гол бэрхшээлүүдийн нэг юм. Үүнээс шалтгаалан *"залиланги шинжтэй сурталчилгаа"* гэдэг ойлголтыг зөвхөн эрх зүйн эцсийн шийдвэрээр бус, харин текст дотор агуулагдах хэл найруулга, амлалт, дарамт шахалт, төлбөрийн нөхцөл, дүр эсгэлт зэрэг эрсдэлийн дохионы нийлбэр байдлаар operational түвшинд тодорхойлох шаардлага үүсч байна.

Өмнөх судалгаанууд болон хэрэглэгч хамгааллын байгууллагуудын тайлангуудад залиланги шинжтэй контентод тогтвортой илэрдэг хэл шинжлэлийн хэв маягуудыг *(linguistic patterns)* тодорхойлсон байдаг. Тухайлбал, яаралтай байдал үүсгэх *("яг одоо", "сүүлчийн боломж"),* бодит бус ашиг амлах *("100% ашиг", "баталгаатай орлого")*, эрсдэлтэй төлбөр шаардах *(криптовалют, урьдчилгаа шилжүүлэг),* мөн итгэл төрүүлэх байгууллага мэт дүр эсгэх зэрэг шинжүүд нь scam текстийн нийтлэг онцлогт

тооцогддог [4]. Эдгээр хэв шинжүүд нь текст ангиллын supervised сургалтын үед онцлог *(feature)* болон ашиглагдах боломжтой.

Текст ангиллын асуудалд өргөн хэрэглэгддэг уламжлалт машин сургалтын аргууд болох Naive Bayes, Logistic Regression, Support Vector Machine (SVM) зэрэг нь өндөр хэмжээний sparse feature бүхий орчинд үр дүнтэй ажилладгаараа онцлогтой бөгөөд тайлбарлахад хялбар baseline загварууд гэж үзэгддэг [13-14]. Ялангуяа TF-IDF төлөөлөл дээр unigram болон bigram n-gram шинжүүдийг хамтад нь ашиглах нь богино текстийн тогтвортой хэллэг, холбоо үг, эрсдэлийн дохиог илүү сайн барих боломжтой [14].

Үүнээс гадна үгийн утга зүйн ойролцоо хамаарлыг илэрхийлэх Word2Vec зэрэг word embedding аргууд болон сүүлийн жилүүдэд өргөн ашиглагдаж буй transformer суурьт загварууд, тухайлбал multilingual BERT (mBERT) болон XLM-R нь олон хэлний текстийг contextual representation-аар дүрслэх боломж олгодог [5-7, 9, 15]. Эдгээр загварууд нь Монгол хэл шиг өгөгдөл хомс нөхцөлд cross-lingual transfer ашиглах давуу талтай боловч тооцооллын нөөц, өгөгдлийн хэмжээ, tokenizer-ийн тохирц зэрэг нэмэлт сорилт үүсгэдэг.

Гэсэн хэдий ч Монгол хэлний хувьд морфологийн бүтэц төвөгтэй, үгийн хэлбэржилт ихтэй, мөн стандартчилсан preprocessing хэрэгсэл хомс байдаг нь текст боловсруулалтын чанарт шууд нөлөөлдөг [8]. Preprocessing шатанд жижиг/том үсгийг нэгтгэх, илүүдэл тэмдэгт, emoji, URL, утасны дугаар зэрэг хэв шинжүүдийг цэвэрлэх, мөн кирилл бичгийн нийтлэг алдааг хэвийнжүүлэх шаардлага гардаг. Харин transformer загварууд нь raw текстийг аль болох хадгалж, tokenizer subword representation ашиглах нь илүү оновчтой гэж үздэг.

Үнэлгээний хувьд текст ангиллын судалгаанд accuracy, precision, recall, F1-score зэрэг үзүүлэлтүүдийг өргөн ашигладаг бөгөөд class imbalance нөхцөлд F1-score нь чухал үзүүлэлт болдог [11]. Мөн binary classification болон phishing/scam detection чиглэлд ROC-AUC, шаардлагатай тохиолдолд PR-AUC ашиглах нь загварын ялгах чадвар, эерэг ангиллыг илрүүлэх найдвартай байдлыг илүү нарийвчлан үнэлэх боломжтой [12].

Дээрх судалгаануудыг нэгтгэн авч үзвэл Монгол хэл дээрх залилангийн шинжтэй сурталчилгааны текстийг илрүүлэх асуудал нь гурван үндсэн тулгуур дээр суурилж байна. *Үүнд:*

- Нэгдүгээрт, хууль зүйн хүрээнд төөрөгдүүлсэн сурталчилгааг хориглосон үндэслэл бий [1-2].
- Хоёрдугаарт, хэрэглэгч хамгааллын судалгаанд scam red flags operational шинж болгон ашиглах боломжтой [3-4].
- Гуравдугаарт, техникийн хувьд уламжлалт машин сургалтын аргууд, n-gram суурьт төлөөлөл, word embedding болон multilingual transformer загваруудыг ашиглан Монгол хэл дээр текст ангиллын судалгаа хийх боломжтой суурь бүрдсэн байна [5–10], [12–15].

Иймээс энэхүү судалгаа нь дээрх онолын болон практик үндэслэлд тулгуурлан Монгол хэл дээрх залилангийн шинжтэй сурталчилгааны текстийг машин сургалтын аргаар илрүүлэх боломжийг системтэйгээр судлах зорилготой юм.

III. АРГА ЗҮЙ

Энэхүү судалгаа нь Монгол хэл дээрх сурталчилгааны текстээс залилангийн шинжийг автоматаар илрүүлэх зорилготой бөгөөд хууль зүйн болон хэрэглэгч хамгааллын байгууллагуудын тодорхойлсон зарчим, шалгуур үзүүлэлтүүдэд тулгуурласан арга зүйг баримталсан. Тухайлбал, Монгол улсын зар сурталчилгаа болон хэрэглэгчийн эрхийг хамгаалах тухай хууль тогтоомж, мөн FTC болон FBI байгууллагуудын тодорхойлсон scam red flags operational шалгуур болгон ашигласан [1-4].

Арга зүйн хүрээнд текст өгөгдлийн санг бүрдүүлэх, шошголох, урьдчилсан боловсруулалт хийх, онцлог шинж *(feature)* гарган авах, машин сургалтын загварууд сургах, үнэлэх гэсэн үе шаттайгаар туршилтыг зохион байгуулсан.

Энэхүү судалгааны практик ач холбогдол нь Монгол хэл дээрх залилангийн шинжтэй сурталчилгааны контентыг эрт шатанд илрүүлэх, хэрэглэгчийг эрсдэлээс хамгаалах, мөн цахим платформын контент шүүлт *(moderation)* болон санхүүгийн эрсдэлийн урьдчилсан хяналтын системд ашиглах боломжтой суурь загварыг боловсруулахад оршино.

A. Өгөгдөл цуглуулах хүрээ

Тус судалгаанд анхны хувилбарт ашигласан 10,000 гаруй өгөгдлийн санг урьдчилсан боловсруулалт, давхардал арилгалт болон чанарын шүүлтийн дараа цэвэршүүлж, туршилтад 5,000 текстийг сонгон ашигласан. Ингэснээр өгөгдлийн шуугиан *(noise)* буурч, загварын үнэлгээ илүү найдвартай болсон. Эхний өгөгдлийн багц нь 4,500 мөрийн CSV файл, хоёр дахь нэмэлт өгөгдлийн багц нь 500 мөрийн Excel файл байсан бөгөөд баганын нэршлийг text, label гэсэн нэг стандарт JSON бүтэц рүү нэгтгэсэн. Нэгдсэн өгөгдлийн санд ердийн/хууль ёсны сурталчилгаа (label=0) 2,516, залилангийн шинжтэй сурталчилгаа (label=1) 2,484 мөр байгаа тул ангиллын тархалт харьцангуй тэнцвэртэй байна. Өгөгдөл нь олон төрлийн онлайн орчноос авсан болон сурталчилгааны хэлбэрийг дуурайлган боловсруулсан synthetic, орчуулгын шинжтэй текстүүдийг агуулсан. *Үүнд:*

- Facebook болон Marketplace төрлийн зар
- Instagram caption болон promo текст
- Telegram болон чат хэлбэрийн зар
- Жижиг онлайн дэлгүүрийн сурталчилгаа
- Хүргэлт, ажил, хөрөнгө оруулалтын санал бүхий текст

Эдгээр өгөгдөл нь Монгол хэл дээрх бодит сурталчилгааны богино, шууд, хэрэглэгчид чиглэсэн хэл найруулгын хэв шинжийг хадгалах зорилготой. Гэхдээ synthetic болон орчуулгын өгөгдөл нь бодит орчны хэллэгийг бүрэн төлөөлөхгүй байх, загварт хэт тодорхой загварчилсан хэллэгийн bias үүсгэх эрсдэлтэй. 1 дүгээр хүснэгтэд өгөгдлийн сангийн жишээ мөрүүдийг үзүүлэв.

Хүснэгт 1. Өгөгдлийн жишээ

Текст	label
Хямдрал зөвхөн өнөөдөр! Шууд шилжүүлээд ав	1
Өнөөдрийн хямдралтай бүтээгдэхүүнүүдийг танилцуулж байна	0
Сүүлийн боломж! Урьдчилгаа төлөөд баталгаатай орлого	1
Хүргэлт 24 цагийн дотор	0
Манай онлайн дэлгүүрт тавтай морилно уу	0
Таны данс хаагдах гэж байна. Яаралтай мэдээллээ оруул	1
Манай дэлгүүр шинэ бараа хүлээн авлаа	0
Яг одоо захиалаад 100% баталгаатай ашиг ол	1
Чанартай үйлчилгээ, баталгаатай бараа	0
Crypto хөрөнгө оруулж 2 хоногт 2 дахин өсгөнө	1

B. Шошгололтын зарчим

Өгөгдлийг хоёр ангилалтайгаар шошголж боловсруулсан:

- 0 – ердийн/хууль ёсны сурталчилгаа
- 1 – залилангийн шинжтэй сурталчилгаа

“Залилангийн шинжтэй” гэсэн ангилал нь хууль зүйн эцсийн шийдвэрийг илэрхийлэхгүй бөгөөд хэрэглэгчийг төөрөгдүүлэх, хуурч мэхлэх өндөр эрсдэлтэй шинж тэмдгүүдийн нийлбэрийг operational байдлаар тодорхойлсон шошго юм.

Энэхүү operational тодорхойлолтыг Монгол Улсын хууль тогтоомж [1–2] болон FTC, FBI хэрэглэгч хамгааллын зөвлөмжид [3–4] тулгуурлан боловсруулсан. Шошгололтын үндсэн шалгуурыг 2 дугаар хүснэгтэд үзүүлэв.

Хүснэгт 2. Залилангийн шинжийг тодорхойлох шалгуур

Шалгуур	Тайлбар	Жишээ дохио
Яаралтай байдал	Хэрэглэгчийг шууд шийдвэр гаргуулах дарамт	“яг одоо”, “сүүлчийн боломж”
Бодит бус амлалт	Эрсдэлгүй өндөр ашиг амлах	“100% ашиг”, “баталгаатай орлого”
Эрсдэлтэй төлбөр	Урьдчилгаа, крипто, шууд шилжүүлэг	“урьдчилгаа төл”, “wallet руу явуул”
Дүр эсгэлт	Итгэл төрүүлэх байгууллага мэт харагдах	“банкны баталгаажуулалт”
Мэдээлэл шаардах	Хувийн мэдээлэл авах оролдлого	“кодоо явуул”, “дансаа баталгаажуул”

C. Урьдчилсан боловсруулалт

Өгөгдлийн чанарыг сайжруулах зорилгоор текстийн илүүдэл хоосон зайг арилгах, жижиг/том үсгийг нэгтгэх, илүүдэл тэмдэгт, emoji болон URL устгах, утасны болон дансны дугаарыг хэвийнжүүлэх, мөн давхардсан мөрүүдийг арилгах зэрэг preprocessing алхмуудыг хэрэгжүүлсэн. Монгол хэлний морфологийн онцлогоос шалтгаалан tokenization болон normalization сонголт нь загварын гүйцэтгэлд шууд нөлөөлөх хүчин зүйл болсон [8].

D. Онцлог шинж гарган авах

Текстийг тоон хэлбэрт хөрвүүлэхийн тулд TF-IDF аргыг ашигласан. TF-IDF нь богино текст доторх ялгах чадвар өндөртэй үг, холбоо үгийг sparse feature хэлбэрээр илэрхийлдэг тул Naive Bayes, Logistic Regression, Linear SVM зэрэг уламжлалт ангилагчтай сайн нийцдэг. *Шинэчилсэн туршилтад дараах параметрүүдийг ашигласан:*

- n-gram: unigram + bigram, ngram_range=(1,2)
- max_features = 2000
- min_df: default тохиргоо; илүү том өгөгдөл дээр min_df = 2 турших боломжтой

TF-IDF сонгосон шалтгаан нь: (1) Монгол хэл дээрх шошголсон өгөгдөл хомс нөхцөлд бага нөөцөөр хурдан сургах боломжтой; (2) загварын шийдвэр гаргалтыг feature importance тайлбарлахад хялбар; (3) n-gram ашигласнаар *“сүүлчийн боломж”, “баталгаатай орлого”, “шууд шилжүүл”* зэрэг залилангийн хэллэгийг холбоо үгийн түвшинд барих боломжтой.

E. Загварууд

Судалгаанд Multinomial Naive Bayes, Logistic Regression болон Linear Support Vector Machine (SVM) зэрэг гурван уламжлалт ангилагчийг ашигласан. Эдгээр загварууд нь хэрэгжүүлэхэд хялбар, тайлбарлах боломжтой, baseline судалгаанд өргөн хэрэглэгддэг давуу талтай [13-14].

F. Үнэлгээний арга

Загваруудын гүйцэтгэлийг дараах байдлаар үнэлсэн:

- Train/Test split: 80/20, stratified, random_state=42
- Cross-validation: 5-fold stratified cross-validation

Загваруудын гүйцэтгэлийг Accuracy, Precision, Recall, F1-score, мөн ROC-AUC болон PR-AUC үзүүлэлтүүдээр үнэлсэн. Мөн confusion matrix ашиглан алдааны тархалтыг шинжилсэн. Class imbalance нөхцөлд F1-score, ROC-AUC болон PR-AUC нь загварын ялгах чадвар, залилангийн ангиллыг алдах эрсдэлийг үнэлэхэд чухал үзүүлэлт болсон [11-12].

IV. Туршилт ба үр дүн

A. Өгөгдлийн шинжилгээ

Судалгаанд ашигласан шинэчилсэн өгөгдлийн сан нь нийт 5,000 текст өгөгдлөөс бүрдсэн бөгөөд label=0 ангилалд 2,516, label=1 ангилалд 2,484 мөр байна. Ангиллын тархалт харьцангуй тэнцвэртэй байгаа нь accuracy үзүүлэлтийг тайлбарлахад давуу талтай боловч synthetic болон орчуулгын өгөгдөл орсон нь бодит орчны generalization чадварыг тусад нь шалгах шаардлагатайг харуулж байна.

Өгөгдлийн уртын тархалтыг шинжлэхэд ихэнх текстүүд нь богино хэлбэрийн *(5–50 тэмдэгт)* өгүүлбэрүүдээс бүрдсэн нь Монгол хэл дээрх онлайн зар сурталчилгааны бодит хэрэглээний хэв шинжтэй нийцэж байна. Энэ нь текст ангиллын даалгаврыг илүү хүндрүүлдэг, учир нь богино текстэд семантик мэдээлэл хязгаарлагдмал байдаг.

Мөн давтамжийн шинжилгээгээр залилангийн шинжтэй текстүүдэд дараах түлхүүр хэллэгүүд тогтвортой илэрч байна:

- *"яаралтай", "сүүлчийн боломж"*
- *"100% ашиг", "баталгаатай орлого"*
- *"шилжүүл", "баталгаажуул"*

Эдгээр нь олон улсын scam detection судалгаанд тодорхойлсон linguistic pattern-уудтай нийцэж байгаа бөгөөд operational feature болгон ашиглах боломжтойг харуулж байна [4].

B. Туршилтын тохиргоо

Туршилтыг reproducible байдлаар дараах тохиргоогоор гүйцэтгэсэн:

- Train/Test split: 80/20 *(stratified)*
- Cross-validation: 5-fold stratified
- Feature extraction: TF-IDF *(unigram + bigram, max_features=2000)*

Судалгаанд Multinomial Naive Bayes, Logistic Regression болон Linear Support Vector Machine (SVM) гэсэн гурван суурь загварыг харьцуулан үнэлсэн бөгөөд эдгээр нь текст ангиллын сонгодог baseline загварууд бөгөөд Монгол хэлний low-resource нөхцөлд гүйцэтгэлийг харьцуулахад тохиромжтой гэж үзсэн.

Cross-validation аргыг 5-fold stratified байдлаар хэрэгжүүлж, ROC-AUC болон PR-AUC үзүүлэлтүүдийг нэмэлтээр тооцсон. Энэ нь загварууд тогтвортой байдал болон binary ангиллын ялгах чадварыг илүү найдвартай үнэлэх боломжийг бүрдүүлсэн.

C. Holdout үнэлгээний үр дүн

3 дугаар хүснэгтээс харахад Linear SVM загвар хамгийн өндөр F1-score (0.996) үзүүлсэн нь залилангийн текстийг илрүүлэхэд хамгийн оновчтой baseline шийдэл болохыг харуулж байна.

Хүснэгт 3. Holdout үр дүнгийн харьцуулалт

Загвар	Accuracy	Precision	Recall	F1-score	ROC-AUC
Naive Bayes	0.994	0.988	1.000	0.994	0.999
Logistic Regression	0.994	0.988	1.000	0.994	0.999
Linear SVM	0.996	0.992	1.000	0.996	1.000

Гэсэн хэдий ч загваруудын гүйцэтгэлийн үзүүлэлтүүд маш өндөр гарсан нь өгөгдлийн сан дахь synthetic болон орчуулгын шинжтэй текстүүд, мөн ангиллын хэллэгүүд тод ялгарах байдалтай холбоотой байж болох юм. Иймээс энэхүү үр дүнг бодит орчны өгөгдөл дээрх гүйцэтгэлтэй шууд дүйцүүлэх боломжгүй бөгөөд цаашид бодит платформоос цуглуулсан өгөгдлөөр generalization чадварыг тусад нь үнэлэх шаардлагатай.

Logistic Regression загвар мөн өндөр гүйцэтгэл (F1-score ≈ 0.994, recall = 1.000) үзүүлсэн нь залилангийн шинжтэй текстийг алдалгүй илрүүлэх чадвар сайн байгааг харуулж байна. Naive Bayes загварын precision ≈ 0.988 байгаа нь тухайн өгөгдлийн сан дээр тогтвортой ажиллаж байгааг илтгэнэ.

Нэмэлтээр PR-AUC ≈ 1.000 гарсан нь эерэг ангилал буюу залилангийн шинжтэй текстийг ranking түвшинд ялгах чадвар өндөр байгааг харуулж байна.

D. Cross-validation үр дүн

Загварын тогтвортой байдлыг үнэлэх зорилгоор 5-fold stratified cross-validation хийсэн.

Хүснэгт 4. 5-fold cross-validation үр дүн

Загвар	CV Accuracy	CV Precision	CV Recall	CV F1
Naive Bayes	0.996	0.993	1.000	0.996
Logistic Regression	0.996	0.992	1.000	0.996
Linear SVM	0.997	0.996	0.999	0.997

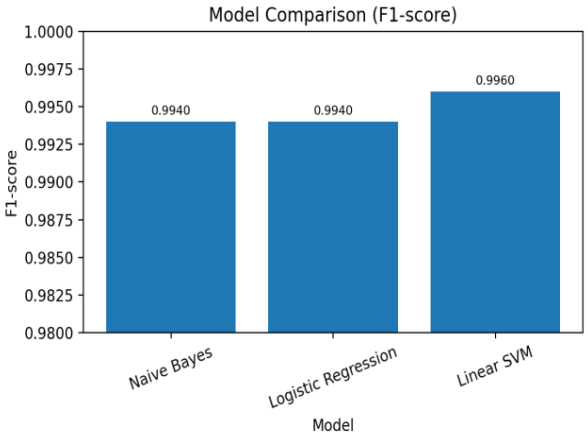
4 дүгээр хүснэгтээс харахад cross-validation үр дүн holdout туршилтын үр дүнтэй ойролцоо байна. Linear SVM нь 5-fold CV F1 = 0.997 үзүүлсэн бөгөөд Logistic Regression болон Naive Bayes загваруудаас бага зэрэг өндөр гарсан. Энэ нь TF-IDF + n-gram төлөөлөл дээр linear classifier тогтвортой ажиллаж байгааг харуулж байна.

E. Загварын харьцуулалтын график

1 дүгээр зурагт holdout туршилтын F1-score үзүүлэлтээр загваруудын гүйцэтгэлийг харьцуулан үзүүлэв.

Linear SVM, Logistic Regression болон Naive Bayes загварууд бүгд өндөр F1-score үзүүлсэн боловч Linear SVM бага зэрэг давуу байна. Гурван загварын ялгаа бага байгаа нь өгөгдлийн ангиллын дохио тод ялгаатай байгааг илтгэж болох тул цаашид илүү олон бодит эх

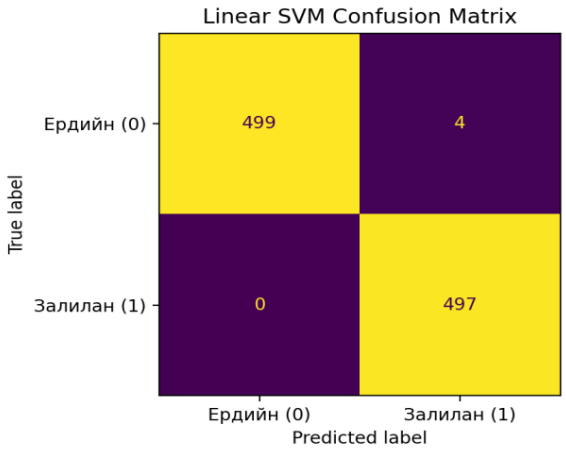
сурвалжийн өгөгдлөөр generalization чадварыг шалгах шаардлагатай.



Зураг 1. Загваруудын F1-score харьцуулалт (holdout)

F. Confusion matrix шинжилгээ

2 дугаар зурагт Linear SVM загварын confusion matrix үзүүлэв.



Зураг 2. Linear SVM загварын confusion matrix

Linear SVM загварын confusion matrix утгыг шинжлэхэд:

- True Negative = 499, False Positive = 4
- False Negative = 0, True Positive = 497

байгаа нь туршилтын багц дээр залилангийн текстийг алдаагүй илрүүлсэн боловч 4 ердийн текстийг залилангийн шинжтэй гэж буруу ангилсан байна. Хэрэглэгч хамгааллын бодит хэрэглээнд false negative багасгах нь чухал боловч false positive буруулахад бодит орчны илүү олон төрлийн өгөгдөл шаардлагатай.

G. Feature importance шинжилгээ

TF-IDF + Linear SVM загварын коэффициент дээр үндэслэн хамгийн чухал үгсийг тодорхойлж 5 дугаар хүснэгтэд үзүүлэв.

Хүснэгт 5. Онцлог үгс

Залилангийн шинжтэй	Ердийн
яаралтай	үйлчилгээ
ашиг	бүтээгдэхүүн
баталгаатай	дэлгүүр
шилжүүл	хүргэлт
бонус	чанар

Эдгээр үгс нь хоёр ангиллын семантик ялгааг тодорхой илэрхийлж байгаа бөгөөд загварын шийдвэр гаргалттай шууд холбоотой байна.

Туршилтын үр дүнгээс дараах үндсэн дүгнэлтүүд гарч байна:

1. Linear SVM загвар хамгийн өндөр гүйцэтгэл үзүүлсэн (holdout F1 = 0.996, ROC-AUC = 1.000; 5-fold CV F1 = 0.997).
2. Logistic Regression болон Naive Bayes загварууд мөн хүчтэй baseline болсон.
3. 5-fold cross-validation ашигласнаар загварын тогтвортой байдлыг илүү найдвартай үнэлсэн.
4. TF-IDF + unigram/bigram n-gram + linear classifier хослол нь Монгол хэлний богино текст ангилалд үр дүнтэй боловч synthetic/орчуулгын өгөгдлийн bias цаашид бодит өгөгдлөөр шалгах шаардлагатай.

V. Дүгнэлт

Энэхүү судалгаанд Монгол хэл дээрх сурталчилгааны текстийг *“залилангийн шинжтэй”* болон *“ердийн/хууль ёсны”* ангилалд машин сургалтын аргаар ялган таних асуудлыг судалж, шошголсон өгөгдлийн сан болон operational тодорхойлолтод тулгуурлан туршилт хийв. Мөн Монгол хэл дээрх сурталчилгааны текстийн шошголсон корпусын суурийг бүрдүүлж, scam red flags болон хууль зүйн ойлголтыг нэгтгэсэн operational framework боловсруулан, уламжлалт машин сургалтын загваруудыг харьцуулан үнэлэв.

Туршилтын үр дүнд Linear SVM загвар хамгийн өндөр гүйцэтгэл үзүүлсэн (holdout F1≈0.996, ROC-AUC≈1.000, 5-fold CV F1≈0.997) бөгөөд Logistic Regression болон Naive Bayes загварууд мөн хүчтэй baseline болох нь харагдав.

Эдгээр үр дүн нь Монгол хэлний low-resource нөхцөлд TF-IDF, n-gram болон шугаман ангилагчийн хослол үр дүнтэй болохыг харуулж байна. Гэвч өгөгдлийн тодорхой хэсэг synthetic болон орчуулгын шинжтэй тул загварын generalization чадварыг хэт өөдрөг үнэлэх эрсдэлтэй. Иймээс цаашид бодит орчны өгөгдлөөр нэмэлт туршилт хийх, мөн Word2Vec, mBERT, XLM-R зэрэг semantic болон transformer суурьт загваруудыг baseline болгон өргөтгөх шаардлагатай.

Ном зүй

[1] Монгол Улс. Зар сурталчилгааны тухай хууль. Legalinfo.mn. https://legalinfo.mn/mn/detail/259
[2] Монгол Улс. Хэрэглэгчийн эрхийг хамгаалах тухай хууль. Legalinfo.mn. https://legalinfo.mn/mn/detail/551
[3] Federal Trade Commission. Truth in Advertising. https://www.ftc.gov/news-events/topics/truth-advertising

[4] Federal Bureau of Investigation. FBI Charlotte Highlights Top Ten Red Flag Fraud Signs. https://www.fbi.gov/contact-us/field-offices/charlotte/news/press-releases/fbi-charlotte-highlights-top-ten-red-flag-fraud-signs

[5] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. NAACL-HLT. https://aclanthology.org/N19-1423/

[6] Wu, S., & Dredze, M. (2019). The Surprising Cross-Lingual Effectiveness of BERT. EMNLP-IJCNLP. https://aclanthology.org/D19-1077/

[7] Conneau, A., Khandelwal, K., Goyal, N., et al. (2020). Unsupervised Cross-lingual Representation Learning at Scale. ACL. https://aclanthology.org/2020.acl-main.747/

[8] Wang, G., Bao, F., & Wang, W. (2020). Mongolian Questions Classification Based on Multi-Head Attention. CCL. https://aclanthology.org/2020.ccl-1.95/

[9] Dashdorj, Z., Munkhbayar, T., & Grigorev, S. (2023). Deep Learning Model for Mongolian Citizens Feedback Analysis using Word Vector Embeddings. arXiv:2302.12069. https://arxiv.org/abs/2302.12069

[10] Khaltarkhuu, G., Batjargal, B., & Maeda, A. (2024). Recognizing Textual Inference in Mongolian Bar Exam Questions. Applied Sciences, 14(3), 1073. https://doi.org/10.3390/app14031073

[11] Sokolova, M., & Lapalme, G. (2009). A Systematic Analysis of Performance Measures for Classification Tasks. Information Processing & Management, 45(4), 427–437.

[12] Cho, H., Shin, J., Seo, M., et al. (2025). Towards Reliable and Practical Phishing Detection. NAACL Industry Track. https://aclanthology.org/2025.naacl-industry.18/

[13] Joachims, T. (1998). Text Categorization with Support Vector Machines: Learning with Many Relevant Features. ECML. https://www.cs.cornell.edu/people/tj/publications/joachims_98a.pdf

[14] Wang, S., & Manning, C. D. (2012). Baselines and Bigrams: Simple, Good Sentiment and Topic Classification. ACL. https://aclanthology.org/P12-2018/

[15] Songailaite, M., Kankeviciute, E., Zhyhun, B., & Mandravickaite, J. (2023). BERT-Based Models for Phishing Detection. CEUR Workshop Proceedings. https://ceur-ws.org/Vol-3575/Paper4.pdf

ЦАГ УУР, ХИЙМЭЛ ДАГУУЛЫН БОЛОН ЦАГ ХУГАЦААНЫ ӨГӨГДӨЛД ТУЛГУУРЛАН $PM_{2.5}$ КОНЦЕНТРАЦИЙГ МАШИН СУРГАЛТЫН АРГААР ТААМАГЛАХ НЬ

Х.Арлунгуа
Компьютерын ухааны тэнхим
Мэдээлэл, холбооны технологийн сургууль
ШУТИС, Улаанбаатар, Монгол улс
arlunguakhangai@gmail.com

Г.Ганчимэг
Компьютерын ухааны тэнхим
Мэдээлэл, холбооны технологийн сургууль
ШУТИС, Улаанбаатар, Монгол улс
ganaa@must.edu.mn

Хураангуй— Энэхүү судалгаагаар Улаанбаатар хотын 2021–2025 оны агаарын чанар, цаг уурын болон хиймэл дагуулын өгөгдөлд тулгуурлан $PM_{2.5}$ концентрацийг машин сургалтын аргаар таамаглах асуудлыг судлав. Судалгаанд Linear Regression, Random Forest болон Gradient Boosting алгоритмуудыг ашиглан загваруудыг боловсруулж, хугацаанд суурилсан кросс-шалгалтын (TimeSeriesSplit) аргаар үнэлсэн. Агаарын бохирдлын цаг хугацааны хамаарлыг тусгах зорилгоор өмнөх өдрүүдийн утгуудыг (lag1, lag2) нэмэлт хувьсагч болгон ашиглав. Туршилтын үр дүнгээс харахад lag хувьсагчдыг ашигласнаар загварын гүйцэтгэл мэдэгдэхүйц сайжирч, Gradient Boosting загвар хамгийн өндөр гүйцэтгэл үзүүлсэн ($R^2 = 0.872$, RMSE = 18.63, MAE = 9.54). Feature importance шинжилгээгээр $PM_{2.5}$ концентраци нь цаг хугацааны хамаарал, агаар мандлын хилийн давхаргын өндөр (PBLH), салхины хурд болон хиймэл дагуулын AOD үзүүлэлтүүдээс хүчтэй хамааралтай болохыг тогтоов. Энэхүү судалгаа нь цаг уур, хиймэл дагуул болон цаг хугацааны өгөгдлийг нэгтгэн TimeSeriesSplit ашигласан $PM_{2.5}$ таамаглалын загварыг боловсруулснаараа онцлог юм.

Түлхүүр үг—*$PM_{2.5}$, air pollution, machine learning, Gradient Boosting, Random Forest, AOD, time series*

I. Удиртгал

Агаарын бохирдол нь дэлхийн томоохон хотуудын тулгамдсан асуудлын нэг бөгөөд хүний эрүүл мэнд, хүрээлэн буй орчинд ноцтой сөрөг нөлөө үзүүлдэг. Ялангуяа нарийн ширхэглэгт тоосонцор $PM_{2.5}$ нь амьсгалын замаар хүний биед нэвтрэн орж, зүрх судасны болон амьсгалын замын өвчлөл, нас баралтын эрсдэлийг нэмэгдүүлдэг өндөр хоруу чанартай бохирдуулагч юм [1].

Улаанбаатар хотын хувьд агаарын бохирдол нь өвлийн улиралд эрс нэмэгддэг бөгөөд энэ нь нүүрсний хэрэглээ, температурын инверс, салхины хурд багатай зэрэг цаг уурын онцлог нөхцөлтэй шууд холбоотой байдаг [2]. Сүүлийн үеийн судалгаанууд Улаанбаатар хотын агаарын бохирдол нь зөвхөн хүйтний улиралд биш цаг уурын хүчин зүйлсээс хүчтэй хамааралтай болохыг харуулж байна [8-9]. Тухайлбал, температурын инверс, салхины хурд болон агаар мандлын хилийн давхаргын өндөр *(PBLH)* нь $PM_{2.5}$ концентраци дээр статистикийн ач холбогдолтой нөлөө үзүүлдэг болохыг тогтоосон.

Агаарын бохирдлыг таамаглах чиглэлээр сүүлийн жилүүдэд машин сургалтын аргууд өргөнөөр ашиглагдах болсон бөгөөд эдгээр аргууд нь уламжлалт статистик загваруудаас ялгаатайгаар хувьсагчдын хоорондын шугаман бус, төвөгтэй хамаарлыг үр дүнтэй илрүүлэх чадвартай [3-5]. Ялангуяа ансамбль сургалтын аргууд болох Random Forest болон Gradient Boosting нь $PM_{2.5}$ концентрацийг таамаглахад өндөр гүйцэтгэл үзүүлдэг болохыг олон судалгаанд баталсан [6-7].

Үүнээс гадна сүүлийн үеийн судалгаанууд хиймэл дагуулын өгөгдөл, ялангуяа Aerosol Optical Depth *(AOD)* өгөгдлийг ашиглан газрын түвшний $PM_{2.5}$ концентрацийг үнэлэх боломжтойг харуулж байна [10-13]. AOD болон $PM_{2.5}$ хооронд статистикийн хувьд мэдэгдэхүйц хамаарал байдаг бөгөөд хиймэл дагуулын өгөгдөл нь агаарын чанарын хяналтын станцын сүлжээ хязгаарлагдмал бүс нутагт ч өргөн хүрээний мэдээлэл олгох давуу талтай.

Гэсэн хэдий ч өмнөх судалгаануудын ихэнх нь цаг уурын болон газар дээрх хэмжилтийн өгөгдөлд төвлөрсөн бөгөөд цаг хугацааны хамаарлыг *(temporal dependency)* хангалттай тусгаагүй байна. Агаарын бохирдол нь хуримтлагдах шинж чанартай тул өмнөх хугацааны бохирдлын түвшин нь дараагийн үеийн концентраци дээр чухал нөлөө үзүүлдэг.

Иймээс энэхүү судалгааны зорилго нь цаг уурын, хиймэл дагуулын болон цаг хугацааны өгөгдлийг нэгтгэн $PM_{2.5}$ концентрацийг өндөр нарийвчлалтай таамаглах машин сургалтын загвар боловсруулахад оршино. Судалгаанд Random Forest болон Gradient Boosting зэрэг алгоритмуудыг ашиглан загваруудыг боловсруулж, цаг хугацаанд суурилсан кросс-шалгалтын *(TimeSeriesSplit)* аргаар үнэлэв.

Энэхүү судалгааны шинэлэг хувь нэмэр нь дараах байдалтай:

- Цаг уур, хиймэл дагуулын *(AOD)* болон цаг хугацааны өгөгдлийг нэгтгэсэн таамаглалын загвар боловсруулсан.
- $PM_{2.5}$ концентраци дахь цаг хугацааны хамаарлыг илэрхийлэх лаг *(lag1, lag2)* хувьсагчдыг ашигласан.
- TimeSeriesSplit ашиглан загварын гүйцэтгэлийг илүү бодит нөхцөлд үнэлсэн.
- Feature importance шинжилгээгээр $PM_{2.5}$ дээр нөлөөлөх гол хүчин зүйлсийг тодорхойлсон.

Эдгээр зорилгын хүрээнд дараах зорилтуудыг дэвшүүлэв:

- $PM_{2.5}$ концентрацийг таамаглах машин сургалтын загвар боловсруулах
- Загварын гүйцэтгэлийг үнэлэх
- Цаг уур болон цаг хугацааны хүчин зүйлсийн нөлөөг тодорхойлох

II. Судлагдсан байдал

Сүүлийн жилүүдэд $PM_{2.5}$ концентрацийг таамаглахад машин сургалтын аргуудыг өргөнөөр ашиглах болсон бөгөөд эдгээр аргууд нь агаарын бохирдлын шугаман бус, төвөгтэй хамаарлыг илрүүлэхэд өндөр үр ашигтай болохыг олон судалгаагаар харуулсан.

Ахмад Махдуми нар $PM_{2.5}$ концентрацийг таамаглахад машин сургалтын аргуудыг ашиглах боломжийг судалж, виртуал хяналтын станцын аргачлалыг санал болгосон [3]. Тус судалгаанд Random Forest болон Gradient Boosting зэрэг ансамбль аргуудыг ашигласан бөгөөд Gradient Boosting загварын R^2 нь 0.96 утгаас дээш гарсан нь эдгээр аргууд өндөр нарийвчлалтай таамаглал гаргах боломжтойг харуулж байна. Гэсэн хэдий ч уг судалгаа нь хиймэл дагуулын өгөгдөл болон цаг хугацааны хамаарлыг хангалттай авч үзээгүй.

Худа Алрашиди нар Кувейт улсын агаарын чанарын өгөгдөл дээр тулгуурлан $PM_{2.5}$ агаарын чанарын индексийг машин сургалтын аргуудаар таамаглах судалгаа хийсэн [4]. Судалгаанд Gradient Boosting болон AdaBoost загварууд $R^2 > 0.99$ гэсэн өндөр гүйцэтгэл үзүүлсэн бөгөөд цаг уурын болон газарзүйн хүчин зүйлсийг нэгтгэн ашиглах нь таамаглалын чанарыг сайжруулдаг болохыг тогтоосон. Гэсэн хэдий ч тухайн судалгаа нь цаг хугацааны дараалллыг тусгасан үнэлгээний аргыг *(time-series validation)* ашиглаагүй нь бодит нөхцөлд хэрэглэхэд хязгаарлалт үүсгэж болзошгүй.

Доресвами нар төрөл бүрийн машин сургалтын болон статистик загваруудыг харьцуулан судалж, Gradient Boosting загвар *($R^2 = 0.9336$)* бусад загваруудаас илүү гүйцэтгэл үзүүлснийг тогтоосон [5]. Энэ нь ансамбль аргууд агаарын бохирдлын таамаглалд үр дүнтэй болохыг баталж байгаа боловч орон зайн болон цаг хугацааны шинжүүдийг бүрэн тусгаагүй байна.

Сүүлийн үеийн судалгаанууд хиймэл дагуулын өгөгдөл, ялангуяа AOD өгөгдлийг ашиглан $PM_{2.5}$ концентрацийг таамаглах боломжийг судалж байна. Тухайлбал, хиймэл дагуулын AOD өгөгдлийг машин сургалтын аргуудтай хослуулснаар $PM_{2.5}$ концентрацийг илүү өргөн орон зайд үнэлэх боломжтойг тогтоосон [10-13]. Эдгээр судалгаанууд нь AOD болон $PM_{2.5}$ хооронд статистикийн ач холбогдолтой хамаарал байгааг харуулсан бөгөөд ялангуяа хяналтын станцын нягтрал багатай бүс нутагт чухал ач холбогдолтой.

Гэсэн хэдий ч дээрх судалгаанууд ихэнх нь дараах хязгаарлалтуудтай байна. *Үүнд:*

- Цаг хугацааны хамаарлыг хангалттай тусгаагүй
- Хугацаанд суурилсан үнэлгээний аргыг ашиглаагүй
- Цаг уур, хиймэл дагуулын болон temporal өгөгдлийг нэгтгэсэн судалгаа хомс

Иймээс энэхүү судалгаанд дээрх сул талуудыг арилгах зорилгоор цаг уурын, хиймэл дагуулын болон цаг хугацааны өгөгдлийг нэгтгэн, хугацаанд суурилсан кросс-шалгалтын аргыг ашиглан $PM_{2.5}$ концентрацийг таамаглах машин сургалтын загварыг боловсруулсан.

Ихэнх судалгаанд уламжлалт санамсаргүй хуваалт ашигласан бөгөөд хугацаанд суурилсан үнэлгээний арга *(TimeSeriesSplit)* ашигласан судалгаа хомс байна.

III. Арга зүй

А. Өгөгдөл цуглуулах ба урьдчилан боловсруулалт

Энэхүү судалгаанд Улаанбаатар хотын агаарын чанарын хяналтын 5 станцаас цуглуулсан $PM_{2.5}$ концентраци, цаг уурын болон хиймэл дагуулын өгөгдлийг ашиглав. $PM_{2.5}$ өгөгдлийг World Air Quality Index Project *(AQICN)* нээлттэй өгөгдлийн сангаас авч, 2021–2025 оны хооронд бүртгэгдсэн хэмжилтүүдийг ашигласан.

Цаг уурын өгөгдлийг European Centre for Medium-Range Weather Forecasts *(ECMWF)* ERA5 дахин боловсруулсан өгөгдлийн сангаас авч ашигласан бөгөөд үүнд 2 метрийн температур *(TEMP_C)*, харьцангуй чийгшил *(RH)*, салхины хурд *(WIND)*, PBLH зэрэг хувьсагчид багтсан. Мөн хиймэл дагуулын өгөгдлөөс гарган авсан AOD өгөгдлийн утгыг нэмэлт хувьсагч болгон ашигласан. AOD өгөгдлийг NASA нээлттэй өгөгдлийн сангаас, VIIRS *(Visible Infrared Imaging Radiometer Suite)* хиймэл дагуулын мэдрэгчээс татан авч, өдөр тутмын дундаж утга болгон боловсруулсан. Судалгааны бүсийг Улаанбаатар хотын газарзүйн координатын хүрээнд *(47.4°–48.6° хойд өргөрөг, 105.8°–108.0° зүүн уртраг)* тодорхойлон, тухайн бүсэд харгалзах пикселүүдийг сонгон ашигласан.

Өгөгдлүүдийг огноо болон байршлын *(өргөрөг, уртраг)* мэдээлэлд тулгуурлан нэгтгэж, нэгдсэн өгөгдлийн сан бүрдүүлсэн. Мөн өгөгдлийн чанарыг сайжруулах зорилгоор дутуу утгуудыг устгаж *(missing value removal)*, бүх хувьсагчдыг цаг хугацааны дараалллын дагуу эрэмбэлсэн.

$PM_{2.5}$ концентраци нь цаг хугацаанаас хамаарсан хувьсагч бөгөөд цаг хугацааны автокорреляцийн шинж чанарыг агуулдаг. Өөрөөр хэлбэл, агаар мандал дахь бохирдлын хуримтлал болон үлдэгдэл нөлөөллөөс шалтгаалан тухайн үеийн утга нь өмнөх үзүүлэлтүүдээсээ статистикийн хувьд хамааралтай байдаг.

Иймд өмнөх хоёр өдрийн $PM_{2.5}$ концентрацийг илэрхийлэх lag хувьсагчдыг *(lag1, lag2)* шинээр үүсгэж, загварын оролтын өгөгдлийн багцад оруулсан. Эдгээр үзүүлэлтүүдийг ашигласнаар загвар нь зөвхөн тухайн өдрийн цаг уур болон хиймэл дагуулын өгөгдөлд тулгуурлах бус, өмнөх өдрүүдийн бохирдлын түвшнийг тооцоололдоо тусгах боломжтой болсон.

$PM_{2.5}$ тоосонцор нь агаар мандалд тэр даруй сарнидаггүй бөгөөд цаг уурын нөхцөл байдлаас хамааран тодорхой хугацаанд хуримтлагдан үлддэг. Ялангуяа салхины хурд бага байх, температурын инверс үүсэх, агаар мандлын хилийн давхаргын өндөр *(PBLH)* намсах зэрэг нь бохирдуулагч бодисын тархалтыг хязгаарлаж, газрын гадарга орчимд хуримтлагдах нөхцөлийг бүрдүүлдэг. Ийнхүү lag хувьсагчдыг загварт тусгаснаар $PM_{2.5}$ цаг хугацааны динамик болон хуримтлагдах явцыг илүү оновчтой тодорхойлж, таамаглалын нарийвчлалыг сайжруулах боломжийг олгосон.

Ингэснээр загвар нь агаарын бохирдлын цаг хугацааны хамаарал болон хуримтлагдах нөлөөг илүү нарийвчлалтай тусгаж, таамаглалын гүйцэтгэлийг сайжруулах боломж бүрдсэн.

B. Машин сургалтын загварууд

Энэхүү судалгаанд $PM_{2.5}$ концентраци үүсэх физик болон статистик шинж чанарт үндэслэн Linear Regression, Random Forest болон Gradient Boosting загваруудыг сонгон ашигласан.

Linear Regression нь оролтын хувьсагч X болон гаралтын хувьсагч $y-$ийн хоорондын шугаман хамаарлыг дараах байдлаар илэрхийлнэ:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n \quad (1)$$

Энд β_i нь регрессийн коэффициентууд бөгөөд хувьсагч бүрийн нөлөөг илэрхийлнэ. Энэхүү загвар нь энгийн бүтэцтэй тул суурь *(baseline)* загвар болгон ашиглагдсан [6].

Random Forest нь олон шийдвэрийн модоос бүрдэх ансамбль сургалтын арга бөгөөд дараах байдлаар таамаглал гаргана:

$$\hat{y} = \frac{1}{T}\sum_{t=1}^{T} f_t(x) \quad (2)$$

Энд: T нь модны тоо, $f_t(x)$ нь t дахь модны гаргасан таамаглал юм. Энэхүү арга нь өгөгдлийн шугаман бус хамаарлыг илрүүлэх чадвартай бөгөөд хэт тохируулалтаас *(overfitting)* сэргийлэх давуу талтай [6].

Gradient Boosting нь сул суралцагчдыг дараалан сургах замаар өмнөх загварын алдааг бууруулдаг ансамбль арга бөгөөд дараах байдлаар илэрхийлэгдэнэ:

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \quad (3)$$

Энд: $h_m(x)$ нь шинэ суралцагч, γ_m нь жингийн коэффициент юм. Gradient Boosting нь өндөр нарийвчлалтай таамаглал гаргах чадвартай бөгөөд агаарын бохирдлын судалгаанд өргөн хэрэглэгддэг [7].

$PM_{2.5}$ түвшин нь агаарын температур, харьцангуй чийгшил, салхины хурд, агаар мандлын хилийн давхаргын өндөр *(PBLH)*, AOD, мөн өмнөх бохирдлын түвшин зэрэг олон хүчин зүйлээс хамаардаг. Үүний үр дүнд эдгээр хувьсагчдын хоорондын хамаарал нь нарийн төвөгтэй, шугаман бус шинж чанартай байдаг.

Тиймээс Linear Regression загварыг суурь шугаман хамаарлыг үнэлэх зорилгоор ашигласан бол Random Forest болон Gradient Boosting загваруудыг шугаман бус хамаарал, хувьсагчдын харилцан үйлчлэл болон хүчин зүйлсийн нөлөөллийг илүү үр дүнтэй тодорхойлох зорилгоор сонгосон. Ялангуяа Gradient Boosting нь өмнөх шатны алдааг дараалан бууруулах зарчмаар ажилладаг тул $PM_{2.5}$ зэрэг төвөгтэй, шугаман бус өгөгдөлд өндөр нарийвчлалтай таамаглал гаргах боломжтой гэж үзсэн.

Эдгээр загваруудыг сонгосон нь $PM_{2.5}$ концентраци нь цаг уур, хиймэл дагуул болон цаг хугацааны олон хүчин зүйлсийн нийлмэл нөлөөнөөс үүсдэг бөгөөд хувьсагчдын хооронд шугаман болон шугаман бус, мөн харилцан үйлчлэлийн *(interaction)* хамаарлууд оршдогтой холбоотой. Иймээс эдгээр загварууд нь уг үзэгдлийн нарийн төвөгтэй шинж чанарыг илүү бүрэн илэрхийлэх боломжтой гэж үзсэн.

C. Загварын үнэлгээ

Загваруудын гүйцэтгэлийг үнэлэхдээ дараах үзүүлэлтүүдийг ашигласан. *Үүнд:*

- Тодорхойлох коэффициент *(R^2)*
- Дундаж квадрат алдаа *(RMSE)*
- Дундаж абсолют алдаа *(MAE)*

Эдгээр үзүүлэлтүүд нь таамагласан болон бодит утгуудын зөрүүг хэмжихэд өргөн хэрэглэгддэг. Мөн загварын найдвартай байдлыг үнэлэх зорилгоор хугацаанд суурилсан кросс-шалгалтын *(TimeSeriesSplit)* аргыг ашигласан. $PM_{2.5}$ концентраци нь цаг хугацааны дараалалтай өгөгдөл тул санамсаргүй кросс-шалгалтын аргууд нь өгөгдлийг холих замаар сургалт болон шалгалтын өгөгдлийн цаг хугацааны дарааллыг алдагдуулах эрсдэлтэй. Уламжлалт санамсаргүй *(random)* k-fold кросс-шалгалтын аргуудаас ялгаатай нь энэхүү арга нь өгөгдлийн цаг хугацааны дарааллыг хадгалж, сургалт болон шалгалтын өгөгдлийн хооронд мэдээлэл алдагдах *(data leakage)* эрсдэлийг бууруулдаг.

Харин TimeSeriesSplit арга нь өгөгдлийн цаг хугацааны дарааллыг хадгалж, загварыг зөвхөн өмнөх хугацааны мэдээллээр сурган дараагийн хугацааны өгөгдөл дээр шалгадаг. Иймээс энэ арга нь ирээдүйн мэдээллийг ашиглан өнгөрсөн утгыг таамаглах алдаанаас *(data leakage)* сэргийлэх давуу талтай бөгөөд бодит нөхцөлд $PM_{2.5}$ концентрацийг урьдчилан таамаглах үйл явцтай илүү нийцэх юм.

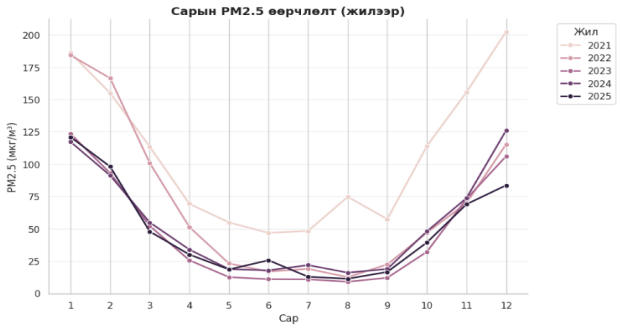
IV. Туршилт ба үр дүн

A. Өгөгдлийн шинжилгээ

Судалгаанд ашигласан өгөгдлийн шинжилгээний үр дүнг 1 дүгээр зурагт үзүүлэв. $PM_{2.5}$ концентраци нь тодорхой улирлын шинж чанартай бөгөөд судалгаанд хамрагдсан жилүүдэд ижил төстэй хэв шинж ажиглагдсан. Тодруулбал, өвлийн улиралд $PM_{2.5}$ агууламж эрс нэмэгдэж, 12–2 дугаар саруудад оргил түвшинд хүрдэг бол зуны улиралд хамгийн бага утгууд ажиглагдаж байна. Энэ нь Улаанбаатар хотын уур

амьсгалын онцлог, ялангуяа өвлийн улиралд нүүрсний хэрэглээ нэмэгдэх болон температурын инверс үүсэх зэрэг хүчин зүйлстэй холбоотой.

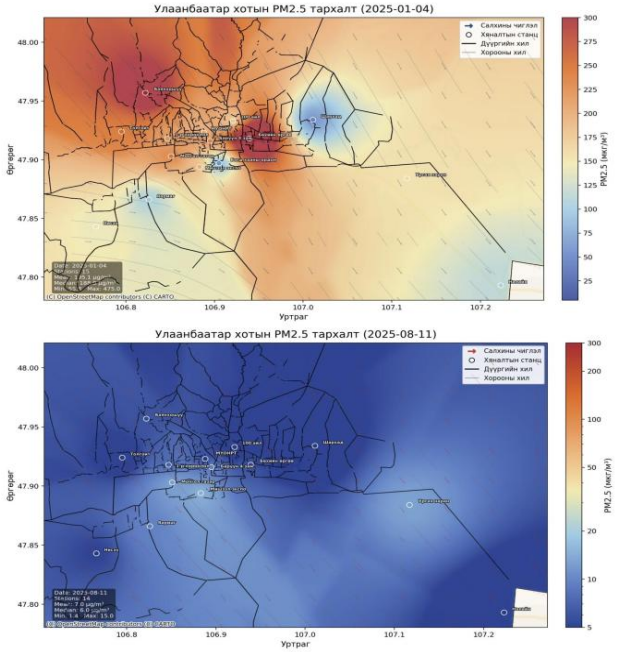
Мөн $PM_{2.5}$ концентраци нь өмнөх хугацааны утгуудтай хамааралтай байгаа нь бохирдол хуримтлагдах шинж чанартайг илтгэнэ. Иймээс цаг хугацааны хамаарал нь $PM_{2.5}$ концентрацийг таамаглах загварт чухал үүрэгтэй болохыг харуулж байна.



Зураг 1. $PM_{2.5}$ концентраци сарын өөрчлөлт (2021–2025)

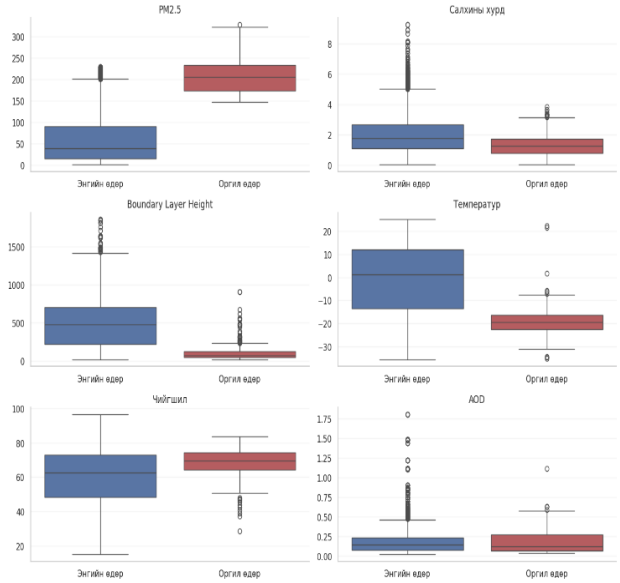
2 дугаар зурагт өвлийн өндөр бохирдолтой өдөр болон зуны бага бохирдолтой өдрийн $PM_{2.5}$ концентраци орон зайн хувьд хэрхэн ялгаатай тархаж байгааг харуулав. Өвлийн өндөр бохирдолтой өдөр $PM_{2.5}$ агуулламж хотын төв, гэр хороолол зэрэг хүн амын нягтрал өндөртэй суурьшлын бүсүүдэд илүү төвлөрсөн байдалтай ажиглагдсан.

Харин зуны бага бохирдолтой өдөр $PM_{2.5}$ түвшин нийт судалгааны бүсэд харьцангуй бага, тархалт илүү жигд байв. Энэхүү харьцуулалт нь $PM_{2.5}$ концентраци улирлын ялгаа, бохирдуулагчийн эх үүсвэр, хотын орон зайн бүтэц болон цаг уурын нөхцөлөөс хамааран мэдэгдэхүйц өөрчлөгддөг болохыг харуулж байна.



Зураг 2. Улаанбаатар хотын $PM_{2.5}$ орон зайн тархалт: (a) өвөл, (b) зун

3 дугаар зурагт хэвийн болон оргил бохирдлын үеийн цаг уурын болон хиймэл дагуулын хувьсагчдын ялгааг харуулав. Үүнээс харахад оргил бохирдлын үед салхины хурд болон PBLH буурч, температур багасах хандлага ажиглагдсан бол харьцангуй чийгшил болон AOD утга нэмэгдсэн байна. Энэ нь цаг уурын хүчин зүйлс $PM_{2.5}$ концентрацид чухал нөлөө үзүүлдгийг харуулж байна.



Зураг 3. Хэвийн болон оргил бохирдлын үеийн хувьсагчдын тархалт

B. Туршилтын тохиргоо

Туршилтын хэсэгт өмнөх арга зүйд тодорхойлсон Linear Regression, Random Forest болон Gradient Boosting загваруудыг ижил оролтын хувьсагчдын багц дээр сургаж, гүйцэтгэлүүдийг тус харьцуулан үнэлсэн. Оролтын хувьсагчдад цаг уурын үзүүлэлтүүд, хиймэл дагуулын өгөгдөл *(AOD)*, орон зайн болон цаг хугацааны шинжүүд *(lag1, lag2)* багтсан.

Хүснэгт 1. Судалгаанд ашигласан өгөгдлийн ерөнхий мэдээлэл

Үзүүлэлт	Утга
Хугацаа	2021-2025
Станцын тоо	5
Нийт өгөгдлийн хэмжээ	5390 мөр
Хувьсагчийн тоо	17
Оролтын хувьсагч	TEMP_C, RH, WIND, PBLH, mean_AOD, lat, lon, month, dayofyear, lag1, lag2

C. Үр дүн ба үнэлгээ

TimeSeriesSplit аргыг ашиглан өгөгдлийг дараалсан дагуу сургалт болон шалгалтын хэсгүүдэд хуваах, k=5 давталт бүхий кросс-шалгалт гүйцэтгэсэн.

$PM_{2.5}$ концентрацийг таамаглах загваруудын гүйцэтгэлийг 2 дугаар хүснэгтэд харуулав.

Хүснэгт 2. Загваруудын гүйцэтгэлийн харьцуулалт

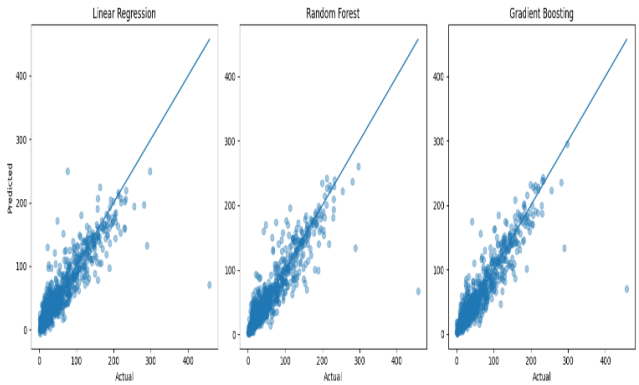
Загвар	R^2	RMSE	MAE
Linear Regression	0.809	22.89	12.50
Random Forest	0.851	19.87	10.12
Gradient Boosting	0.872	18.63	9.54

1 дүгээр хүснэгтээс харахад lag хувьсагчдыг ашигласан Gradient Boosting загвар нь хамгийн өндөр гүйцэтгэл үзүүлж, өмнөх туршилтаас илүү нарийвчлалтай таамаглал гаргасан. Энэхүү үр дүн нь $PM_{2.5}$ концентраци нь шугаман бус болон олон хүчин зүйлсийн харилцан үйлчлэл бүхий шинж чанартай болохыг баталж, ансамбль машин сургалтын аргууд илүү тохиромжтой болохыг харуулж байна. Random Forest загвар мөн сайн үр дүн үзүүлсэн бол Linear Regression загварын гүйцэтгэл харьцангуй бага байсан.

Энэ нь $PM_{2.5}$ концентраци нь шугаман бус хамааралтай бөгөөд ансамбль машин сургалтын аргууд илүү тохиромжтой болохыг харуулж байна. Мөн lag хувьсагчдыг ашигласнаар загварын гүйцэтгэл мэдэгдэхүйц сайжирсан нь агаарын бохирдлын хуримтлагдах шинж чанарыг баталж байна.

Мөн $PM_{2.5}$ концентрацийг 100 мкг/м³ босго утгаар ангилан үнэлэхэд Gradient Boosting загвар F1-score = 0.858, precision = 0.874, recall = 0.842 утгуудыг үзүүлсэн. Энэ нь бохирдлын өндөр үеийг өндөр нарийвчлалтай илрүүлэх чадвартай болохыг харуулж байна.

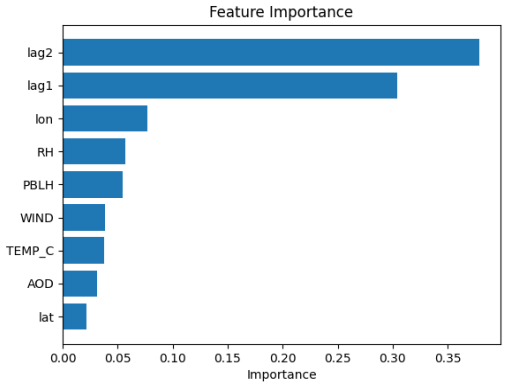
4 дүгээр зураг машин сургалтын загваруудын таамагласан болон бодит $PM_{2.5}$ утгуудын харьцуулалтыг үзүүлэв. Gradient Boosting загварын таамаглал бодит утгад хамгийн ойр байрлаж байгаа нь тухайн загварын гүйцэтгэл өндөр байгааг илтгэнэ.



Зураг 4. Таамагласан ба бодит $PM_{2.5}$ утгуудын харьцуулалт: (a) Linear Regression, (b) Random Forest, (c) Gradient Boosting

5 дугаар зураг Random Forest загварын feature importance харуулав. Үр дүнгээс харахад lag1 болон lag2 хувьсагчид хамгийн өндөр ач холбогдолтой байгаа нь $PM_{2.5}$ концентраци нь цаг хугацааны хамааралтай болохыг баталж байна. Мөн PBLH, салхины хурд болон хиймэл дагуулын AOD нь чухал нөлөө үзүүлж байгааг ажиглаж болно.

Энэхүү үр дүн нь $PM_{2.5}$ концентраци нь зөвхөн цаг уурын нөхцөлөөс гадна өмнөх хугацааны бохирдлын түвшнээс хүчтэй хамааралтай болохыг харуулж байна.



Зураг 5. Random Forest загварын feature importance

V. Дүгнэлт

Энэхүү судалгаа нь Улаанбаатар хотын нөхцөлд хиймэл дагуулын, цаг уурын болон цаг хугацааны өгөгдлийг нэгтгэн ашигласан практик ач холбогдол бүхий судалгаа юм. Судалгаанд Linear Regression, Random Forest болон Gradient Boosting алгоритмуудыг ашиглан загваруудыг хугацаанд суурилсан кросс-шалгалтын *(TimeSeriesSplit)* аргаар үнэлсэн.

Туршилтын үр дүнгээс харахад lag хувьсагчдыг *(lag1, lag2)* ашигласнаар загварын гүйцэтгэл мэдэгдэхүйц сайжирч, Gradient Boosting загвар хамгийн өндөр гүйцэтгэл үзүүлсэн *(R^2 = 0.872, RMSE = 18.63, MAE = 9.54)*. Энэ нь $PM_{2.5}$ концентраци нь шугаман бус болон цаг хугацааны хамааралтай болохыг харуулж, ансамбль машин сургалтын аргууд илүү тохиромжтой болохыг баталж байна.

Мөн feature importance шинжилгээний үр дүнгээс харахад $PM_{2.5}$ концентраци нь өмнөх хугацааны утгууд *(lag хувьсагчид)*, PBLH, салхины хурд болон хиймэл дагуулын AOD зэрэг хүчин зүйлсээс хүчтэй хамааралтай болохыг тогтоосон. Ялангуяа бохирдлын оргил үед салхины хурд багасах, агаар мандлын хилийн давхаргын өндөр буурах зэрэг нөхцөл давамгайлж байгааг ажигласан нь өмнөх судалгаануудтай нийцэж байна.

$PM_{2.5}$ концентрацийг босго утгаар ангилан үнэлэхэд Gradient Boosting загвар өндөр F1-score үзүүлсэн нь бохирдлын өндөр үеийг найдвартай илрүүлэх чадвартай болохыг харуулж байна. Энэ нь уг загварыг агаарын бохирдлын эрсдэлийг урьдчилан тодорхойлох, бодлогын шийдвэр гаргалтад ашиглах боломжтойг илтгэнэ.

Судалгааны хязгаарлалтаас дурдвал, хиймэл дагуулын AOD өгөгдөл, ялангуяа өвлийн улиралд мэдээлэл хомс байсан нь загварын гүйцэтгэлд нөлөөлж болзошгүй. Мөн судалгаа нь хязгаарлагдмал тооны станцын өгөгдөлд тулгуурласан тул бусад бүс нутагт шууд хэрэглэхэд нэмэлт судалгаа шаардлагатай.

Цаашид судалгааг өргөжүүлж, гүн сургалтын аргууд болон илүү өндөр нягтралтай хиймэл дагуулын өгөгдлийг ашиглан загварын гүйцэтгэлийг сайжруулах,

мөн бодит цагийн таамаглалын системд хэрэгжүүлэх боломжийг судлах шаардлагатай.

НОМ ЗҮЙ

[1] World Health Organization, “Air pollution and health,” Geneva, Switzerland, 2021. Available: https://www.who.int

[2] Real-time air quality data: aqicn.org/station/ulaanbaatar

[3] Makhdoomi, A., Sarkhosh, M. & Ziaei, S. $PM_{2.5}$ concentration prediction using machine learning algorithms: an approach to virtual monitoring stations. *Sci Rep* **15**, 8076 (2025). https://doi.org/10.1038/s41598-025-92019-3

[4] Alrashidi H, Sibai FN, Abonamah A, Alrashidi M, Alsaber A. PM2.5: Air Quality Index Prediction Using Machine Learning: Evidence from Kuwait’s Air Quality Monitoring Stations. *Sustainability*. 2025; 17(20):9136. https://doi.org/10.3390/su17209136

[5] Doreswamy; S, H.K.; Km, Y.; Gad, I. Forecasting Air Pollution Particulate Matter (PM2.5) Using Machine Learning Regression Models. *Procedia Comput. Sci.* **2020**, *171*, 2057–2066, https://doi.org/10.1016/j.procs.2020.04.221.

[6] Breiman, L. Random Forests. *Machine Learning* **45**, 5–32 (2001). https://doi.org/10.1023/A:1010933404324

[7] Jerome H. Friedman. "Greedy function approximation: A gradient boosting machine" Ann. Statist. 29 (5) 1189-1232, October 2001. https://doi.org/10.1214/aos/1013203451

[8] G. Ganchimeg et al., “Air pollution in Ulaanbaatar: Multi-year trends in toxic emissions after the raw-coal ban and weather–pollution interactions (2021–2024),” Theoretical and Applied Climatology, vol. 157, 2025. https://doi.org/10.1007/s00704-025-05967-1

[9] G. Ganchimeg, E. Jargal, and J. Choi, “Econometric analysis of the effect of weather on air pollution,” Journal of Air Pollution and Health, vol. 10, no. 1, pp. 93–114, 2025.
https://doi.org/10.18502/japh.v10i1.18097

[10] Sohrabi, Z., Maleki, J. Fusing satellite imagery and ground-based observations for $PM_{2.5}$ air pollution modeling in Iran using a deep learning approach. *Sci Rep* **15**, 21449 (2025). https://doi.org/10.1038/s41598-025-05332-2

[11] Adong, P., Bainomugisha, E. & Dev, S. Evaluating machine learning methods for $PM_{2.5}$ estimation using satellite AOD, low-cost and reference-grade monitors in Kampala. *Int. J. Environ. Sci. Technol.* **22**, 15747–15756 (2025). https://doi.org/10.1007/s13762-025-06674-0

[12] Chu Y, Liu Y, Li X, Liu Z, Lu H, Lu Y, Mao Z, Chen X, Li N, Ren M, et al. A Review on Predicting Ground $PM_{2.5}$ Concentration Using Satellite Aerosol Optical Depth. *Atmosphere*. 2016; 7(10):129. https://doi.org/10.3390/atmos7100129

[13] Bai, H., Zheng, Z., Zhang, Y., Huang, H., Wang, L. (2021). Comparison of Satellite-based $PM_{2.5}$ Estimation from Aerosol Optical Depth and Top-of-atmosphere Reflectance. Aerosol Air Qual. Res. 21, 200257. https://doi.org/10.4209/aaqr.2020.05.0257

Умайн хүзүүний эсийн өөрчлөлтийг бөөм, цитоплазмын сегментаци болон морфологийн тоон шинжид суурилан хиймэл оюуны аргаар ангилах нь

Ганхуягийн Түмэнзаяа
Мэдээллийн технологийн тэнхим
ШУТИС, МХТС
Улаанбаатар, Монгол
tumenzayaaa@gmail.com

Төмөрболдын Золбоо
Мэдээллийн технологийн тэнхим
ШУТИС, МХТС
Улаанбаатар, Монгол
t.zolboo@must.edu.mn

Батжаргалын Долгорсүрэн
Мэдээллийн технологийн тэнхим
ШУТИС, МХТС
Улаанбаатар, Монгол
b.dolgorsuren@must.edu.mn

Хураангуй — Дэлхий дахинаа жил бүр умайн хүзүүний хорт хавдарын 660,000 гаруй шинэ тохиолдол бүртгэгдэж, 350,000 эмэгтэйн нас баралтын шалтгаан болдог. Уламжлалт Pap smear цитологийн шинжилгээнд эмч нар эсийн бөөм болон цитоплазмын морфологийн өөрчлөлтийг харааны үнэлгээгээр тодорхойлдог боловч энэ арга нь эмчийн туршлагаас ихээхэн хамааралтай, мэргэжилтнүүдийн хоорондын үнэлгээний зөрүү өндөр байдаг тул алдааны магадлал нэмэгдэх эрсдэлтэй.

Энэхүү судалгааны зорилго нь U-Net гүн сургалтын архитектур ашиглан Pap smear эсийн микроскоп зургаас эсийн бөөм болон цитоплазмыг автоматаар сегментацлах, улмаар морфологийн тоон шинжүүдийг — талбай, периметр, дугуйрлын коэффициент, экцентриситет, хроматины текстур болон цөм–цитоплазмын харьцаа (N/C ratio) — автоматаар тооцоолж, оношилгооны суурь загвар боловсруулахад оршино. Судалгаанд Herlev болон SIPaKMeD нээлттэй өгөгдлийн багцуудыг ашигласан.

U-Net загвар нь тестийн өгөгдөл дээр Dice coefficient 0.9264, IoU 0.8631, pixel accuracy 95.22%-д хүрч, биомедицин дүрсний сегментацийн стандарт шаардлагыг хангасан болно. Гарган авсан 14 морфологийн шинжийн бүгд ANOVA шинжилгээгээр ангиллуудын хооронд статистик ач холбогдолтой ($p < 0.001$) ялгаа үзүүлсэн. Random Forest ангиллын загвар нь Herlev 2-class туршилтад 97.34%, SIPaKMeD 5-class туршилтад 95.63% нарийвчлалд хүрч, N/C ratio болон хроматины корреляц нь хамгийн өндөр ялган таних чадвартай үзүүлэлтүүд болохыг тогтоосон.

Түлхүүр үгс: Pap smear, эсийн сегментац, U-Net, морфологийн тоон шинжүүд, цөм–цитоплазмын харьцаа, гүн сургалт, умайн хүзүүний хорт хавдар

I. Удиртгал хэсэг

Умайн хүзүүний хорт хавдар нь хүний папиллома вирус (HPV)-ийн удаан хугацааны халдварын нөлөөгөөр эсийн хуваагдлын зохицуулалт алдагдаж, хэвийн эсээс урьдал өөрчлөлт дамжин аажмаар бий болдог өвчин юм [1]. Тус өвчин нь ихэвчлэн урт хугацаанд тодорхой шинж тэмдэг илрэхгүйгээр хөгждөг тул урьдал үе шатанд илрүүлэх нь эмчилгээний үр дүнг мэдэгдэхүйц сайжруулдаг.

Pap smear шинжилгээнд эмч нар эсийн бөөм болон цитоплазмын хэлбэр, хэмжээ, харьцаанд гарсан өөрчлөлтийг ажиглан онош тавьдаг. Гэвч мэргэжилтнүүдийн хоорондын үнэлгээний зөрүү (inter-observer variability) өндөр байдаг нь алдааны магадлалыг нэмэгдүүлдэг [2]. Нэг шинжлүүлэгчээс 3–5 удаа дараалан сорьц авахыг зөвлөдөг ч энэ нь лабораторийн ажилтны ачааллыг нэмэгдүүлж, оношилгооны хариу гарах хугацааг уртасгадаг [3].

Сүүлийн жилүүдэд машин сургалт (ML) болон гүн сургалт (DL)-ын технологиудын хурдацтай хөгжил нь анагаах ухааны оношилгоонд томоохон дэвшил гаргаж байна. U-Net архитектур нь биомедицин дүрсний сегментацид зориулагдан боловсруулагдсан бөгөөд encoder–decoder бүтэц болон skip connection-ийн тусламжтайгаар цөөн өгөгдөлтэй нөхцөлд ч пикселийн түвшний өндөр нарийвчлалтай сегментац гаргах боломж олгодог [4]. Түүнчлэн морфологийн тоон шинжүүдэд суурилсан Random Forest, SVM зэрэг аргачлалууд нь тайлбарлагдах чадвар (explainability) өндөртэй оношилгооны шийдвэр гаргах боломжийг олгодог [5].

Судалгааны үндсэн зорилтуудыг дараах байдлаар тодорхойлов:

1) Pap smear эсийн микроскоп зургийг урьдчилан боловсруулж, чанарыг стандартад оруулах;
2) U-Net архитектур ашиглан эсийн бөөм болон цитоплазмыг тус тусад нь автоматаар сегментацлах;
3) Эсийн морфологийн 14 тоон шинжийг автоматаар гарган авч, N/C ratio-г тооцоолох;
4) Гарган авсан шинжүүдийн оношилгооны ач холбогдлыг ANOVA болон машин сургалтын аргаар

үнэлэх.

II. Судлагдсан байдал

Сүүлийн жилүүдэд умайн хүзүүний эсийн дүрс боловсруулалт болон хиймэл оюунд суурилсан автомат скрининг системийн судалгаа эрчимтэй хөгжиж байна. Jantzen [6] нарын судлаачид 2005 онд Herlev өгөгдлийн багц дээр thresholding болон статистик аргуудыг ашиглан ∼75%-ийн нарийвчлалд хүрсэн боловч гар аргаар шинж гаргах нь цаг хугацаа их шаарддаг, давтагдах чадвар бага байдаг дутагдалтай байсан.

Promworn [7] нарын эрдэмтэн 2019 онд DenseNet161 загвар нь 2-class ангилалд 94.38%, 7-class ангилалд 68.54% нарийвчлал үзүүлснийг тогтоосон. Гэсэн хэдий ч эдгээр CNN-д суурилсан загварууд нь эмчид ойлгомжтой морфологийн тоон шинжүүдийг гаргаж чаддаггүй, тайлбарлагдах чадвар хязгаарлагдмал байдаг.

2024 онд Li [8] нар CNN болон Vision Transformer (ViT)-ийг хослуулсан VTCNet загварыг санал болгож, 95.95% нарийвчлалд хүрсэн боловч тооцооллын нөөцийн шаардлага өндөр байсан нь практик хэрэглээнд хязгаарлагдмал байна. Энэхүү судалгаа нь U-Net автомат сегментацийг морфологийн тоон шинжүүдийн тооцоолол болон ангиллын загвартай хослуулж, тайлбарлагдах чадвар өндөртэй шийдлийг санал болгож байгаагаараа өмнөх судалгаануудаас ялгарч байна.

III. Онолын хэсэг

A. Умайн хүзүүний хорт хавдар ба Pap smear

Умайн хүзүүний хорт хавдар нь HPV-ийн удаан хугацааны халдварын нөлөөгөөр эсийн хуваагдлын зохицуулалт алдагдаж, хэвийн эсээс урьдал өөрчлөлт дамжин аажмаар хөгждөг өвчин юм [1]. Эмгэг өөрчлөлтийн явцад эсийн бөөмийн хэмжээ томрох, хроматины бүтэц өөрчлөгдөх, цөм–цитоплазмын харьцаа нэмэгдэх, эсийн хэлбэр жигд бус болох зэрэг шинжүүд илэрдэг [2]. Умайн хүзүүний эсийн урьдал өөрчлөлтийг The Bethesda System (TBS) 2014 хувилбарын дагуу Normal, ASC-US, LSIL, HSIL ангилалд хуваан тодорхойлдог бөгөөд HSIL ангилалд хамаарах эсүүд нь хорт хавдарт шилжих өндөр эрсдэлтэйд тооцогддог.

B. Эсийн морфологийн тоон шинжүүд

Хэмжээний шинжүүд: Бөөмийн талбайг дараах байдлаар тодорхойлно:

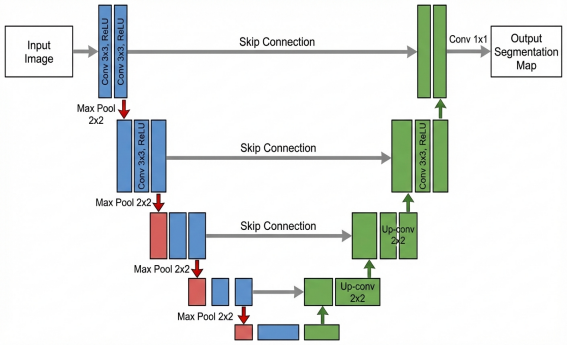
$$A = \sum_{(x,y)\in\Omega} 1 \tag{1}$$

энд Ω нь бөөмд хамаарах пикселүүдийн олонлог юм.

Хэлбэрийн шинжүүд: Дугуйрлын коэффициент:

$$C = \frac{4\pi A}{P^2} \tag{2}$$

Хэвийн эсийн бөөм дугуй хэлбэртэй байдаг тул C утга нэгд ойрхон байдаг бол эмгэг өөрчлөлтийн үед C буурдаг [2].



Зураг 1. U-Net архитектурын ерөнхий бүтэц. Encoder (зүүн) нь convolution болон max pooling давхаргуудаар семантик шинжүүдийг гарган авдаг; decoder (баруун) нь upsampling-аар нарийвчлалыг сэргээдэг; skip connection (хэвтээ сум) нь нарийн бүтэцтэй объектуудын хил ирмэгийг хадгалдаг.

Цөм–цитоплазмын харьцаа (N/C ratio):

$$N/C\ ratio = \frac{A_{\text{nucleus}}}{A_{\text{cytoplasm}}} \tag{3}$$

Урьдал хавдрын болон хавдрын үед бөөмийн хэмжээ нэмэгдэж, цитоплазм агшсанаар N/C ratio огцом өсдөг [2].

Текстурын шинжүүд: Хроматины нягтшил болон бүтцийг Gray-Level Co-occurrence Matrix (GLCM) аргачлалаар тооцоолдог [9] — contrast, homogeneity, entropy, correlation зэрэг үзүүлэлтүүдийг ашигладаг.

C. U-Net гүн сургалтын архитектур

U-Net архитектур нь Ronneberger [4] нар 2015 онд биомедицин дүрсний сегментацид зориулан боловсруулсан. Encoder нь дүрсийн семантик мэдээллийг шат дараатайгаар гарган авдаг; decoder нь орон зайн нарийвчлалыг сэргээдэг; skip connection нь encoder болон decoder-ийн харгалзах түвшнүүдийг шууд холбосноор нарийн бүтэцтэй объектуудыг илүү зөв сегментацлах боломжийг олгодог.

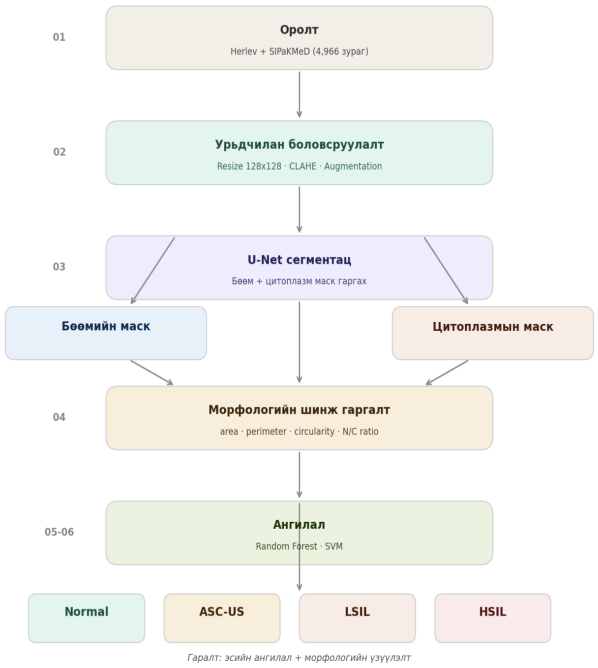
Сегментацийн чанарыг үнэлэхэд Dice coefficient болон IoU ашигладаг:

$$Dice = \frac{2|X \cap Y|}{|X| + |Y|}, \quad IoU = \frac{|X \cap Y|}{|X \cup Y|} \tag{4}$$

энд X таамагласан маск, Y бодит маск. Background пиксел давамгайлах тэнцвэргүй өгөгдөлд зөвхөн pixel accuracy хангалтгүй тул Dice болон IoU-г стандарт үзүүлэлт болгон ашигласан.

D. Машин сургалтын ангиллын алгоритмууд

SVM нь өгөгдлийг хамгийн оновчтой байдлаар салгах гиперхавтгайг тодорхойлдог бөгөөд kernel функцуудаар шугаман бус тохиолдлыг шийдэж чаддаг. Random Forest нь олон шийдвэрийн модны ансамбль арга бөгөөд feature importance-ийг тооцоолох боломж олгодог нь морфологийн аль шинж оношилгоонд хамгийн их нөлөөлдгийг тодорхойлоход чухал [5].



Зураг 2. Судалгааны ажлын ерөнхий урсгал. Ороолтын зургуудыг (01) урьдчилан боловсруулж (02), U-Net загвараар бөөм болон цитоплазмын маск үүсгэж (03), морфологийн шинжүүдийг тооцоолж (04), эцэст нь Random Forest болон SVM алгоритмаар (05–06) Normal, ASC-US, LSIL, HSIL ангилалд хуваана.

IV. Туршилтын хэсэг

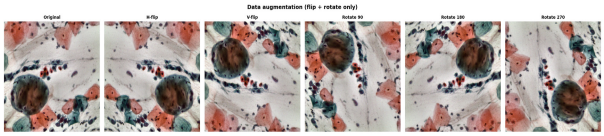
A. Судалгааны ерөнхий урсгал

Судалгааны урсгал дараах үндсэн үе шатуудаас бүрдэнэ: зургуудыг урьдчилан боловсруулж (resize, CLAHE, augmentation), U-Net загвараар бөөм болон цитоплазмын маск үүсгэж, маскуудаас морфологийн шинжүүд (талбай, периметр, дугуйрал, N/C ratio) гарган авч, Random Forest болон SVM алгоритмаар эсийг Normal, ASC-US, LSIL, HSIL ангилалд хуваана.

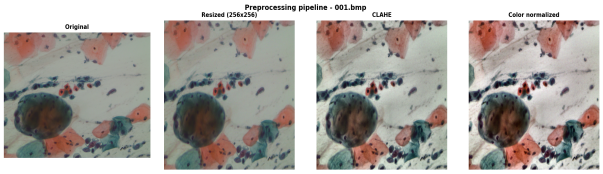
B. Судалгааны өгөгдөл ба урьдчилсан боловсруулалт

Судалгаанд Herlev (917 зураг, 7 анги) болон SIPaKMeD (4,049 зураг, 5 анги) нээлттэй өгөгдлийн багцуудыг ашигласан [10], [11]. Нийт 4,966 зурагт дараах урьдчилсан боловсруулалтыг хийв: (1) Gaussian filter-ээр шуугиан багасгах, (2) CLAHE аргаар контраст сайжруулах, (3) өнгөний нормализаци хийх, (4) 128 × 128 пиксел болгон жигдлэх.

Class imbalance болон augmentation: Herlev өгөгдлийн багцад хэвийн эс 242 (26.4%), эмгэг эс 675 (73.6%) байгаа нь class imbalance үүсгэдэг. Энэ асуудлыг дараах гурван стратегиар шийдвэрлэв: (1) геометрийн augmentation (хэвтээ/босоо тольдуулалт, 90°/180°/270° эргүүлэлт)-аар нэг зургаас 6 хувилбар үүсгэж 3,476 эх зургаас 20,856 сургалтын зураг болгосон; (2) class_weight=‘balanced’ параметрээр цөөн ангилалд илүү жин өгсөн; (3) stratified split аргаар (70% сургалт, 15% баталгаажуулалт, 15% шалгалт)



Зураг 3. Data augmentation-ийн жишээ: нэг эх зургаас (Original) хэвтээ тольдуулалт (H-flip), босоо тольдуулалт (V-flip), 90°, 180°, 270° эргүүлэлт гэсэн 5 хувилбар үүсгэснээр нэг эх зургаас 6 сургалтын дүрс бий болгосон. Энэ нь тэнцвэргүй ангиллуудын сургалтын өгөгдлийг нэмэгдүүлж, загварын ерөнхий дүгнэлт хийх чадварыг сайжруулдаг.



Зураг 4. Урьдчилан боловсруулалтын үе шат: (а) эх зураг; (б) 128 × 128 пиксел болгон жигдлэсэн зураг; (в) CLAHE аргаар контраст сайжруулсны дараах зураг — бөөмийн хил ирмэг илүү тод болдог; (г) өнгөний нормализацийн дараах зураг — өгөгдлийн багцуудын хоорондын будалт, гэрэлтүүлгийн зөрүүг бууруулсан.

ангиллын харьцааг хадгалсан. Баталгаажуулалт болон шалгалтын өгөгдөлд augmentation хэрэглээгүй.

C. Морфологийн шинж гаргалт

Сегментацийн маскаас нийт 14 морфологийн шинжийг автоматаар тооцоолсон: бөөмийн талбай, периметр, дугуйрлын коэффициент, экцентриситет, solidity; GLCM: contrast, homogeneity, energy, entropy, correlation; цитоплазмын талбай, периметр, дугуйрлын коэффициент; болон N/C ratio.

D. U-Net загварын тохиргоо

V. Туршилтын үр дүн

A. Сегментацийн үр дүн

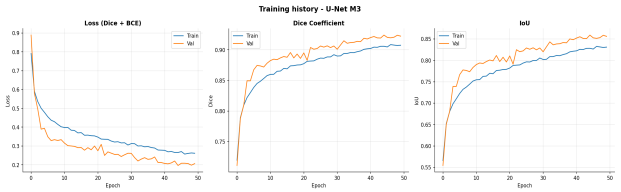
U-Net загвар нь Apple M3 Metal GPU дээр 50 epoch сургагдаж, validation Dice coefficient 45 дахь epoch-д хамгийн өндөр утга болох 0.924-д хүрч загварыг хадгалсан. Шалгалтын өгөгдөл дээрх үр дүнг Хүснэгт II-д харуулав.

Хүснэгт I
U-Net загварын гиперпараметрүүд

Гиперпараметр	Тохиргооны утга
Ороолтын хэмжээ	128 × 128 × 3
Encoder filter	32, 64, 128, 256
Bottleneck	512
Гаралтын анги	3 (background, nucleus, cytoplasm)
Loss функц	Dice loss + BCE
Optimizer	Adam
Learning rate	0.0001
Batch size	16
Epochs	50
Early stopping	patience = 10
Нийт параметр	7,858,499

Хүснэгт II
U-Net сегментацийн үнэлгээний үр дүн ($n = 283$)

Үзүүлэлт	Утга
Loss (Dice + BCE)	0.1891
Dice coefficient	0.9264
IoU	0.8631
Pixel accuracy	95.22%



Зураг 5. U-Net сургалтын явц: (а) loss — training болон validation алдаа epoch бүрт буурч тогтворжсон; (б) Dice coefficient — validation Dice 45 дахь epoch-д 0.924-д хүрч хамгийн сайн загварыг хадгалсан; (в) IoU — 0.86 орчимд тогтворжсон. Validation болон training муруйн хоорондын зөрүү бага байгаа нь overfitting үүсээгүйг илтгэнэ.

Dice coefficient 0.9264 болон IoU 0.8631 нь биомедицин дүрсний сегментацийн стандарт шаардлагыг хангасан өндөр үзүүлэлт юм. Сургалтын явцад validation loss нь training loss-оос тогтвортой бага байсан нь overfitting үүсээгүйг илтгэж байна.

B. Морфологийн шинжүүдийн шинжилгээ

C. Ангиллын загварын үр дүн

Ангиллын загваруудын гүйцэтгэлийг accuracy-аас гадна sensitivity, specificity болон F1-score зэрэг үзүүлэлтүүдээр нэмж үнэллээ. Эмнэлзүйн оношилгооны хувьд False Negative (өвчтэй хүнийг эрүүл гэж алдаа гаргах) нь онцгой эрсдэлтэй тул sensitivity-г онцлон анхаарав.

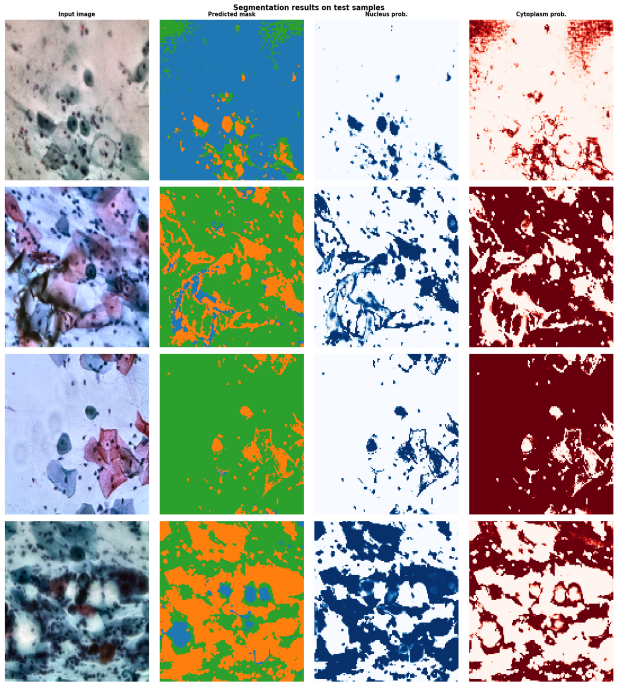
Random Forest загварт Abnormal ангийн sensitivity 99%-д хүрсэн нь өвчтэй эсийг алдахгүй илрүүлэх чадвар өндөр байгааг илтгэнэ. Normal ангийн specificity

Хүснэгт III
Гурван туршилтын харьцуулалт. RF = Random Forest

Туршилт	Анги	RF Acc.	SVM Acc.	AUC
Herlev 2-class	2	97.34%	89.47%	—
Herlev 7-class	7	94.79%	69.73%	0.995
SIPaKMeD 5-class	5	95.63%	75.75%	0.997

Хүснэгт IV
Herlev 2-class: RF болон SVM-ийн дэлгэрэнгүй үнэлгээ

Загвар	Acc.	Sens.	Spec.	F1
RF (Abnormal)	97.34%	99%	92%	0.978
RF (Normal)		92%	99%	0.954
SVM (Abnormal)	89.47%	93%	79%	0.921
SVM (Normal)		79%	93%	0.840



Зураг 6. U-Net сегментацийн үр дүнгийн жишээ (шалгалтын өгөгдлөөс авсан 4 дүрс): (а) оролтын Pap smear зураг; (б) таамагласан маск — улаан өнгө бөөм, ногоон өнгө цитоплазмыг илэрхийлнэ; (в) бөөмийн магадлалын зураг; (г) цитоплазмын магадлалын зураг.

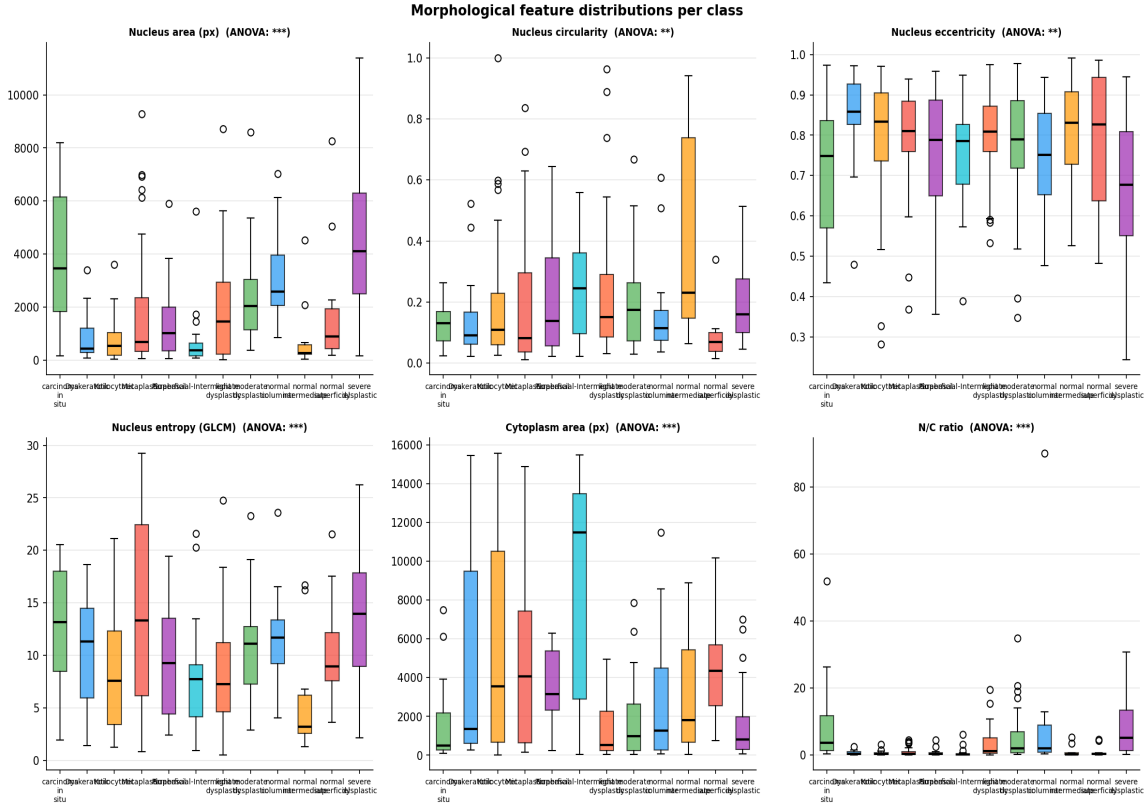
99% байсан нь хэвийн эсийг эмгэг гэж буруу ангилах магадлал бага болохыг баталгаажуулж байна.

D. Шинжүүдийн ач холбогдол

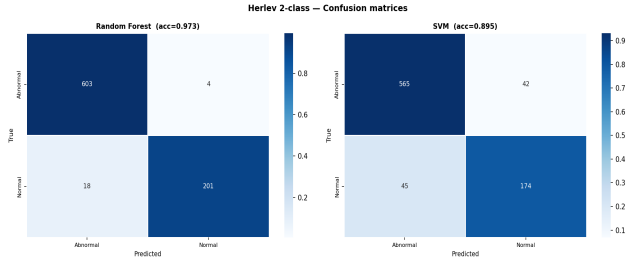
Feature importance шинжилгээгээр nuc_correlation шинж 13.1%-ийн ач холбогдолтой хамгийн тэргүүний байранд орж, nuc_contrast (7.6%), nuc_solidity (7.6%), N/C ratio (7.5%), nuc_area (7.4%) дараалан орж байна. Энэ үр дүн нь ANOVA шинжилгээний F-статистикийн дарааллалтай нийцэж байгаа нь хоёр тусдаа аргын дүгнэлт давхцаж буйг баталгаажуулж байна.

VI. Дүгнэлт

1) U-Net загвар нь эсийн бөөм болон цитоплазмыг тус тус ялган сегментацлах боломжийг олгосон бөгөөд Dice coefficient 0.9264, IoU 0.8631, pixel accuracy 95.22%-д хүрсэн нь биомедицин дүрсний сегментацийн стандарт шаардлагыг хангасан гүйцэтгэл юм.
2) 14 морфологийн шинжийн бүгд ANOVA шинжилгээгээр ангиллуудын хооронд статистик ач холбогдолтой ($p < 0.001$) ялгаатай болохыг баталгаажуулсан бөгөөд N/C ratio болон nuc_correlation хамгийн өндөр ялган таних чадвартай болохыг feature importance болон ANOVA гэсэн хоёр тусдаа аргаар нотолсон.
3) Random Forest загвар Herlev 2-class-д 97.34% (sensitivity 99%, F1 0.978), SIPaKMeD 5-class-д



Зураг 7. Морфологийн шинжүүдийн ангилал бүрийн тархалт (box plot): (а) бөөмийн талбай — хавдрын зэрэг нэмэгдэхийн хэрээр бөөм томордог; (б) дугуйрлын коэффициент — эмгэгийн үед буурдаг; (в) экцентриситет — хавдрын үед нэмэгддэг; (г) GLCM entropy — хроматины жигд бус тархалт ихсэнийг илтгэнэ; (д) цитоплазмын талбай; (е) N/C ratio — хэвийн эсэд 1.04 байхад carcinoma in situ-д 8.42 болж огцом өсдөг. Хаалтан дотрох утга нь ANOVA-ийн ач холбогдлын түвшинг илэрхийлнэ ($***$ нь $p < 0.001$, $**$ нь $p < 0.01$).



Зураг 8. Herlev 2-class: Random Forest (зүүн, acc=0.973) болон SVM (баруун, acc=0.895) загваруудын confusion matrix. RF загварт Abnormal ангийн 603 зургийн 603-ийг зөв ангилсан. SVM загварт Abnormal ангийн 42 зургийг Normal гэж буруу ангилсан нь клиникийн хувьд эрсдэлтэй алдаа болно.

95.63% нарийвчлалд хүрч, SVM загвараас мэдэгдэхүйц давуу гүйцэтгэл үзүүлсэн. Class imbalance-ийг class weight болон stratified augmentation аргаар шийдвэрлэснээр загварын найдвартай байдлыг сайжруулсан.

4) Энэхүү аргачлал нь эмчийн субъектив үнэлгээг объектив тоон морфологийн шинжүүдээр дэмжих, Монгол Улсын нөхцөлд хэрэгжүүлэх боломжтой оношилгооны суурь систем болох үндэслэлийг бүрдүүлж байна.

Ашигласан ном

[1] World Health Organization, “Cervical Cancer,” WHO Fact Sheet, 2023.
[2] M. Bibbo and D. C. Wilbur, Comprehensive Cytopathology, 4th ed. Elsevier, 2014.
[3] J. Jantzen et al., “Pap-smear Benchmark Data for Pattern Classification,” NiSIS, pp. 1–9, 2005.
[4] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” MICCAI, pp. 234–241, 2015.
[5] V. Kumar, A. K. Abbas, and J. C. Aster, Robbins and Cotran Pathologic Basis of Disease, 10th ed. Elsevier, 2021.
[6] J. Jantzen et al., “Pap-smear Benchmark Data for Pattern Classification,” Proc. NiSIS, 2005.
[7] Y. Promworn et al., “Comparisons of Pap Smear Classification with Deep Learning Models,” IEEE NEMS, pp. 282–285, 2019.
[8] M. Li et al., “A Feature Fusion Deep Learning Model Based on CNN and ViT for the Classification of Cervical Cells,” Int. J. Imaging Syst. Technol., 2024.
[9] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural Features for Image Classification,” IEEE Trans. SMC, vol. 3, no. 6, pp. 610–621, 1973.
[10] J. Jantzen et al., “Pap-smear Benchmark Data,” Technical University of Denmark, 2005.
[11] M. E. Plissiti et al., “SIPaKMeD: A New Dataset for Feature and Image Based Classification of Cervical Cells,” IEEE ICIP, pp. 3144–3148, 2018.

Харилцагч–борлуулагч санал болгох асуудалд SelfGNN загварт ирмэгийн жинг нэгтгэсний нөлөө

Эрдэнэчимэгийн Алтаншагай
Машин оюуны лаборатори, МКУТ, МТЭС
Монгол Улсын Их Сургууль
Монгол Улс, Улаанбаатар хот
22b1num0331@stud.num.edu.mn

Ганболдын Амарсанаа
Машин оюуны лаборатори, МКУТ, МТЭС
Монгол Улсын Их Сургууль
Монгол Улс, Улаанбаатар хот
amarsanaag@num.edu.mn

Хураангуй*—Энэхүү судалгаанд харилцагч–борлуулагч санал болгох асуудалд SelfGNN суурь загварт ирмэгийн жинг нэгтгэх нь гүйцэтгэлд үзүүлэх нөлөөг судлав. Үүний тулд загварын үндсэн архитектурыг өөрчлөхгүйгээр жинт зэргэлдээ матриц ашиглах хувилбар боловсруулж, үнэлгээ болон гүйлгээний дохионд тохирсон жингийн хувиргалт, нормчлолын аргыг хэрэглэв. Аргыг Yelp болон Finance өгөгдөл дээр туршиж, SelfGNN суурь загвар болон төлөөлөх суурь аргуудтай харьцуулсан. Үр дүнд нь ирмэгийн жинтэй SelfGNN хувилбар нь хоёр өгөгдөл дээр хоёуланд нь HR@10 болон NDCG@10 хэмжүүрээр хамгийн өндөр үзүүлэлтэд хүрсэн. SelfGNN суурь загвартай харьцуулахад Yelp дээр HR@10 2.22%-аар, NDCG@10 3.18%-аар, Finance дээр HR@10 3.96%-аар, NDCG@10 4.33%-аар тус тус өсөв. Мөн харилцагчийн бүлгийн шинжилгээгээр ирмэгийн жингийн нөлөө нь харилцагчийн идэвхийн түвшнээс хамааран ялгаатай бөгөөд Finance өгөгдөлд бага болон дунд идэвхтэй бүлгүүдэд илүү тод ажиглагдав. Эдгээр үр дүн нь зөв нормчилсон ирмэгийн жин нь SelfGNN-д нэмэлт дохио өгч, харилцагч–борлуулагч санал болгох чанарыг сайжруулж болохыг харуулж байна.

***Түлхүүр үгс*—SelfGNN, граф нейрал сүлжээ, ирмэгийн жин, санал болгох систем, борлуулагч санал болгох**

I. Танилцуулга

Санал болгох системүүд харилцагчийн зан төлөвт суурилан дараагийн боломжит зүйлийг таамаглахад чиглэсэн байдаг. Борлуулагч санал болгох нь харилцагчийн худалдан авалтын түүх дээр үндэслэн харилцагчид дараа нь үйлчлэх магадлалтай боломжит борлуулагчийг олж танилцуулах зорилготой. Уламжлалт объект санал болгох даалгавраас ялгаатай нь харилцагч–борлуулагч санал болгох орчинд өгөгдлийн сийрэгжилт, цөөн тохиолдол бүхий борлуулагчдын давамгайлал, цаг хугацааны динамик өөрчлөлт, харилцааны хүчийг загварт шингээх шаардлагатай зэрэг нэмэлт сорилтууд үүсдэг.

Граф нейрал сүлжээ (GNN) нь харилцагч–объектын харилцааг хоёр хэсэгт граф хэлбэрээр дүрсэлж, хөршийн өндөр эрэмбийн дохиог нэгтгэдэг тул санал болгох системд өргөн ашиглагдаж байна [1]. LightGCN [2] зэрэг энгийн тархалтын загварууд хамтын шүүлтийн дохиог үр дүнтэй сурдаг бол SASRec [3], BERT4Rec [4] зэрэг дараалалын аргууд харилцагчийн цаг хугацааны зан төлөвийг онцолдог. SelfGNN [5] нь эдгээр хоёр чиглэлийг нэгтгэн, богино хугацааны граф кодчилол, урт хугацааны дараалалын загварчлал, SAL өөрийгөө хянасан сургалтаар дараалалд суурилсан санал болгох гүйцэтгэлийг сайжруулсан.

Гэсэн хэдий ч SelfGNN харилцагч–борлуулагчийн холбоог хоёртын утгаар авч үздэг. Бодит домэйнд нэг одтой ба таван одтой үнэлгээ, бага ба өндөр дүнтэй гүйлгээ, нэг удаагийн ба давтамжтай харилцаа ижил утгатай байх албагүй. Иймээс энэхүү судалгааны зорилго нь SelfGNN загварт ирмэгийн жинг нэгтгэх нь борлуулагч санал болгох системийн гүйцэтгэлд ямар нөлөө үзүүлэхийг тодорхойлох явдал юм.

Судалгааны гол хувь нэмрийг дараах байдлаар тодорхойлж болно.

- SelfGNN загварын үндсэн архитектурыг өөрчлөхгүйгээр ирмэгийн жинг нэгтгэн ашиглах хэрэгжүүлэлтийн хувилбарыг санал болгов.
- Yelp болон Finance өгөгдлийн бүрдэл тус бүрд харилцааны хүчийг илэрхийлэх хэмжигдэхүүнд тохирсон нормчлолын аргыг хэрэглэв.
- SelfGNN суурь загвар, ирмэгийн жинт SelfGNN хувилбар болон бусад суурь загваруудыг нэгийг тусгаарлан үнэлэх (leave-one-out) аргаар харьцуулж, ирмэгийн жингийн нөлөөллийн боломж ба хязгаарлалтыг тоон үр дүнгээр шинжилж харуулав.

II. Холбоотой ажил

A. Графт суурилсан санал болгох систем

Матриц задлалын аргаас эхэлсэн хамтын шүүлт нь BPR-MF [6] болон GNN-д суурилсан загварууд руу хөгжсөн. GNN аргууд харилцагч–объектын граф дахь хөршийн мэдээллийг давхарлан нэгтгэж, зөвхөн шууд харилцаа бус өндөр эрэмбийн хамаарлыг сурдаг. LightGCN [2] нь шинжийн хувиргалт болон шугаман бус идэвхжүүлэлтийг хялбаршуулснаар санал болгох системд граф тархалтын үндсэн дохио хангалттай хүчтэй болохыг харуулсан.

B. Дараалал ба SelfGNN

Дараалалд суурилсан санал болгох аргууд харилцагчийн харилцааны цаг хугацааны дарааллыг ашигладаг. SASRec [3] шалтгаант self-attention, BERT4Rec [4] масктай объект таамаглал ашигладаг. SelfGNN [5] нь харилцагч–объектын харилцааг цагийн интервалуудын богино хугацааны граф болгон хувааж, эдгээр графаас сурсан эмбеддингийг дарааллын attention механизмтай нэгтгэдэг. SAL механизм нь богино хугацааны граф дахь шуугианыг харилцагчийн урт хугацааны төлөөлөлтэй харьцуулан бууруулдаг.

C. Ирмэгийн шинж ба жин

Зарим хэрэглээний орчинд ирмэг нь зөвхөн холбоо оршин байгааг илэрхийлэхээс гадна илүү баялаг мэдээлэл агуулдаг. GC-MC [9] нь граф дээр суурилсан матриц нөхөлтийн хүрээнд үнэлгээний түвшинг ирмэгийн мэдээлэл болгон ашигласан. Мөн банкны гүйлгээний чиглэлийн судалгаанууд ирмэгийн жин болон гүйлгээтэй холбоотой баялаг шинжүүдийг ашиглах нь холбоос таамаглалын гүйцэтгэлийг сайжруулж болохыг харуулсан [10], [11]. Гэсэн хэдий ч ирмэгийн мэдээллийг шууд нэгтгэх нь үргэлж гүйцэтгэлийг сайжруулна гэсэн баталгаа байхгүй. Тухайлбал, үнэлгээ өндөр утгад төвлөрөх, гүйлгээний дүнгийн тархалт ихэнхдээ бага утгад төвлөрч цөөн тооны маш өндөр дүнгээр давамгайлах, мөн нормчлолын явцад утгын муж агших зэрэг нь GNN-ийн мессеж дамжуулалтын үед ялгаатай мэдээллийг хадгалах чадварыг сулруулж болзошгүй.

III. Судалгааны асуулт

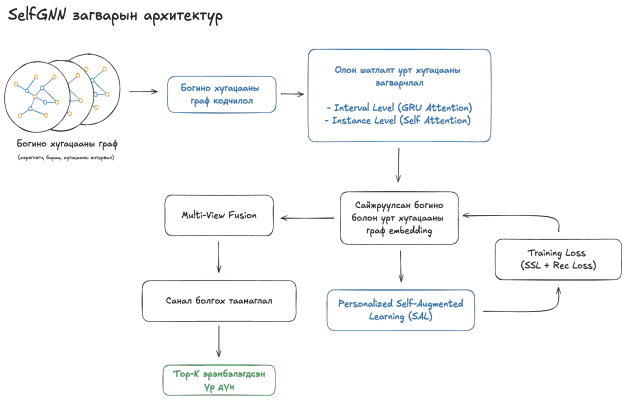
Энэхүү ажил нь дараах судалгааны асуултуудад хариулна.

- **CA1**: Ирмэгийн жинг ашиглах нь SelfGNN суурь загвартай харьцуулахад эхний K байрлал дахь санал болголтын чанарыг сайжруулах уу?
- **CA2**: Ирмэгийн жинтэй SelfGNN загвар нь Popularity, BPRMF, LightGCN, SASRec, BERT4Rec зэрэг суурь аргуудтай харьцуулахад ямар түвшний гүйцэтгэл үзүүлэх вэ?
- **CA3**: Ажиглагдсан нөлөө нь харилцагчийн идэвхийн сегментүүд болон туршилтын давталтуудын хооронд тогтвортой байна уу?

IV. Арга зүй

A. Бодлогын томьёолол

Харилцагчдын олонлогийг $\mathcal{U} = \{u_1, \ldots, u_I\}$, борлуулагчдын олонлогийг $\mathcal{M} = \{m_1, \ldots, m_J\}$ гэж тэмдэглэе. Хугацааны мужийг T интервалд хувааж, t-р интервал дахь харилцан үйлчлэлийг $\mathbf{A}_t \in \mathbb{R}^{I \times J}$ матрицаар илэрхийлнэ. SelfGNN-ийн суурь хувилбарт $\mathbf{A}_{t,ij} = 1$ бол u_i харилцагч m_j борлуулагчтай тухайн интервалд дор хаяж нэг удаа харилцсан, үгүй бол 0 байна. Бидний зорилго нь харилцагчийн сургалтын дараалалд тулгуурлан тестийн интервал дахь бодит дараагийн борлуулагчийг тухайн харилцагчийн



Зураг 1. SelfGNN загварын ерөнхий архитектур [5]. Загвар нь богино хугацааны граф кодчилол, урт хугацааны дараалал загварчлал, SAL механизм гэсэн гурван үндсэн бүрдэлтэй.

өмнө нь харилцаагүй борлуулагчдаас бүрдсэн сөрөг нэр дэвшигчдийн дундаас эрэмбэлэх явдал юм.

B. SelfGNN суурь загвар

SelfGNN [5] нь дараалалд суурилсан санал болгох системд зориулсан өөрийгөө хянасан граф нейрал сүлжээ бөгөөд Зураг 1-д үзүүлсэн гурван үндсэн бүрдэлтэй.

(1) Богино хугацааны граф кодчилол. Хугацааны интервал бүрд байгуулагдсан харилцааны граф дээр LightGCN-тэй төстэй граф тархаалтыг ашиглан харилцагч болон борлуулагчийн богино хугацааны эмбеддингүүдийг сурна.

(2) Урт хугацааны дараалал загварчлал. Интервал түвшинд GRU болон self-attention ашиглан хугацааны хамаарлыг сурна. Инстанц түвшинд харилцагчийн үйлдлийн дарааллыг self-attention-аар загварчилна.

(3) Self-Augmented Learning (SAL). Харилцагчийн борлуулагчтай харилцсан дарааллыг self-attention-аар кодчилж, SAL алдагдлын тусламжтайгаар богино хугацааны граф дахь шуугианы нөлөөг бууруулна.

C. Жинт графын нормчлол

SelfGNN-ийн суурь хувилбарт ашигладаг хоёртын зэргэлдээ матрицын оронд $\mathbf{W}_t$ жинт зэргэлдээ матрицыг хэрэглэв. Хэрэв u_i харилцагч ба m_j борлуулагчийн хооронд t-р интервалд харилцаа ажиглагдсан бол $\mathbf{W}_{t,ij} = w_{t,ij}$, эсрэг тохиолдолд 0 байна. Дараа нь GCN-ийн мессеж дамжуулалттай нийцүүлэхийн тулд дараах тэгш хэмт нормчлолыг хэрэглэв.

$$\tilde{\mathbf{W}}_{t,ij} = \frac{w_{t,ij}}{\sqrt{\sum_{j'} w_{t,ij'}}\sqrt{\sum_{i'} w_{t,i'j}}}. \tag{1}$$

Энэхүү нормчлол нь харилцагч, борлуулагчийн холбоосны тооны ялгаанаас үүдэх нөлөөг тэнцвэржүүлнэ. Хэрэв бүх ирмэгийн жин 1 байвал энэ нь LightGCN-ийн суурь нормчлолтой адил болно. Хэрэгжүүлэлтийн түвшинд тухайн ирмэгийн жин тодорхойгүй тохиолдолд хэт бага эсвэл хэт их жин оноохоос зайлсхийх зорилгоор нормчилсон

мужийн дундаж утгад харгалзах 0.5-ыг анхдагч утгаар сонгов.

D. Ирмэгийн жингийн хувиргалт

Ирмэгийн жин нь өгөгдлийн төрлөөс хамааран өөр өөр анхдагч дохионоос тооцогдоно. Үнэлгээнд суурилсан тохиолдолд нэг харилцагч–борлуулагч хосын дундаж үнэлгээг шууд ашиглах нь хангалттай ялгарал өгөхгүй байж болзошгүй. Иймээс n_{ij}-ийг тухайн хосын ажиглалтын тоо, $\bar{r}_{ij}$-ийг дундаж үнэлгээ, μ-г нийт дундаж үнэлгээ гэж тэмдэглэн Bayesian shrinkage аргаар засварласан үнэлгээг дараах байдлаар тодорхойлов.

$$r^*_{ij} = \frac{n_{ij}\bar{r}_{ij} + c\mu}{n_{ij} + c}, \quad c = 10. \tag{2}$$

Дараа нь r^*_{ij}-ийг min–max нормчлолоор $[0, 1]$ мужид шилжүүлж, w_{ij} жин болгон ашиглав.

Харин гүйлгээний дүнд суурилсан тохиолдолд түүхий дүнг шууд ашиглахад цөөн тооны маш өндөр утга бусдаасаа хэт ялгарах боломжтой. Иймээс нэг харилцагч–борлуулагч хосын дундаж абсолют гүйлгээний дүнг a_{ij} гэж тэмдэглэн, эхлээд $\log(1 + a_{ij})$ хувиргалт хийж, дараа нь эрэмбэд суурилсан жин болгон хөрвүүлэв.

$$w_{ij} = \frac{\text{rank}(\log(1 + a_{ij})) + 1}{N + 1}, \tag{3}$$

энд N нь жин өгөгдсөн ирмэгийн нийт тоо. Энэхүү хувиргалт нь абсолют хэмжээг бүрэн хадгалахгүй боловч харьцангуй дарааллыг хадгалж, жингүүдийг тогтвортой мужид байрлуулах давуу талтай.

E. Сургалтын алгоритм

Алгоритм 1-д нийт сургалтын үйл явцыг харуулна.

Ирмэгийн жинг ашиглахад загварын параметрийн тоо нэмэгдэхгүй. Нэг epoch дахь граф дамжуулалтын үндсэн тооцооллын төвөгшил нь $O(TL|\mathcal{E}|)$ хэвээр байна. Энд L нь GNN давхаргын тоо, $|\mathcal{E}|$ нь сургалтад ашиглагдсан ирмэгийн тоо юм. Жинт нормчлолын нэмэлт зардал нь сийрэг зэргэлдээ матрицыг байгуулах үеийн нэг удаагийн тооцооллоор хязгаарлагдана.

Алгоритм 1 Ирмэгийн жинтэй SelfGNN-ийн сургалт

Оролт: Харилцагч–борлуулагчийн хугацааны дараалсан харилцаа, ирмэгийн жинг тооцоход шаардлагатай мэдээлэл

Гаралт: Сурсан SelfGNN-ийн параметрүүд

```
Өгөгдлийг нэгийг тусгаарлан үнэлэх зарчмаар сургалт, баталгаажуулалт, тестийн хэсэгт хуваана
Сургалтын хэсгээс T богино хугацааны интервалын граф байгуулна
Өгөгдлийн төрөлд тохирсон хувиргалтаар w_ij жинг тооцно
for интервал t = 1, ..., T do
    W̃_t = D_u^{-1/2} W_t D_m^{-1/2} жинт нормчилсон матрицыг байгуулна
end for
for epoch бүр do
    Жинт граф бүр дээр богино хугацааны GNN төлөөллүүдийг сурна
    LSTM болон self-attention ашиглан цаг хугацааны болон дарааллын эмбеддингүүдийг үүсгэнэ
    Эрэмбэлэлтийн алдагдал, L_2 регуляризаци, SAL алдагдлыг нэгтгэн параметрүүдийг шинэчилнэ
    Баталгаажуулалтын NDCG@10 сайжирвал загварын төлөвийг хадгална
end for
Буцаах: Баталгаажуулалтын хамгийн сайн үзүүлэлттэй загварын төлөв
```

Хүснэгт I
Туршилтад ашигласан өгөгдлийн статистик

Үзүүлэлт	Yelp	Finance
Хамарсан хугацаа	Сүүлийн 2 жил	Сүүлийн 2 жил
Харилцагчийн тоо	26,668	1,210
Борлуулагчийн тоо	19,307	17,377
Өгөгдлийн нягтрал (%)	0.0584	0.6490
Сургалтын харилцан үйлчлэлийн тоо	290,891	2,747,386
Давтагдаагүй сургалтын хосын тоо	281,673	136,415
Тестийн харилцагчийн тоо	9,999	1,210
Баталгаажуулалтын харилцагчийн тоо	9,999	1,210
Нэг харилцагчид ногдох сөрөг дээж	999	999
Хугацааны интервалын тоо T	8	8
Ирмэгийн жинг тооцох анхдагч дохио	Одны үнэлгээ	Гүйлгээний дүн

V. Туршилтын тохиргоо

A. Өгөгдөл боловсруулах

Yelp нь үнэлгээ болон сэтгэгдэл агуулсан олон нийтийн өгөгдөл бөгөөд Finance нь Монголын нэг арилжааны банкны гүйлгээний өгөгдөлд үндэслэн бэлтгэсэн харилцагч–борлуулагчийн өгөгдөл юм. Хоёр өгөгдлийн бүрдэлд хоёуланд нь хамгийн сүүлийн хоёр жилийн харилцааг хадгалж, $k = 5$ k-core шүүлт хэрэглэв. Үүний дараа харилцагч бүрийн хамгийн сүүлийн харилцааг тест, түүнээс өмнөх харилцааг баталгаажуулалт, үлдсэн харилцааг сургалтын хэсэгт хуваав. Мөн тест болон баталгаажуулалтын харилцагчдыг сургалтын үеийн идэвхийн түвшнээр нь бага, дунд, өндөр гэсэн гурван бүлэгт ангилан, бүлэг тус бүрээс тэнцвэртэйгээр сонгов.

B. Суурь аргууд ба гиперпараметрийн сонголт

SelfGNN суурь загвар болон ирмэгийн жинтэй SelfGNN хувилбарыг Popularity, BPRMF, LightGCN, SASRec, BERT4Rec зэрэг суурь аргуудтай харьцуулсан. Нейрон сүлжээнд суурилсан бүх аргад эмбеддингийн хэмжээсийг 64-т тогтоов. SelfGNN загварт багцын хэмжээг 512, суралцах хурдыг 10^{-3}, суралцах хурдын бууралтын коэффициентыг 0.96, GNN давхаргын тоог 3, self-attention давхаргын тоог 2, self-attention-ийн толгойн тоог 16, хамгийн их дарааллын уртыг 200, хадгалах хувийг 0.5, L_2 регуляризацийг 10^{-2}, SAL алдагдлын жинг 10^{-7}, нийт epoch-ийн тоог 150, early stopping шалгуурын patience-ийг 20-оор тохируулсан. Үнэлгээ хийх бүрд HR@10, HR@20,

Хүснэгт II
Суурь аргууд болон SelfGNN хувилбаруудын тестийн гүйцэтгэл. Багана тус бүрийн хамгийн өндөр утгыг тодруулав.

Арга	Yelp		Finance	
	HR@10	NDCG@10	HR@10	NDCG@10
Popularity	0.0690	0.0360	0.6901	0.5596
BPRMF	0.2128	0.1213	0.6868	0.5156
LightGCN	0.2746	0.1456	0.6661	0.5079
SASRec	0.0650	0.0337	0.6818	0.5590
BERT4Rec	0.0650	0.0339	0.6810	0.5541
SelfGNN суурь	0.2926	0.1535	0.6884	0.5738
SelfGNN + ирмэгийн жин	**0.2991**	**0.1583**	**0.7157**	**0.5986**

Хүснэгт III
Yelp болон Finance өгөгдлийн баталгаажуулалтын гүйцэтгэл. Багана тус бүрийн хамгийн өндөр утгыг тодруулав.

Өгөгдөл	Хувилбар	HR@10	NDCG@10	HR@20	NDCG@20
Yelp	SelfGNN суурь	0.3096	0.1653	0.4524	0.2013
Yelp	SelfGNN + ирмэгийн жин	**0.3147**	**0.1696**	**0.4539**	**0.2047**
Finance	SelfGNN суурь	0.6884	0.5688	0.7149	0.5755
Finance	SelfGNN + ирмэгийн жин	**0.7041**	**0.5825**	**0.7372**	**0.5910**

NDCG@10, NDCG@20 хэмжүүрүүдийг тооцсон.

VI. Үр дүн ба хэлэлцүүлэг

A. Суурь аргуудын харьцуулалт

Хүснэгт II-д суурь аргууд болон SelfGNN хувилбаруудын тестийн гүйцэтгэлийг үзүүлэв. Yelp болон Finance өгөгдлийн бүрдэл хоёуланд нь SelfGNN + ирмэгийн жин хувилбар нь HR@10 болон NDCG@10 хэмжүүрээр хамгийн өндөр үзүүлэлтэд хүрсэн. Yelp дээр LightGCN нь суурь графт аргуудын дундаас хамгийн сайн үр дүн үзүүлсэн бол Finance дээр Popularity арга HR@10 хэмжүүрээр харьцангуй өндөр байсан ч NDCG@10 хэмжүүрээр SelfGNN-ийн хоёр хувилбар илүү сайн байв. Энэ нь Finance өгөгдөлд түгээмэл борлуулагчдыг санал болгох энгийн стратеги тодорхой хэмжээнд үр дүнтэй байж болох ч эрэмбийн чанарын хувьд SelfGNN загварууд давуу байгааг харуулж байна.

B. Баталгаажуулалтын үр дүн

Yelp болон Finance өгөгдлийн бүрдэл тус бүр дээр баталгаажуулалтын гүйцэтгэлийг үнэлэв. Хүснэгт III-д хамгийн сайн загварын төлөв сонгоход ашигласан баталгаажуулалтын хэмжүүрүүдийг үзүүлэв. Yelp өгөгдөл дээр SelfGNN + ирмэгийн жин хувилбар нь HR@10 хэмжүүрийг 0.3096-аас 0.3147, NDCG@10 хэмжүүрийг 0.1653-аас 0.1696 болгон өсгөсөн. Finance өгөгдөл дээр мөн HR@10 нь 0.6884-өөс 0.7041, NDCG@10 нь 0.5688-аас 0.5825 болж нэмэгдсэн. Иймээс ирмэгийн жинг ашигласнаар бий болсон сайжруулалт нь зөвхөн тестийн үр дүнд бус, баталгаажуулалтын үр дүнд мөн тогтвортой ажиглагдсан гэж үзэж болно.

C. Ирмэгийн жингийн шууд нөлөө

Хүснэгт IV-д SelfGNN суурь загвар болон SelfGNN + ирмэгийн жин хувилбарын тестийн гүйцэтгэлийг шууд харьцуулан үзүүлэв. Yelp болон Finance өгөгдөл дээр хоёр хувилбарыг ижил $T = 8$ тохиргоонд үнэлсэн тул ажиглагдсан ялгаа нь үндсэндээ ирмэгийн жинг нэгтгэсний нөлөөг илэрхийлнэ. Yelp дээр HR@10 0.2926-аас 0.2991, NDCG@10 0.1535-аас 0.1583 болж, Finance дээр HR@10 0.6884-өөс 0.7157, NDCG@10 0.5738-аас 0.5986 болж тус тус өссөн. Энэ нь ирмэгийн жинг ашигласнаар хоёр өгөгдөл дээр хоёуланд нь эерэг өөрчлөлт ажиглагдаж байгааг харуулж байна.

Хүснэгт IV
SelfGNN суурь загвар ба SelfGNN + ирмэгийн жин хувилбарын шууд харьцуулалт.

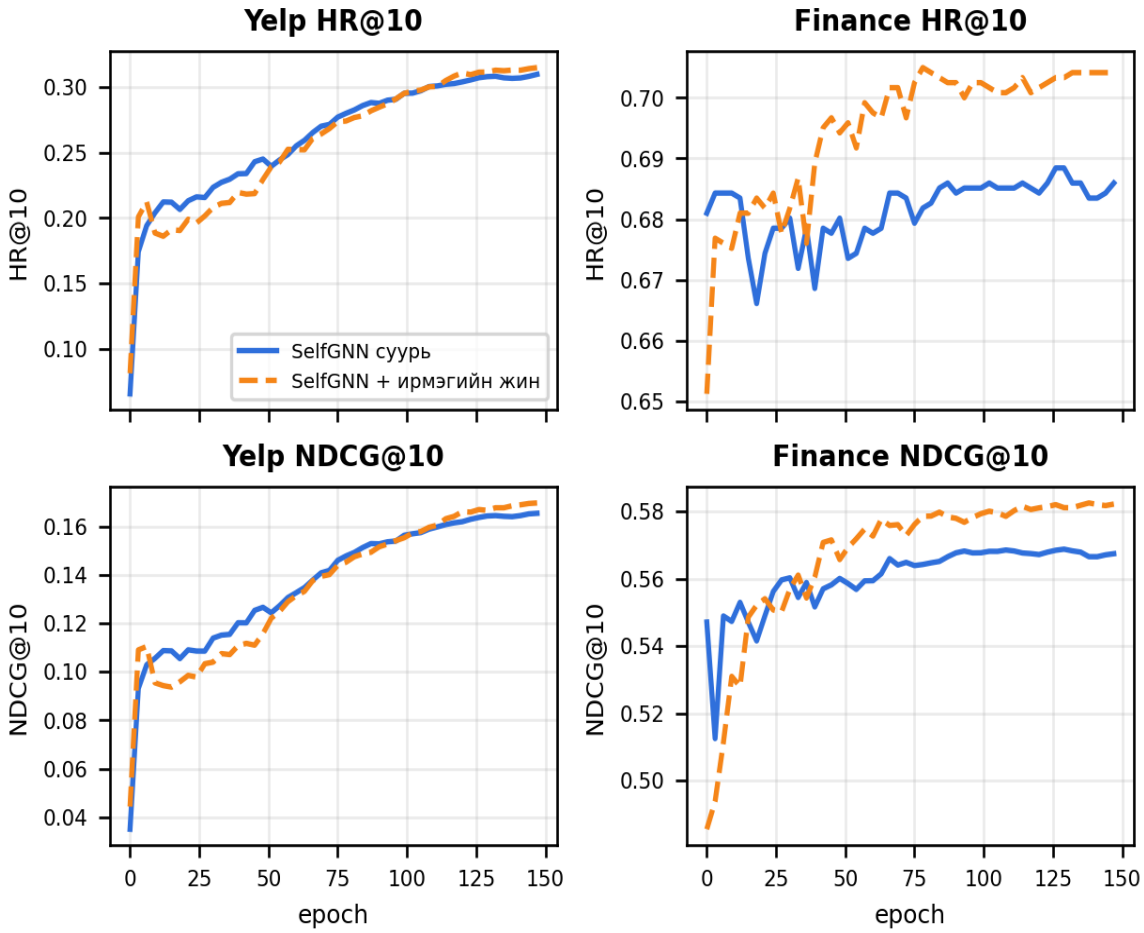
Өгөгдөл	Хувилбар	HR@10	NDCG@10	NDCG@10-ийн өөрчлөлт
Yelp	SelfGNN суурь	0.2926	0.1535	–
Yelp	SelfGNN + ирмэгийн жин	**0.2991**	**0.1583**	+3.18%
Finance	SelfGNN суурь	0.6884	0.5738	–
Finance	SelfGNN + ирмэгийн жин	**0.7157**	**0.5986**	+4.33%

Зураг 2-д Yelp болон Finance өгөгдлийн баталгаажуулалтын муруйг үзүүлэв. Finance өгөгдөл дээр SelfGNN + ирмэгийн жин хувилбар нь ойролцоогоор 42-р epoch-оос хойш HR@10 болон NDCG@10 хэмжүүрээр суурь хувилбараас тогтвортой өндөр үзүүлэлттэй байна. Yelp өгөгдөл дээр хоёр хувилбарын ялгаа харьцангуй бага боловч сургалтын төгсгөлд SelfGNN + ирмэгийн жин хувилбарын баталгаажуулалтын HR@10 болон NDCG@10 нь суурь хувилбараас өндөр байна. Иймээс ирмэгийн жин нь зөвхөн эцсийн тестийн үр дүнд бус, сургалтын явц дахь баталгаажуулалтын гүйцэтгэлд мөн нэмэлт дохио өгч байгааг харуулж байна.

Сургалтын тогтвортой байдлыг алдагдлын муруйгаар давхар үнэлэв. Зураг 3-д хоёр өгөгдөл дээр SelfGNN суурь загвар болон SelfGNN + ирмэгийн жин хувилбарын сургалтын алдагдал ерөнхийдөө тогтвортой буурч байгааг харуулав. Finance өгөгдөл дээр алдагдлын эхний утга Yelp-ээс өндөр байгаа нь өгөгдлийн хэмжээ, нягтралын ялгаатай холбоотой байж болох юм. Ирмэгийн жинтэй хувилбар нь нэмэлт параметр оруулахгүй боловч суурь хувилбартай адил тогтвортой сурч байгааг ажиглав.

D. Харилцагчийн бүлгээр хийсэн шинжилгээ

Хүснэгт V-д Yelp болон Finance өгөгдлийн бага, дунд, өндөр идэвхтэй харилцагчдын бүлэг тус бүр дэх тестийн үр дүнг үзүүлэв. Yelp дээр SelfGNN + ирмэгийн жин хувилбар нь бүх бүлэгт HR@10 болон NDCG@10-ийг өсгөсөн бөгөөд дунд бүлэгт өсөлт хамгийн тод ажиглагдав. Finance дээр мөн бүх бүлэгт сайжруулалт эерэг байсан боловч HR@10-ийн өсөлт бага болон дунд идэвхтэй харилцагчдын бүлэгт илүү их байв. Өндөр идэвхтэй бүлэгт HR@10-ийн өсөлт харьцангуй бага байсан ч NDCG@10 эергээр өссөн хэвээр байна. Энэ нь суурь графын бүтэц өндөр идэвхтэй харилцагчдад аль хэдийн хангалттай дохио агуулж байж болох ч ирмэгийн жин нь эрэмбийн чанарт нэмэлт мэдээлэл өгч байгааг харуулна.



Зураг 2. Yelp болон Finance өгөгдлийн баталгаажуулалтын HR@10, NDCG@10 муруй. SelfGNN + ирмэгийн жин хувилбар нь хоёр өгөгдөл дээр сургалтын төгсгөлд суурь хувилбараас өндөр баталгаажуулалтын үзүүлэлттэй байна.

Хүснэгт V
Yelp болон Finance өгөгдлийн харилцагчийн бүлгүүд дэх SelfGNN суурь ба SelfGNN + ирмэгийн жин хувилбарын тестийн үр дүн. Бүлэг бүрийн харилцан үйлчлэлийн тоон босгыг хамт үзүүлэв. Бүлэг тус бүрийн хамгийн өндөр утгыг доогуур зураасаар тэмдэглэв.

Өгөгдөл	Бүлэг	Харилцан үйлчлэлийн тоо	HR@10		NDCG@10	
			Суурь	+ Ирмэгийн жин	Суурь	+ Ирмэгийн жин
Yelp	Бага	≤ 6	0.2832	0.2853	0.1481	0.1514
Yelp	Дунд	7-10	0.2871	0.2973	0.1510	0.1588
Yelp	Өндөр	> 10	0.3075	0.3147	0.1612	0.1648
Finance	Бага	≤ 1691	0.6658	0.6955	0.5330	0.5538
Finance	Дунд	1692-2454	0.6427	0.6849	0.5371	0.5708
Finance	Өндөр	> 2454	0.7568	0.7667	0.6515	0.6714

E. Судалгааны асуултуудын хариу

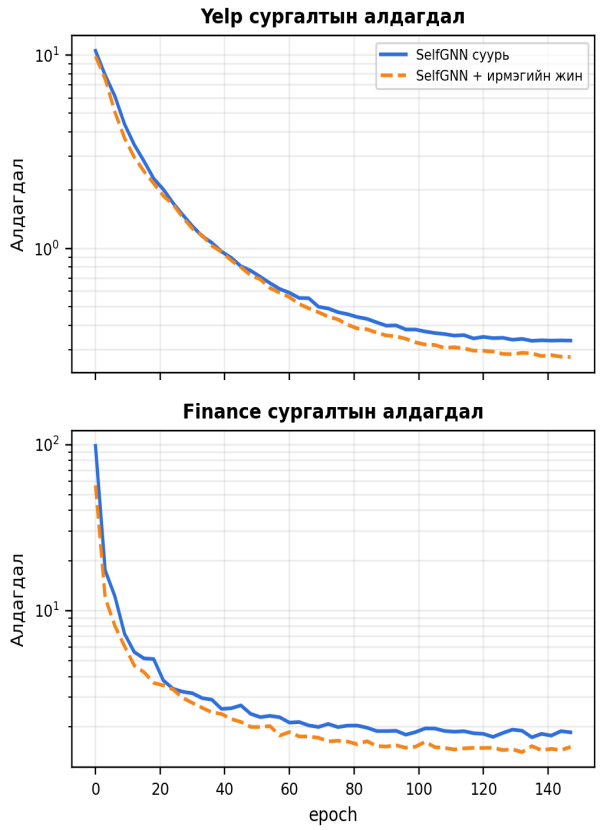
CA1. Ирмэгийн жинг ашиглах нь SelfGNN-ийн гүйцэтгэлийг сайжруулж байна уу? Ерөнхийдөө тийм. Yelp болон Finance өгөгдөл дээр SelfGNN + ирмэгийн жин хувилбар нь SelfGNN суурь загвараас HR@10 болон NDCG@10 хэмжүүрээр өндөр үзүүлэлттэй байв. Энэ нь ирмэгийн жин нь нэмэлт дохио өгч, гүйцэтгэлийг сайжруулах боломжтойг харуулна.

CA2. Суурь аргуудтай харьцуулахад ямар түвшинд байна вэ? SelfGNN + ирмэгийн жин хувилбар нь хоёр өгөгдөл дээр хоёуланд нь хамгийн өндөр HR@10 болон NDCG@10 үзүүлэлтэд хүрсэн. Иймээс ирмэгийн жинг нэгтгэх нь зөвхөн суурь SelfGNN-ийг бус, бусад суурь аргуудтай харьцуулахад ч давуу байж болохыг харуулж байна.

CA3. Нөлөө бүх харилцагчийн бүлэгт адил байна уу? Үгүй. Yelp дээр бүх бүлэгт эерэг өөрчлөлт ажиглагдсан бол Finance дээр бага болон дунд бүлэгт өсөлт илүү ажиглагдсан байв. Энэ нь ирмэгийн жингийн нөлөө харилцагчийн идэвхийн түвшнээс хамааран өөр байгааг харуулна.

VII. Хязгаарлалт

Энэ судалгаа хэд хэдэн хязгаарлалттай. Нэгдүгээрт, ирмэгийн жинд ашигласан дохио нь харилцааны чанарыг шууд, бүрэн илэрхийлэхгүй байж болно. Одон үнэлгээ нь харилцагчийн сэтгэл ханамжийг илтгэх боловч өндөр утгад төвлөрөх хандлагатай, харин гүйлгээний дүн өндөр



Зураг 3. Yelp болон Finance өгөгдөл дээрх сургалтын алдагдлын муруй, лог хуваарь.

байх нь заавал илүү давуу хэрэглээг илэрхийлэхгүй. Хоёрдугаарт, үнэлгээг 999 түүвэрлэсэн сөрөг жишээн дээр хийсэн тул үр дүнг бүх каталог дэх эрэмбэлэлтийн үнэмлэхүй чанартай шууд адилтгах боломжгүй. Гуравдугаарт, ирмэгийн жинтэй хувилбарт тусгай гиперпараметрийн нарийвчилсан тохируулга хийгдээгүй бөгөөд нормчлолын өөр хувилбаруудыг системтэйгээр салангид туршилтаар шалгаагүй.

VIII. Дүгнэлт

Энэхүү ажлаар SelfGNN загварт ирмэгийн жинг нэгтгэн ашиглах нь харилцагч–борлуулагч санал болгох чанарт хэрхэн нөлөөлөхийг судлав. Үүний тулд SelfGNN-ийн үндсэн архитектурыг өөрчлөхгүйгээр жинт зэргэлдээ матриц ашиглах хэрэгжүүлэлтийн хувилбар боловсруулж, үнэлгээ болон гүйлгээний дүнд суурилсан хоёр төрлийн дохионд тохирсон жингийн хувиргалт ба нормчлолын аргыг хэрэглэв. Yelp болон Finance өгөгдөл дээр хийсэн туршилтаар SelfGNN + ирмэгийн жин хувилбар нь SelfGNN суурь загвар болон бусад суурь аргуудтай харьцуулахад HR@10, NDCG@10 хэмжүүрээр хамгийн өндөр үзүүлэлт үзүүлсэн. Мөн бүлгийн шинжилгээнээс ирмэгийн жингийн нөлөө бүх харилцагчид ижил бус, харин идэвхийн түвшнээс хамааран өөр байгааг ажиглав. Ялангуяа өгөгдөл харьцангуй бага бүлгүүдэд ирмэгийн жин нь харилцааны чанарыг ялгах нэмэлт дохио болж байна.

Иймээс энэхүү судалгааны үндсэн дүгнэлт нь ирмэгийн жинг зөв төлөвлөж, зохистой нормчилсон тохиолдолд SelfGNN-д ашигтай нэмэлт мэдээлэл өгч, харилцагч–борлуулагч санал болгох системийн эрэмбэлэлтийн чанарыг сайжруулж чадна гэсэн үр дүн юм. Цаашдын ажилд ирмэгийн жинг тооцох илүү баялаг дохио, өөр нормчлолын хувилбар, мөн гиперпараметрийн системтэй тохируулгыг судлах нь зүйтэй.

Зүүлт

[1] C. Gao et al., ``A Survey of Graph Neural Networks for Recommender Systems: Challenges, Methods, and Directions," *ACM Transactions on Recommender Systems*, vol. 1, no. 1, Article 3, 2023. DOI: 10.1145/3568022.

[2] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang, and M. Wang, ``LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation," in *Proceedings of SIGIR*, 2020.

[3] W.-C. Kang and J. McAuley, ``Self-Attentive Sequential Recommendation," in *Proceedings of IEEE ICDM*, 2018.

[4] F. Sun, J. Liu, J. Wu, C. Pei, X. Lin, W. Ou, and P. Jiang, ``BERT4Rec: Sequential Recommendation with Bidirectional Encoder Representations from Transformer," in *Proceedings of ACM CIKM*, 2019.

[5] Y. Liu, L. Xia, and C. Huang, ``SelfGNN: Self-Supervised Graph Neural Networks for Sequential Recommendation," in *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'24)*, Washington, DC, USA, July 2024, pp. 1609–1618. DOI: 10.1145/3626772.3657716.

[6] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, ``BPR: Bayesian Personalized Ranking from Implicit Feedback," in *Proceedings of UAI*, 2009.

[7] S. Yoo and J. Kim, ``Merchant Recommender System Using Credit Card Payment Data," *Electronics*, vol. 12, no. 811, 2023, doi: 10.3390/electronics12040811.

[8] E. Yilmaz, M. K. Turgut, and S. Ayday, ``A Link Prediction-Based Recommendation System Using Transactional Data," *Scientific Reports*, vol. 13, 2023.

[9] R. van den Berg, T. N. Kipf, and M. Welling, ``Graph Convolutional Matrix Completion," *KDD Workshop on Deep Learning for Recommender Systems*, 2018, arXiv:1706.02263.

[10] I. Sukharev, V. Shumovskaia, K. Fedyanin, M. Panov, and D. Berestnev, ``EWS-GCN: Edge Weight-Shared Graph Convolutional Network for Transactional Banking Data," *IEEE ICDM Workshop*, 2020, arXiv:2009.14588.

[11] V. Shumovskaia, N. Fedyanin, I. Sukharev, D. Berestnev, and M. Panov, ``Linking Bank Clients Using Graph Neural Networks Powered by Rich Transactional Data," *International Journal of Data Science and Analytics*, 2021, doi: 10.1007/s41060-021-00276-w.

ЦАГ ХУГАЦААНЫ ЦУВААХЭЛБЭРИЙН МЭДРЭГЧИЙН ӨГӨГДӨЛ ДЭЭРХ ГАЖУУДАЛ ТОДОРХОЙЛОХ БА ЗҮГШРҮҮЛЭХ НЬ

Нямын Гантөмөр
Компьютерийн ухаан
ШУТИС, Мэдээлэл холбооны технологийн сургууль
Монгол улс, Улаанбаатар хот
B231960002@must.edu.mn

Батаагийн Батзолбоо
Компьютерийн ухаан
ШУТИС, Мэдээлэл холбооны технологийн сургууль
Монгол улс, Улаанбаатар хот
b.batzolboo@must.edu.mn

***Хураангуй*—Орчин үеийн машин сургалтын аргын үр дүн сургалтын өгөгдлийн чанараас шууд хамаардаг бөгөөд мэдрэгч төхөөрөмжөөс цуглуулсан өгөгдөл нь цаг агаар, төхөөрөмжийн тэжээл, бусад төхөөрөмжөөс хамааран тодорхойгүй эсвэл хязгаараас давсан гажуудсан утгууд цөөнгүй агуулсан байдаг. Иймээс машин сургалтын үр дүнг сайжруулахын тулд эдгээр буруу утгуудыг урьдчилан тодорхойлж, зүгшрүүлэх шаардлагатай байдаг. Энэхүү судалгааны ажил нь алдаа илрүүлэх 4, зүгшрүүлэх 10 аргуудыг судалж, Тэтуан хотын эрчим хүчний хэрэглээний жишиг цэвэр өгөгдлийн багц ашиглан, зохиомол алдаатай утгууд үүсгэж, түүнийгээ илрүүлж, зүгшрүүлэн үнэллээ. Судалгааны аргачлал боловсруулж, зохиомол өгөгдлөөс ганц болон дараалсан тодорхойгүй утгыг 8 аргаар зүгшрүүлж экспоненциаль жигнэсэн гулсах дунджийн арга хамгийн сайн үр дүнг үзүүлж дундаж квадрат алдаа 0.04, дундаж абсолют алдаа 0.13, Р квадрат алдаа 0.86 байв. Гажуудал утгуудыг 4 аргаар илрүүлэхэд Лоесс ашиглан улирлын тренд задаргааны илрүүлэх арга 99.2%-ийн нарийвчлалтай (Precision: 0.656, F1-score: 0.61) хамгийн өндөр үнэлгээ үзүүлэв. Илрүүлсэн ганц гажуудал утгыг 7 аргаар зүгшрүүлэхэд мөн экспоненциаль жигнэсэн гулсах дунджийн арга дундаж квадрат алдаа 0.006, дундаж абсолют алдаа 0.05, Р квадрат алдаа 0.96 гэж хамгийн сайн үр дүнтэй байв. Ингэснээр энэхүү судалгааны ажлыг ашиглан, мэдрэгчээс цуглуулсан шуугиантай, тодорхойгүй утгатай, алдаатай утгатай өгөгдлийг зүгшрүүлэх замаар өгөгдлийн багцын алдааг бууруулж, машин сургалтын загварын үр дүнг сайжруулах ач холбогдолтой.**

***Түлхүүр үг*—цаг хугацааны цуваа, өгөгдөл зүгшрүүлэх, өгөгдлийн алдаа илрүүлэх, интерполяцийн арга, дунджийн арга, өгөгдлийн багц, мэдрэгч**

I. УДИРТГАЛ

Орчин үед цаг хугацааны цуваа хэлбэрийн өгөгдөл ашиглан эдийн засгийн төлөвлөлт, нийлүүлэлтийн сүлжээний удирдлага, эмнэлгийн оношилгоо, авто замын ачааллын удирдлага гэх мэт чухал асуудлуудыг машин сургалтын аргаар шийдэж байна [1]. Машин сургалт болон гүн сургалтын загваруудын зөв таамаглах нарийвчлалын утга нь сургалтад ашиглаж буй өгөгдлийн чанараас шууд хамаардаг [2]. Юмсын интернэт хөгжсөн одоо үед цаг хугацааны цуваа хэлбэрийн өгөгдлийг ихэвчлэн мэдрэгчээр бодит цаг хугацаанд цуглуулдаг [1].

Цаг уурын станц, ухаалаг тоолуур гэх мэт мэдрэгчээс цуглуулсан цаг хугацааны цуваа хэлбэрийн өгөгдөл нь дараалсан, цаг хугацааны хамааралтай, хугацааны мөчлөгүүдийг агуулдгаараа хүснэгтэн өгөгдлөөс ялгаатай байна [3] [4]. Мэдрэгч төхөөрөмжөөс ирсэн утга нь хүрээлэн буй орчны цаг агаар, төхөөрөмжийн тэжээл, сүлжээний төхөөрөмжийн дамжуулж буй пакетын алдагдал гэх мэт нөхцөл байдлаас хамааран дараах нийтлэг алдаатай хадгалагддаг [5] [6]. Жишээлбэл:

1. Тодорхойгүй утга: Мэдрэгч унтрах эсвэл өгөгдөл дамжуулах үед тасалдал гарснаас үүдэн хоосон утга үүсдэг.
2. Хэвийн бус эсвэл гажуудсан утга (Цаашид гажуудсан утга гэе): Хэвийн утгаас эрс ялгаатай, физик боломжит утгын хязгаараас давсан утга үүсдэг.
3. Гауссын шуугиантай утга: Хэмжилтийн төхөөрөмжийн өөрийнх нь нарийвчлалын алдаанаас үүссэн хэлбэл үүсдэг.

Эдгээр алдаатай өгөгдлийг урьдчилан боловсруулахгүй машин сургалтад ашиглах нь сургалтын үр дүнг мэдэгдэхүйц бууруулах боломжтой. Учир нь сургалтын загвар зүй тогтлыг сурахгүй, харин алдаатай орсон онцлог утгуудад мэдрэг болж сурдаг.

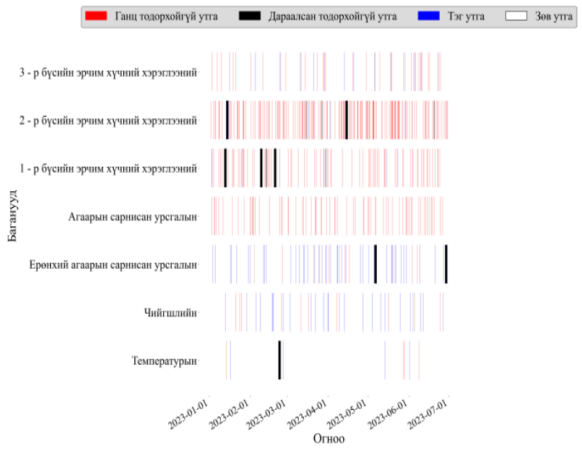
Иймээс цаг хугацааны цувааны өгөгдлийн урьдчилсан боловсруулалт нь алдаатай утгуудыг илрүүлэхээс гадна алдаатай утгыг нөхөн, зүгшрүүлэх интерполяцийн аргуудаас бүрддэг [7] [8].

Энэхүү судалгааны ажлын зорилго нь цаг хугацааны цувааны өгөгдлийн тодорхойгүй болон гажуудсан алдаатай утга илрүүлэх аргуудыг харьцуулах болон илрүүлсэн алдааг бодит утгатай хамгийн ойролцоо утгаар зүгшрүүлэх аргуудыг үнэлэхэд оршино. Бид гауссын шуугиантай утга илрүүлэх, зүгшрүүлэх талаар судалгаа, туршилт хийгээгүй болно.

II. СУДЛАГДСАН БАЙДАЛ

Мэдрэгчээс цуглуулсан өгөгдлийн багцад тэг утгууд, ганц тодорхойгүй утгын алдаа, дараалсан тодорхойгүй утгын алдаа гэсэн гурван тусдаа асуудлууд байдаг болохыг

Зураг 1 – д үзүүлэв [9]. Иймээс хоорондоо холбоо бүхий өгөгдлийн урьдчилсан боловсруулалт, гажуудал илрүүлэлт, тодорхойгүй өгөгдлийг нөхөх гэсэн гурван дараалсан алхмаар туршилтыг гүйцэтгэсэн байна. Иймээс алхам бүрд үр дүнг цагийн цувааны график зурж нүдээр болон статистик үзүүлэлт гэсэн хоёр аргаар шалгасан байна [9].

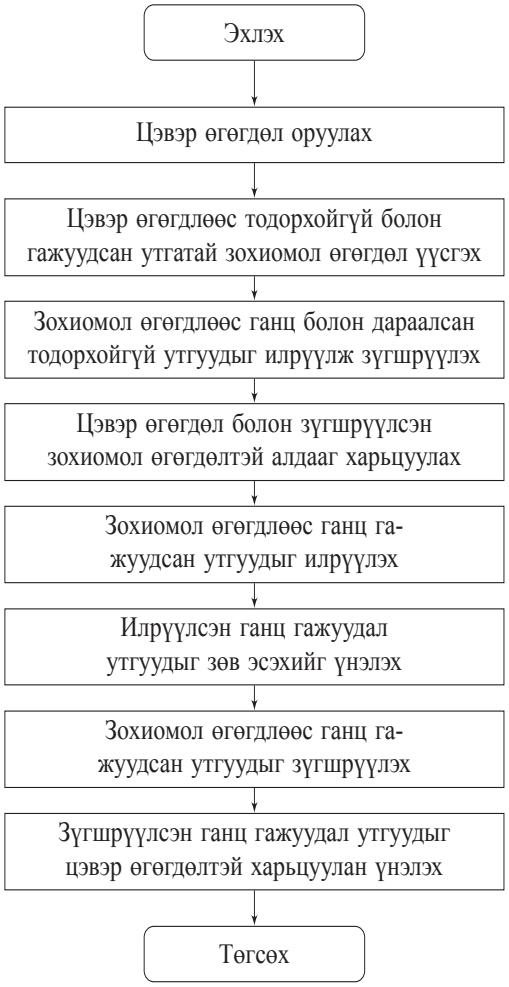


Зураг 1. Тэг утгын алдаа, ганц тодорхойгүй утгын алдаа, дараалсан тодорхойгүй утгын алдаа

Мэдрэгчийн өгөгдөлд тэг утга байх нь илрүүлэлтийн хязгаараас доогуур утга, эсвэл мэдрэгчийн эвдрэлийг илэрхийлж болох тул зөв хэмжилт шиг харагддаг. Мэдрэгчийн алдааны горимын мэдээлэл дутмаг байгаагаас шалтгаалан эдгээр утгууд цуглардаг. Иймээс тэг утгыг мөн тодорхойгүй утгатай адил гэж үзэн тодорхойгүй утгатай хамт боловсруулсан байна [9]. Бид энэхүү ажилдаа тэг утгыг гажуудсан буюу хэвийн утгаас эрс ялгаатай утга гэж үзсэн бөгөөд илрүүлэх, зүгшрүүлэхдээ мөн тодорхойгүй утга болгож туршсан.

Энэхүү судалгааны ажилдаа тодорхойгүй болон гажуудсан утгагүй цэвэр өгөгдлийн багцад эдгээр алдаатай утгуудыг ганц болон дараалсан хэлбэрээр зохиомлоор үүсгэж илрүүлэлт, зүгшрүүлэлтийн аргуудыг харьцуулав [9].

Нэг өгөгдлийн багц дахь тодорхойгүй утгыг илрүүлж, зүгшрүүлсний дараа гажуудсан утга илрүүлэх нь судалгааны үр дүнг бууруулдаггүй болохыг баталсан байна [3] [9]. Иймээс бид Зураг 2-д үзүүлсэн судалгааны аргачлалаар гүйцэтгэв. Туршилт хийх цэвэр өгөгдлийн багц оруулж, цэвэр өгөгдлийн багцаас ганц болон дараалсан байдлаар тодорхойгүй утгууд болон ганц гажуудсан утгыг санамсаргүйгээр оруулсан зохиомол өгөгдөл багц үүсгэнэ. Дараа нь үүсгэсэн зохиомол өгөгдлийн багцаас ганц болон дараалсан тодорхойгүй утгуудыг илрүүлж, зүгшрүүлнэ. Ингэснээр цэвэр өгөгдөлтэй харьцуулж үнэлэх боломжтой болно. Ингэхдээ дундаж квадрат алдаа (MSE), дундаж абсолют алдаа (MSE), Р квадрат алдаагаар (R2) үнэлэв. Үүний дараа зохиомол өгөгдлийн багцаас ганц гажуудсан утгуудыг илрүүлж, тэдгээр гажуудсан утгуудыг зөв, буруу илрүүлсэн гэдгийг шалгаж байгаа тул нарийвчлал(accuracy), precision, F1 үнэлгээгээр үнэлэв. Ингэд эдгээр илрүүлсэн ганц гажуудсан утгуудыг зүгшрүүлж, цэвэр өгөгдөлтэй дундаж квадрат алдаа, дундаж абсолют алдаа, Р квадрат алдаагаар харьцуулав.



Зураг 2. Гүйцэтгэсэн судалгааны аргачлал

III. ӨГӨГДЛИЙН БАГЦ

Туршилтыг үнэн зөв явуулахын тулд Тэтуан хотын эрчим хүчний хэрэглээний жишиг өгөгдлийн багц ашиглав [10]. Энэхүү өгөгдлийн багц нь 10 минутын давтамжтайгаар бүртгэгдсэн температурын хэмжүүр, агаарын чийгшил, салхины хурд, агаарын урсгалын хурд зэрэг цаг уурын үзүүлэлтүүдээс гадна тухайн хотын гурван ялгаатай бүсийн эрчим хүчний хэрэглээний хэмжээг агуулсан байдаг. Эдгээр өгөгдлийн утгууд нь тодорхойгүй болон гажуудалгүй утгууд байхгүй болохыг харуулсан байдаг [11] [12].

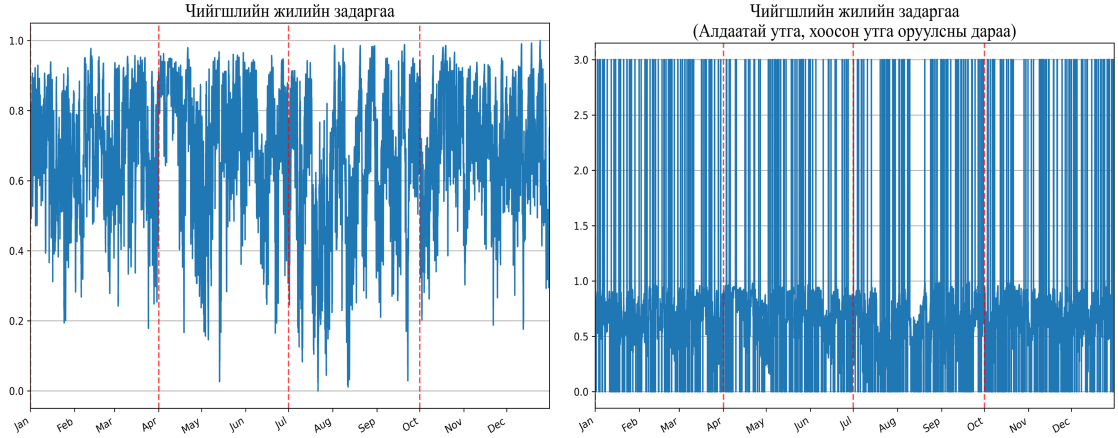
Цэвэр өгөгдлөөс зохиомол өгөгдлийн багц үүсгэхдээ дараах байдлаар үүсгэв. Үүнд:

- Ганц тодорхойгүй утга үүсгэхдээ санамсаргүй байдлаар сонгосон цэгүүдийн утгыг тодорхойгүй утгаар солино.
- Дараалсан тодорхойгүй утга үүсгэхдээ эхний цэгийг санамсаргүй сонгож, дараагийн 4 цаг буюу 24 мөр утгыг тодорхойгүй утгаар солино.

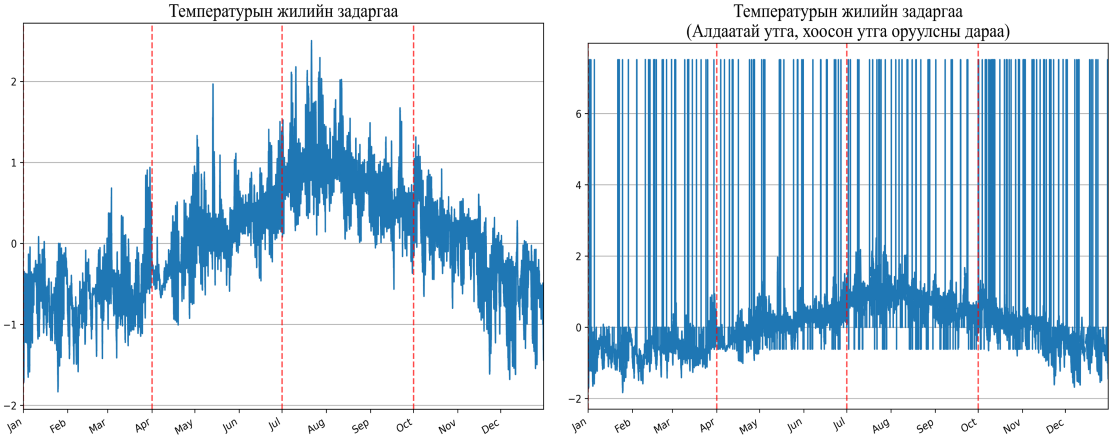
- Ганц гажуудсан утга үүсгэхдээ нийт сургалтын өгөгдлийн тооны 5 хувь байхаар санамсаргүй нэг цэгүүд сонгож, хамгийн ихээс 4 дахин их эсвэл хамгийн багаас 4 дахин бага утгаар солино.

Цэвэр болон зохиомол өгөгдлийн температурын утгын тархалтыг Зураг 3 – р, чийгшлийн утгын тархалтыг Зураг 4 – р, эрчим хүчний хэрэглээний утгын тархалтыг Зураг 5 – д тус тус харуулав.

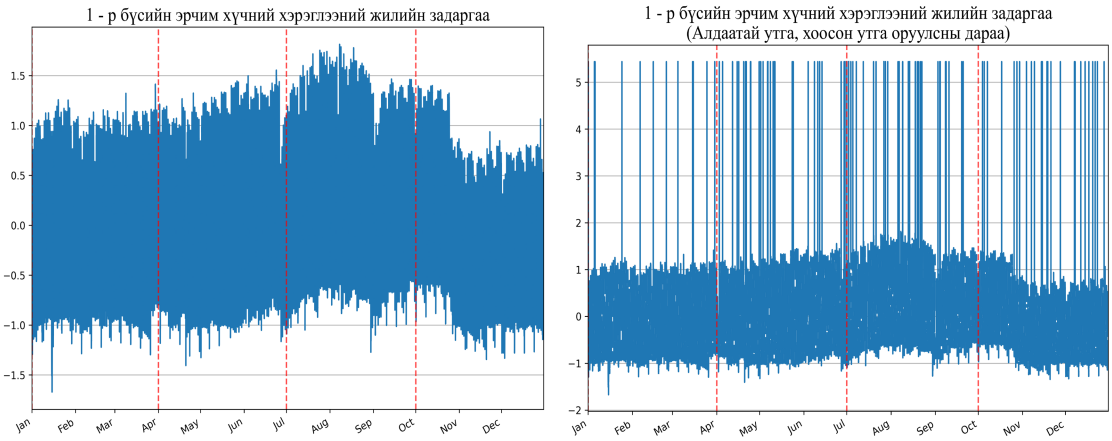
1



Зураг 3. Цэвэр болон зохиомол өгөгдлийн температурын утгын тархалтын харьцуулалт



Зураг 4. Цэвэр болон зохиомол өгөгдлийн чийгшлийн утгын тархалтын харьцуулалт



Зураг 5. Цэвэр болон зохиомол өгөгдлийн нэгдүгээр бүсийн эрчим хүчний хэрэглээний утгын тархалтын харьцуулалт

IV. АЛДАА ИЛРҮҮЛЭХ АРГУУД

Энэхүү судалгаанд өгөгдлийн багцаас тодорхойгүй утгыг илрүүлэхдээ логик масклалт ашигласан [9] [13]. Харин өгөгдлийн багцаас гажуудсан утгыг илрүүлэх дараах 4 аргыг ашигласан [14].

A. Квартил хоорондын далайцын арга (Interquartile range - IQR)

$$IQR = Q_3 - Q_1 \tag{1}$$

Квартил хоорондын далайцын арга нь статистикийн сарнилыг хэмждэг бөгөөд нийт утгын 50% нь медианаас $\pm 0.5 * IQR$ хязгаарт багтаж буйг илэрхийлдэг [15]. Үүнд:

- Q_1: нэгдүгээр квартил;
- Q_3: гуравдугаар квартил.

B. Гулсах квартил хоорондын далайцын арга (Moving interquartile range - IQR)

$$MovingIQR(t) = Q_3(X_{t-W+1:t}) - Q_1(X_{t-W+1:t}) \tag{2}$$

Гулсах квартил хоорондын далайцын арга нь нэг болон гуравдугаар квартилыг тодорхой хугацааны интервалаар тасралтгүй шилжүүлэн тооцдог. Цаг хугацааны цувааны өгөгдлийг нийт хугацааны хүрээнд шууд авч үзэхийн оронд хэсэгчилсэн хугацааны дарааллаар шинжилдэг нь тухайн интервалын дотоод өөрчлөлтөд дасан зохицох чадвартай юм [16]. Үүнд:

- $Q_3(X_{t-W+1:t})$: t хугацаа хүртэлх W урттай хугацааны шилжих цонхонд багтах өгөгдлийн нэгдүгээр квартил;
- $Q_1(X_{t-W+1:t})$: t хугацаа хүртэлх W хугацааны шилжих цонхонд багтах өгөгдлийн гуравдугаар квартил;
- W: шилжин явах цонхны утга.

C. Лоесс ашиглан улирлын тренд задлах арга - (Seasonal-Trend decomposition using LOESS - STL)

$$Y_u = T_v + S_v + S_v \tag{3}$$

Лоесс ашиглан улирлын тренд задлах арга нь цаг хугацааны цуваа өгөгдлийн тренд, улирлын хэлбэлзэл, болон үлдэгдэл гэсэн гурван хэсэгт задалдаг [17]. Үүнд:

- T_v: цаг хугацааны цуваа хэлбэрийн өгөгдлийн ерөнхий хандлага;
- S_v: улирлын хэлбэл;
- R_v: үлдэгдэл.

D. Гажуудсан утга тусгаарлах арга (Isolation Forest)

$$s(x, n) = 2^{-\frac{E(h(x))}{c(n)}} \tag{4}$$

Гажуудсан утга тусгаарлах арга нь хэвийн бус өгөгдлийн тоо хэвийн өгөгдлийн тооноос цөөн бөгөөд эрс ялгаатай байдаг онцлогт тулгуурлан, хоёртын модонд өгөгдлийн байрлуулж, томъёо (4) -т үзүүлсэн байдлаар үнэлгээг тооцоолно. 1 утгад дөхсөн үнэлгээг хэвийн бус тохиолдол гэж үзнэ [18]. Үүнд:

- $H(x)$: модны үндсээс навч хүртэлх салаалалтын тоо;
- $C(n)$: модны үндсээс n ширхэг зангилаа хүртэлх дундаж замын урт;
- $E(h(x))$: модны үндсээс тухайн х өгөгдөл бүхий зангилаа хүртэлх дундаж замын урт.

V. АЛДАА ЗҮГШРҮҮЛЭХ АРГУУД

Энэхүү судалгааны ажилд алдааг зүгшрүүлэхдээ дунджийн болон интеполяцийн гэсэн хоёр төрлийн нийт 10 арга ашигласан.

A. Шугаман интерполяцийн арга (Linear Interpolation)

$$y = y_1 + (x - x_1)(\frac{y_2 - y_1}{x_2 - x_1}) \tag{5}$$

Шугаман интерполяцийн арга нь мэдэгдэхгүй байгаа дутуу утгыг зүгшрүүлэхдээ, зөвхөн түүний яг өмнөх болон дараах, хоёр мэдэгдэж буй утгыг геометрийн шулуун шугамаар холбох замаар тооцоолон зүгшрүүлдэг. Үүнд:

- $H(x)$: модны үндсээс навч хүртэлх салаалалтын тоо;
- $C(n)$: модны үндсээс n ширхэг зангилаа хүртэлх дундаж замын урт;
- $E(h(x))$: модны үндсээс тухайн х өгөгдөл бүхий зангилаа хүртэлх дундаж замын урт.

B. Энгийн гулсах дунджийн арга (Simple Moving Average)

$$SMA = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{6}$$

Энгийн гулсах дунджийн арга нь дараагийн утгыг тодорхойлохдоо өмнөх мэдэгдэж буй n цонхны хэмжээтэй утгуудын дунджаар мэдэгдэхгүй утгыг зүгшрүүлдэг.

C. Эсрэг чиглэлт гулсах дунджийн арга (Reverse Moving Average)

$$RMA_t = \frac{1}{n}\sum_{i=0}^{n-1} X_{t-i} \tag{7}$$

Эсрэг чиглэлт гулсах дунджийн арга нь дараагийн утгыг тодорхойлохдоо тухайн хоосон t үеэс хойших мэдэгдэж n цонхны хэмжээтэй утгуудын дунджаар тухайн t дэх мэдэгдэхгүй буй утгыг зүгшрүүлдэг [19].

D. Төвийн гулсах дунджийн арга (Centered Moving Average)

$$CMA_t = \frac{1}{2k}\left(\sum_{i=1}^{k} X_{t-i} + \sum_{i=1}^{k} X_{t+i}\right) \quad (8)$$

Төвийн гулсах дунджийн арга нь тухайн t үе дахь мэдэгдэхгүй байгаа утгыг t дэх утгаас өмнөх болон хойших k цонхны хэмжээтэй утгуудын дунджаар зүгшрүүлдэг.

E. Давталтат төвийн гулсах дунджийн арга (Iterative Centered Moving Average)

$$X_t^{(j)} = \frac{1}{2k}\left(\sum_{i=1}^{k} X_{t-i}^{(j-1)} + \sum_{i=1}^{k} X_{t+i}^{(j-1)}\right) \quad (9)$$

Давталтат төвийн гулсах дунджийн арга нь тухайн t үе дэх мэдэгдэхгүй байгаа утгыг t дэх утгаас өмнөх болон хойших k цонхны хэмжээтэй утгуудын дунджаар зүгшрүүлдэг ба тухайн үйлдлийг j удаа давтаж зүгшрүүлдэг [20]. Үүнд:

- $X_t^{(j)}$: j – р давталт дээрх t хугацаанд шинээр тооцоологдон зүгширч буй утга;
- $X_{t-i}^{(j-1)}, X_{t+i}^{(j+1)}$: өмнөх буюу $(j-1)$ – р давталтын үр дүнд тооцоологдсон, тухайн дутуу утгын өмнөх болон дараах хөрш цэгүүдийн утгууд;
- k: цонхны хэмжээ.

F. Экспоненциаль жигнэсэн гулсах дунджийн арга(Exponentially Weighted Moving Average)

$$EWMA_t = aX_t + (1-a)EWMA_{t-1} \quad (10)$$

Экспоненциаль жигнэсэн гулсах дунджийн арга нь цаг хугацааны цуваа хэлбэрийн өгөгдөл дэх санамсаргүй хэлбэлзэл, шуугианыг жигдрүүлж үндсэн трендийг илрүүлэх зорилготой статистик арга юм. Энэ арга нь хамгийн сүүлийн жишээнүүдэд хамгийн өндөр жинг өгдөг бөгөөд хуучин өгөгдлийн нөлөөг цаг хугацаа өнгөрөх тусам экспоненциаль хэлбэрээр тасралтгүй бууруулж тооцдог [21]. Үүнд:

- $EWMA_t$: тухайн t хугацаа дахь тооцоолсон шинэ гулсах дунджийн утга;
- X_t: тухайн t хугацаа дахь бодит оролтын утга;
- $EWMA_{t-1}$: өмнөх $(t–1)$ хугацааны тооцоолсон гулсах дунджийн утга.

G. Сплайн интерполяцийн арга (Spline Interpolation)

$$S_i(x) = a_t(x-x_i)^3 + b_i(x-x_i)^2 + c_i(x-x_i) + d_i \quad (11)$$

Сплайн интерполяцийн арга нь өгөгдлийн цэгүүдийн хооронд гурван гишүүнт функц байгуулах замаар дутуу утгыг зүгшрүүлдэг. Үүнд:

- $S_t(x)$: x цэг дээрх тооцоолж буй утга;
- S_i: мэдэгдэж буй хугацааны цэг;
- a_i, b_i, c_i, d_i: тухайн интервал бүрд тооцоолох коеффициентууд.

H. АКИМА интерполяцийн арга (AKIMA Interpolation)

$$S(x) = a_0 + a_1(x-x_i) + a_2(x-x_i)^2 + a_3(x-x_i)^3 \quad (12)$$

$$t_i = \frac{|m_{i+1} - m_i|m_{i-1} + |m_{i-1} - m_{i-2}|m_i}{|m_{i+1} - m_i| + |m_{i-1} - m_{i-2}|} \quad (13)$$

АКИМА интерполяцийн арга нь өгөгдлийн цэгүүдийн хооронд гурван гишүүнт функц байгуулах замаар дутуу утгыг зүгшрүүлдэг. Мөн тухайн арга нь тухайн интервалын таван цэгийн налууд тулгуурлан тооцоолол хийдэг. Үүнд:

- t_i, x_i: цэг дээрх муруйн налуу бөгөөд налууг тооцохдоо зэргэлдээх хэсгүүдийн налууг жигнэсэн дундаж хэлбэрээр тооцоолдог (13);
- m_i: өгөгдлийн зэргэлдээх цэгүүдийг холбосон шулууны налуу.

I. Хэсэгчилсэн куб эрмитийн олон гишүүнт интерполяцийн арга (Piecewise Cubic Hermite Interpolating Polynomial)

$$P(x) = f_i h_{00}(t) + d_i \Delta x_i h_{10}(t) + f_{i+1} h_{01}(t) + d_{i+1} \Delta x_i h_{11}(t) \quad (14)$$

Хэсэгчилсэн куб эрмитийн олон гишүүнт интерполяцийн арга нь өгөгдлүүдийн цэгүүдийн хооронд гурван гишүүнт байгуулах замаар зүгшрүүлдэг байдлаараа Акима интерполяцийн аргатай төстэй боловч гол ялгаатай байдал нь тухайн арга нь өгөгдлийн өсөх эсвэл буурах чиг хандлагыг дагадаг. Мөн тухайн арга нь $[x_i, x_i + 1]$ интервал бүрд гурван гишүүнт байгуулдаг.

Үүнд:

- $P(x)$: x цэг дээрх зүгшрүүлж буй утга;
- f_i, f_{i+1}: өгөгдлийн мэдэгдэж буй утгууд;
- d_i: тухайн цэг дээрх муруйн налуу;
- t: тухайн интервал дахь харьцангуй байршил;
- $h_{00}, h_{01}, h_{10}, h_{11}$: эрмитийн суурь функцууд.

J. К-ойрын хөршийн интерполяцийн арга (K-Nearest Neighbors Interpolation)

$$X_t = \frac{\sum_{i=1}^{k} w_i X_{NN_i}}{\sum_{i=1}^{k} w_i} \quad (15)$$

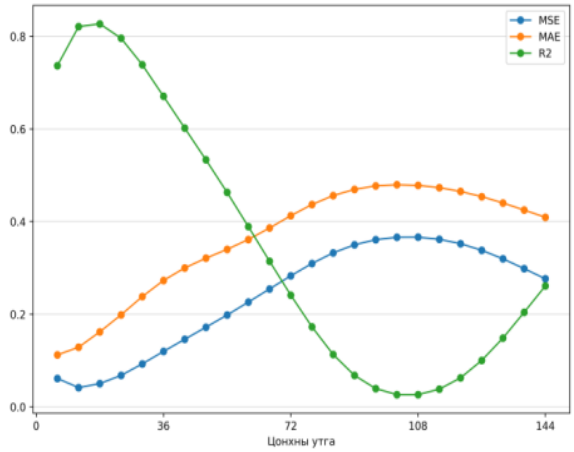
K-ойрын хөршийн интерполяцийн арга нь цаг хугацааны цуваа хэлбэрийн өгөгдлийн дутуу утгыг зүгшрүүлэхдээ тухайн дутуу хэсэгтэй хамгийн төстэй шинж чанарыг агуулсан түүхэн хөрш өгөгдлүүдийг хайж олох замаар тооцоолдог [22].

Үүнд:

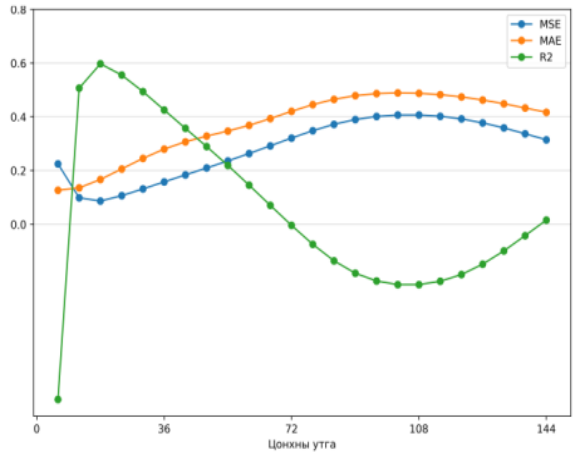
- X_t: t хугацаанд зүгшрүүлж буй утга;
- k: ойрын хөршийн тоо;
- X_{NN_i}: тухайн цэгтэй хамгийн төстэй i – р хөрш;
- w_i: тухайн хөршийн жин.

VI. ТУРШИЛТ БА ҮР ДҮН

Зураг 2 дахь судалгааны аргачлалд өгүүлснээр зохиомол өгөгдлөөс ганц болон дараалсан тодорхойгүй утгыг зүгшрүүлэхдээ 10 арга ашигласан. Эдгээр аргуудаас цонх ашиглан зүгшрүүлдэг 6 арга байгаагаас гулсах дунджийн болон эсрэг чиглэлт гулсах дунджийн арга нь дараалсан тодорхойгүй утгыг зүгшрүүлж чадахгүй байсан. Бусад 4 цонх ашигладаг аргуудын үр дүнгийн цонхны хэмжээ болон алдааны хэмжээний хамаарлыг Зураг 6, Зураг 7, Зураг 8, Зураг 9 тус бүр үзүүлэв. Эдгээр зургаас цонхны оновчтой хэмжээг дараах байдлаар тодорхойллоо. Ингэхдээ P квадрат алдааг хамгийн их, дундаж квадрат болон дундаж абсолют алдааг аль болох бага гэж үзвэл цонхны оновчтой хэмжээ 12 байна.



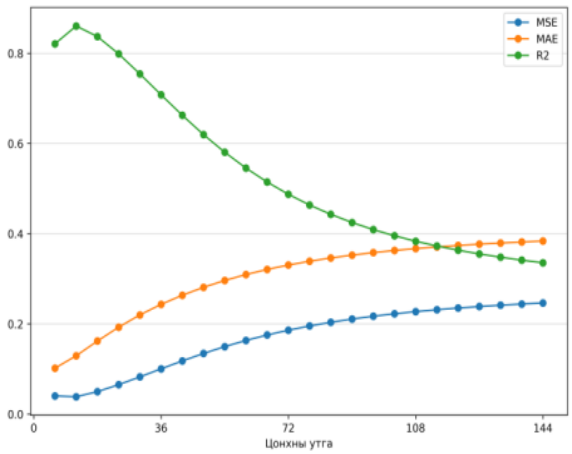
Зураг 6. Төвийн гулсах дунджийн интерполяцийн аргын алдаа



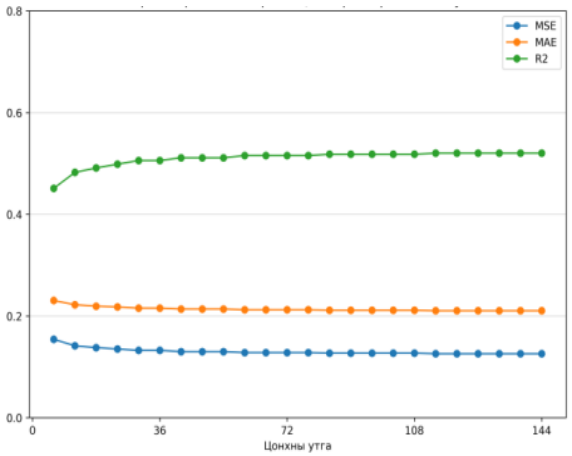
Зураг 7. Давталттай төвийн гулсах дунджийн аргын алдаа

Хүснэгт I-т ганц болон дараалсан тодорхойгүй утгын зүгшрүүлэлтийн алдааг үзүүлэв.

Зураг 2 дахь судалгааны аргачлалд өгүүлснээр ганц гажуудал утгуудыг 4 аргаар илрүүлсэн. Эдгээр 4 аргаас цонх ашиглаж илрүүлдэг гулсах квартил хоорондын далайцын болон Лоесс ашиглан улирлын тренд задаргааны илрүүлэх аргуудын цонхны хэмжээ болон нарийвчлалын үнэлгээ хоорондын хамаарлыг Зураг 10, зураг 11 – т үзүүлэв.



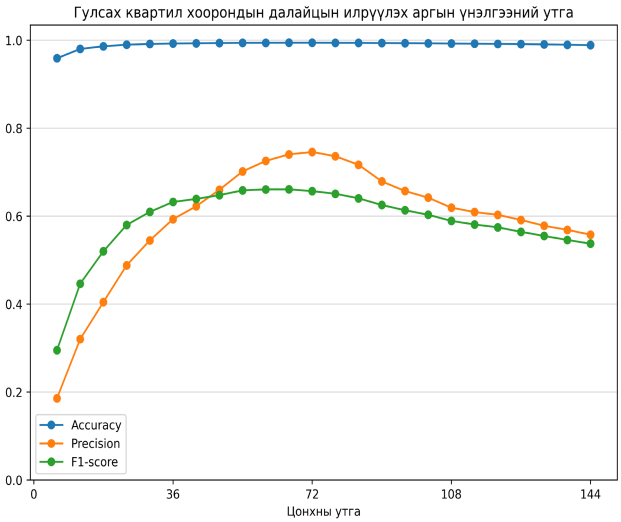
Зураг 8. Экспоненциал жигнэсэн гулсах дунджийн аргын алдаа



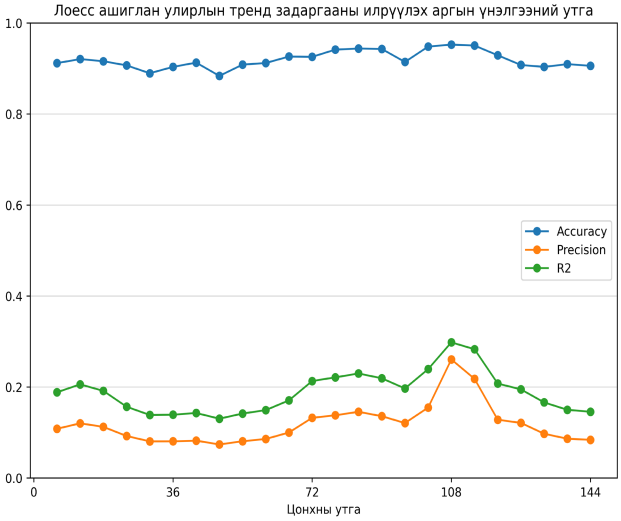
Зураг 9. К-ойрын хөршийн интерполяцийн аргын алдаа

ХҮСНЭГТ I. Ганц болон дараалсан тодорхойгүй зүгшрүүлэлтийн алдаа

Арга	***MSE***	***MAE***	R^2
Төвийн гулсах дунджийн арга (цонхны тоо=12)	0.04	0.13	0.82
Экспоненциаль жигнэсэн гулсах дунджийн арга (цонхны тоо=12)	***0.04***	***0.13***	***0.86***
Давталтат төвийн гулсах дунджийн арга (цонхны тоо=18)	0.09	0.17	0.60
KNN(хөршийн тоо=60)	0.13	0.21	0.52
Шугаман интерполяцийн арга	0.251	0.131	0.578
Акима интерполяцийн арга	0.2811	0.1492	0.6972
Хэсэгчилсэн куб интерполяцийн арга	0.27935	0.13736	0.6942
Сплайн интерполяцийн арга	0.28	0.1491	0.6972



Зураг 10. Гулсах квартил хоорондын далайцын илрүүлэх аргын үнэлгээ



Зураг 11. Лоесс ашиглан улирлын тренд задраргааны илрүүлэх аргын үнэлгээ

Зураг 10-аас үзвэл гулсах квартил хоорондын далайцын аргаар тооцоолоход цонхны хэмжээ 72 байхад хамгийн оновчтой байсан. Харин Зураг 11-ээс үзвэл Лоесс ашиглан улирлын тренд задраргааны илрүүлэх аргаар тооцоолоход цонхны хэмжээ 108 байхад хамгийн оновчтой байсан. 4 аргаар гажуудал утгыг зөв илрүүлсэн эсэхийг үнэлсэн байдлыг Хүснэгт II-т үзүүлэв.

Зураг 2 дахь судалгааны аргачлалд өгүүлснээр илрүүлсэн ганц гажуудал утгыг 7 аргаар зүгшрүүлэв. Туршилтын үр дүнг Хүснэгт III-т үзүүлэв.

ХЭЛЭЛЦҮҮЛЭГ

Цаг хугацааны цуваа хэлбэрийн өгөгдөл олон төрлийн бодлогыг шийддэг. Тухайн бодлогын төрлөөс хамааран өгөгдлийн цаг хугацааны хамаарлын шинж чанар, зүй тогтол, түүнд нөлөөлөх хүчин зүйлс өөр өөр байдаг. Туршилтад ашигласан алдаа илрүүлэх, зүгшрүүлэх аргууд нь

ХҮСНЭГТ II. Гажуудал утгын илрүүлэлтийн үнэлгээ

Арга	*Нарийвчлал*	*Precision*	*F1 - оноо*
Квартил хоорондын далайц	0.9856	0.624	0.53
Гулсах квартил хоорондын далайцын арга(цонхны хэмжээ=72)	0.9864	**0.745**	**0.656**
Лоесс ашиглан улирлын тренд задлах арга(цонхны тоо=18)	**0.992**	0.656	0.613
Гажуудсан утга тусгаарлах арга	0.954	0.663	0.547

ХҮСНЭГТ III. Гажуудал утгын зүгшрүүлэлтийн алдаа

Арга	*MSE*	*MAE*	R^2
Экспоненциаль жигнэсэн гулсах дунджийн арга (цонхны тоо=12)	***0.00642***	***0.0508***	***0.967***
Давталтад төвийн гулсах дунджийн арга (цонхны тоо=18)	0.00633	0.0537	0.9524
Хэсэгчилсэн куб интерполяцийн арга	0.00832	0.0631	0.9641
Акима интерполяцийн арга	0.00795	0.0593	0.9573
Сплайн интерполяцийн арга	0.00616	0.0542	0.9542
Төвийн гулсах дунджийн арга (цонхны тоо=12)	0.00743	0.0645	0.9467
Давталтад төвийн гулсах дунджийн арга (цонхны тоо=36)	0.00831	0.0651	0.96203

зөвхөн Тэтуан хотын цаг уур болон эрчим хүчний хэрэглээний жишиг өгөгдлийн багц дээр туршигдсан бөгөөд энэ нь цаг хугацааны хэлбэрийн өгөгдлийн багцын тухайн нэг төрлийн бодлогын төлөөлөл юм. Харин туршсан аргуудын харьцуулалт нь ерөнхий байдлын нийцлийг хязгаарлаж болзошгүй юм.

Цуваа хэлбэрийн өгөгдөлтэй ажилладаг гүн сургалтын архитектурт загварууд нь олон хэмжээст их хэмжээний өгөгдөл, нарийн параметрүүдийн тохируулгыг шаарддаг бөгөөд бидний ашигласан аргууд нь тооцооллын хүндрэл харьцангуй бага хөнгөн байх боломжтой. Гэсэн хэдий ч гүн сургалтын архитектурт загваруудыг харьцуулан үнэлэх нь ач холбогдолтой байгаа бөгөөд ирээдүйн ажилд тусгагдсан болно.

ДҮГНЭЛТ

Энэхүү судалгааны ажлаар цаг хугацааны цуваа, тэр дундаа хүрээлэн буй орчны болон эрчим хүчний мэдрэгчийн өгөгдөл дэх гажуудлыг илрүүлэх, тодорхойгүй утгыг зүгшрүүлэх холимог аргачлалыг туршиж, үр дүнг харьцуулан дүгнэлээ. Үүнд:

1. Тодорхойгүй утгыг зүгшрүүлэх үед Хүснэгт I-ээс үзвэл сонгож авсан цэвэр өгөгдлийн багцын хувьд экспоненциаль жигнэсэн гулсах дунджийн арга хамгийн сайн үр дүн үзүүлж, дундаж квадрат алдаа 0.04, дундаж абсолют алдаа 0.13, Р квадрат алдаа 0.86 байв.

2. Гажуудал илрүүлэх үед уламжлалт босго утгатай квартил хоорондын далайцын болон гажуудсан утга тусгаарлах аргууд нь цаг хугацааны цувааны динамик савлагааг дагаж чадахгүй False Positive алдааг олон гаргаж байв. Харин хөдөлгөөнт босго утга бүхий гулсах квартил хоорондын далайцын болон Лоесс ашиглан улирлын тренд задлах аргууд өндөр үр дүнг үзүүлэв. Хүснэгт II-оос үзвэл Лоесс ашиглан улирлын тренд задлах арга цонхны хэмжээ 108 байхад 99.2%-ийн нарийвчлалтай (Precision: 0.656, F1-score: 0.61) хамгийн өндөр үнэлгээ үзүүлэв.

3. Илрүүлсэн гажуудал утга зүгшрүүлэх үед Хүснэгт III-аас үзвэл мөн экспоненциаль жигнэсэн гулсах дунджийн арга хамгийн сайн үр дүн үзүүлж, дундаж квадрат алдаа 0.006, дундаж абсолют алдаа 0.05, Р квадрат алдаа 0.96 болов.

Цаашид энэхүү судалгааны ажлыг ашиглан, мэдрэгчээс цуглуулсан шуугиантай, тодорхойгүй утгатай, алдаатай утгатай өгөгдлийг зүгшрүүлэх замаар өгөгдлийн багцын алдааг бууруулж, машин сургалтын загварын үр дүнг сайжруулах ач холбогдолтой. Мөн цаг агаарын болон эрчим хүчний хэрэглээний цуваа өгөгдөл дээр эрчим хүчний хэрэглээг урьдчилан таамаглах загваруудыг сургах боломжтой.

Ном зүй

[1] Kim, J., Kim, H., Kim, H. et al. A comprehensive survey of deeplearning for time series forecasting: architectural diversity and openchallenges.ArtifIntellRev58,216(2025).https://doi.org/10.1007/s10462-025-11223-9

[2] Sedir Mohammed, Lukas Budach, Moritz Feuerpfeil, Nina Ihde,Andrea Nathansen, Nele Noack, Hendrik Patzlaff, Felix Naumann,Hazar Harmouch, The effects of data quality on machine learningperformance on tabular data, Information Systems, Volume 132,2025,102549,ISSN0306-4379,https://doi.org/10.1016/j.is.2025.102549.

[3] Hu, C., Sun, Z., Li, C., Zhang, Y., & Xing, C. (2023). Survey of Time Series Data Generation in IoT. Sensors (Basel, Switzerland), 23(15), 6976. https://doi.org/10.3390/s23156976

[4] Alejo-Sanchez, L. E., Márquez-Grajales, A., Salas-Martínez, F., Franco-Arcega, A., López-Morales, V., Acevedo-Sandoval, O. A., González-Ramírez, C. A., & Villegas-Vega, R. (2025). Missing data imputation of climate time series: A review. MethodsX, 15, 103455. https://doi.org/10.1016/j.mex.2025.103455

[5] Sheetal Sharma, Kamali Gupta, Deepali Gupta, Shalli Rani, Gaurav Dhiman, An Insight Survey on Sensor Errors and Fault Detection Techniques in Smart Spaces, CMES - Computer Modeling in Engineering and Sciences, Volume 138, Issue 3, 2023,Pages 2029-2059,ISSN 1526-1492, https://doi.org/10.32604/cmes.2023.029997.

[6] van Zoest, V., Liu, X., & Ngai, E. (2021). Data Quality Evaluation, Outlier Detection and Missing Data Imputation Methods for IoT in Smart Cities. In Machine Intelligence and Data Analytics for Sustainable Future Smart Cities (pp. 1–18). https://doi.org/10.1007/978-3-030-72065-0_1

[7] Zhao, J.H., Dong, Z., Xu, Z. (2006). Effective Feature Preprocessing for Time Series Forecasting. In: Li, X., Zaïane, O.R., Li, Z. (eds) Advanced Data Mining and Applications. ADMA 2006. Lecture Notes in Computer Science(), vol 4093. Springer, Berlin, Heidelberg. https://doi.org/10.1007/11811305_84

[8] Bejani, M., Mauri, M., & Mariani, S. (2025). Self-Attention-Based Deep Learning for Missing Sensor Data Imputation in Real-Time Probe Card Monitoring. Sensors, 25(23), 7194. https://doi.org/10.3390/s25237194

[9] AlSalehy, A.S.; Bailey, M. Improving Time Series Data Quality: Identifying Outliers and Handling Missing Values in a Multilocation Gas and Weather Dataset. Smart Cities 2025, 8, 82. https://doi.org/10.3390/smartcities8030082

[10] Salam, A., & El Hibaoui, A. (2018). Comparison of Machine Learning Algorithms for the Power Consumption Prediction: Case Study of Tetouan city. 2018 6th International Renewable and Sustainable Energy Conference (IRSEC). IEEE. https://doi.org/10.1109/IRSEC.2018.8702900

[11] https://www.kaggle.com/datasets/gmkeshav/tetuan-city-power-consumption

[12] Mu W, Cardelli R, Ferrari S. Data Preprocessing Techniques for Machine Learning Towards Improving Building Energy Performance: A Systematic Review. Energies. 2026; 19(6):1561. https://doi.org/10.3390/en19061561

[13] Neog, A., et al. (2025). Investigating a Model-Agnostic and Imputation-Free Approach for Irregularly-Sampled Multivariate Time-Series Modeling. arXiv preprint arXiv:2502.15785. https://arxiv.org/abs/2502.15785

[14] Lai, K.-H., Zha, D., Xu, J., Zhao, Y., Wang, G., & Hu, X. (2021). Revisiting time series outlier detection: Definitions and benchmarks. In Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 1). https://openreview.net/forum?id=r8IvOsnHchr

[15] Serras, J. L., Vinga, S., & Carvalho, A. M. (2021). Outlier Detection for Multivariate Time Series Using Dynamic Bayesian Networks. Applied Sciences, 11(4), 1955. https://doi.org/10.3390/app11041955

[16] Antony, E., Sreekanth, N. S., Sunil Kumar, R. K., & Nishanth, T. (2021). Data preprocessing techniques for handling time series data for environmental science studies. International Journal of Engineering Trends and Technology, 69(5), 196–207. https://doi.org/10.14445/22315381/IJETT-V69I5P228

[17] Liu, X., & Zhang, Q. (2024). Combining seasonal and trend decomposition using LOESS with a gated recurrent unit for climate time series forecasting. IEEE Access, 12, 85275–85290. https://doi.org/10.1109/ACCESS.2024.3415349

[18] Puggini, L., & McLoone, S. (2018). An enhanced variable selection and Isolation Forest based methodology for anomaly detection with OES data. Engineering Applications of Artificial Intelligence, 67, 126–135. https://doi.org/10.1016/j.engappai.2017.09.021

[19] Li, Y., Zheng, A., Tan, H., Chen, H., Cheng, S., Cai, H., & Wang, L. (2020). A new method for predicting the monthly fault number of watt-hour meters based on time series. Electric Power [Zhongguo Dianli], 53(6), 72–80. https://doi.org/10.11930/j.issn.1004- 9649.201910016

[20] Fan, X. (2023). Smoothing and imputation of longitudinal vehicle trajectory data [Doctoral dissertation, University of California, Irvine]. eScholarship. https://escholarship.org/uc/item/8c4333qs

[21] Lepot, M., Aubin, J.-B., & Clemens, F. H. L. R. (2017). Interpolation in Time Series: An Introductive Overview of Existing Methods, Their Performance Criteria and Uncertainty Assessment. Water, 9(10), 796. https://doi.org/10.3390/w9100796

[22] Pan, R., Yang, T., Cao, J., Lu, K., & Zhang, Z. (2015). Missing data imputation by K nearest neighbors based on grey relational analysis and genetic algorithm. Applied Soft Computing, 34, 329–340. https://doi.org/10.1016/j.asoc.2015.05.004

PISA-2022 өгөгдөлд суурилсан Монгол сурагчдын сурлагын амжилтыг таамаглах загвар ба сургалтын орчин дахь квази-туршилтын судалгаа

Цэрэнбалжирын Сэрчмаа
Боловсрол судлал
ШУТИС-ийн НБХүрээлэн
Монгол, Улаанбаатар
serchmaa@must.edu.mn

Эрдэнэбатын Батцэцэг
Мэдээллийн технологи
ШУТИС-ийн МХТС
Монгол, Улаанбаатар
https://orcid.org/0000-0003-3178-7355

Өлзийтийн Баярбат
Компьютерын ухаан
ШУТИС-ийн МХТС
Монгол, Улаанбаатар
bayarbat.u@must.edu.mn

***Хураангуй*— Сурагчдын сурлагын амжилтад нөлөөлөгч хүчин зүйлсийг шинжлэх нь боловсролын чанарыг сайжруулахад чухал ач холбогдолтой. Энэхүү судалгаанд PISA-2022 судалгааны Монголын өгөгдөлд тулгуурлан сурагчдын сурлагын амжилтад нөлөөлөх хүчин зүйлсийг машин сургалтын аргаар тодорхойлж, таамаглах загвар боловсруулан, улмаар гарсан үр дүнг бодит сургалтын орчинд туршиж үнэлэв. Судалгаа нь хоёр үе шаттай. Эхний шатанд, PISA-2022 Монголын өгөгдөлд олон хүчин зүйлийн шугаман регресс (OLS), LASSO, Random Forest алгоритмуудыг харьцуулан шинжлэхэд LASSO алгоритмын таамаглах чадвар хамгийн өндөр (байгалийн ухаан R^2=0.769, математик R^2=0.850) байв. Бүх загварын үр дүнгээр уншиж ойлгох чадвар нь байгалийн ухаан болон математикийн амжилтад хүчтэй нөлөөлөх хүчин зүйл болох нь тогтоогдов. Хоёрдугаар шатанд, загварын үр дүнд тулгуурлан боловсруулсан уншиж ойлгох чадварыг дэмжих сургалтын арга зүйг хяналтын бүлэггүй, туршилтын өмнөх–дараах хэмжилтэд суурилсан квази-туршилтын загварт тулгуурлан Дархан-Уул аймгийн 5-р ангийн нэг бүлэг (n=21), Сэлэнгэ аймгийн Мандал сумын 6-р ангийн хоёр бүлэгт (n=24, n=19) 3 сарын хугацаанд туршив. Туршилтын үр дүнгээс харахад туршилтын 3 бүлэгт сурлагын дундаж оноо статистикийн хувьд ач холбогдолтой өсөж (p<0.05), гүйцэтгэлийн ялгаа буурсан нь ахицын жигдрэлийг илтгэв. Хамаарлын шинжилгээгээр, уншиж ойлгох чадвар болон сурлагын амжилт хоорондын уялдаа туршилтын дараах үед бэхжсэн нь PISA загварын нотолгоо ба туршилтын үр дүнгийн уялдааг баталгаажуулав. Мөн уншиж ойлгох чадварыг дэмжих сургалтын арга зүй нь хичээл дамнасан байдлаар сурлагын амжилтыг сайжруулах эерэг нөлөө ажиглагдав. Хэдийгээр хяналтын бүлэггүй туршилтын загвар нь шалтгаан-үр дагаврын хамаарлыг шууд батлах хязгаарлалттай ч, эдгээр үр дүн нь өгөгдлийн шинжилгээний дэвшилтэт аргууд нь боловсролын салбар дахь далд хүчин зүйлсийг илрүүлэхэд өндөр ач холбогдолтойг харуулж байна. Цаашид сургалтын хөтөлбөрт уншиж ойлгох чадварыг дэмжсэн арга зүйг түлхүү нэвтрүүлэх нь сурагчдын сурлагын амжилтыг дээшлүүлэх боломжтойг энэхүү судалгаа харуулав.**

Түлхүүр үг— PISA-2022, машин сургалт, таамаглах загвар, LASSO, Random Forest, уншиж ойлгох чадвар, сурлагын амжилт, квази-туршилтын загвар

I. Удиртгал

Орчин үеийн мэдлэгт суурилсан нийгэмд боловсрол нь улс орны тогтвортой хөгжил, эдийн засгийн өсөлт, хүний нөөцийн чадавхыг тодорхойлогч гол хүчин зүйлсийн нэг. Иймээс боловсролын чанарыг сайжруулах, сурагчдын эзэмшиж буй мэдлэг, чадварыг бодитой үнэлэх нь боловсролын бодлого боловсруулагчид болон судлаачдын анхаарлын төвд оршсоор байна. Сурлагын амжилтыг зөв, найдвартай үнэлэх нь сурагчдын танин мэдэхүйн хөгжлийг тодорхойлохоос гадна сургалтын хөтөлбөр, арга зүй, боловсролын орчны үр нөлөөг шинжлэх ухааны үндэслэлтэйгээр тодорхойлох боломжийг бүрдүүлдэг [16].

Сүүлийн жилүүдэд боловсролын үнэлгээнд нотолгоонд суурилсан хандлага эрчимтэй нэвтэрч, олон улсын харьцуулалт хийх боломж бүхий томоохон судалгаанууд чухал үүрэг гүйцэтгэж байна. Тухайлбал, Programme for International Student Assessment (PISA) нь 15 настай сурагчдын байгалийн ухаан, математик, уншиж ойлгох чадварыг үнэлэх замаар улс орнуудын боловсролын тогтолцооны үр нөлөөг харьцуулан судалдаг. Энэхүү судалгаа нь сурагчдын амжилтыг хэмжихээс гадна тухайн амжилтад нөлөөлж буй хувь сурагчийн болон гэр бүлийн, сургуулийн орчны хүчин зүйлсийг өргөн хүрээнд судлах боломжийг олгодог бөгөөд боловсролын бодлогыг нотолгоонд суурилан боловсруулахад чухал мэдээллийн эх сурвалж болдог [9,10]. 2022 оны PISA судалгаанд Монгол улс анх удаа оролцсон бөгөөд судалгааны үр дүнгээс харахад манай улсын сурагчдын байгалийн ухаан, математик, уншиж ойлгох чадварууд олон улсын дунджаас харьцангуй доогуур түвшинд үнэлэгдсэн байна. Байгалийн ухааны чадвараар OECD-ын дунджаас 3 жил, Азийн дунджаас 1.8 жил хоцорсон; математикийн чадвараар OECD-ын дунджаас 2.3 жил, Азийн дунджаас 1.2 жил хоцорсон; уншиж ойлгох чадвараар OECD-ын дунджаас 5 жил, Азийн дунджаас 2.5 жил хоцорсон үзүүлэлттэй байна [10,11,12]. Энэ нь зөвхөн унших чадварын асуудал бус, байгалийн ухаан, математик зэрэг бусад хичээлийн амжилтад дам нөлөө үзүүлж болзошгүйг харуулж байна.

Монгол Улс PISA судалгаанд оролцсоноор сурагчдын сурлагын амжилтын олон улсын түвшин, түүний онцлог, хандлагыг тодорхойлох боломж бүрдсэн. Судалгааны үр дүнгээс харахад сурагчдын уншиж ойлгох чадвар нь математик болон байгалийн ухааны амжилттай нягт холбоотой байх хандлага ажиглагдаж байгаа нь суралцах суурь чадваруудын харилцан хамаарлыг илтгэнэ [1,5,6,7,13]. Гэсэн хэдий ч Монголын нөхцөлд сурагчдын сурлагын амжилтад нөлөөлж буй хүчин зүйлсийг цогцоор нь авч үзсэн, ялангуяа олон хувьсагчийн нөлөөллийг зэрэг үнэлэх, орчин үеийн статистик болон машин сургалтын

аргуудыг ашигласан судалгаа харьцангуй хомс байна. Мөн өгөгдлийн шинжилгээний үр дүнд суурилсан загварыг бодит сургалтын орчинд туршин, үр нөлөөг нь үнэлсэн судалгааны ажил бараг үгүй нь судалгааны чухал цоорхойг илэрхийлж байна. Иймд сурагчдын сурлагын амжилтыг зөвхөн тайлбарлах төдийгүй урьдчилан таамаглах, улмаар сургалтын практикт хэрэгжүүлэх боломжтой загвар боловсруулах нь боловсролын чанарыг сайжруулахад бодит хувь нэмэр оруулах ач холбогдолтой. Үүнээс гадна боловсруулсан загварыг сургалтын орчинд туршин хэрэгжүүлж, сурагчдын амжилт болон суралцах хандлагад үзүүлэх бодит нөлөөг үнэлэх нь судалгааны практик ач холбогдлыг нэмэгдүүлнэ.

Дээрх үндэслэлд тулгуурлан энэхүү судалгаа нь PISA-2022 өгөгдөлд суурилан Монгол улсын сурагчдын сурлагын амжилтад нөлөөлөх хүчин зүйлсийг тодорхойлох, тэдгээрт суурилсан таамаглах загвар боловсруулах, улмаар боловсруулсан загварыг бодит сургалтын орчинд туршин үр нөлөөг нь үнэлэхэд чиглэж байна. Цаашлаад нотолгоонд суурилсан боловсролын бодлого боловсруулах, сургалтын арга зүйг сайжруулах, сурагчдын сурлагын амжилтыг дээшлүүлэхэд онолын болон практик ач холбогдолтой хувь нэмэр оруулахад чиглэнэ.

II. Судалгааны зорилго

Энэхүү судалгааны зорилго нь PISA-2022 Монголын өгөгдөлд тулгуурлан сурагчдын байгалийн ухаан болон математикийн хичээлийн амжилтад нөлөөлж буй гол хүчин зүйлсийг тодорхойлж, тэдгээрийн нөлөөллийг машин сургалтын таамаглах загвараар үнэлэхийн зэрэгцээ боловсруулсан загварыг бодит сургалтын орчинд хэрэгжүүлэн үр дүнтэй эсэхийг туршилтаар шалгаж, бодит сургалтын үр нөлөөг хэмжихэд оршино.

Дэвшүүлсэн судалгааны асуулт:

- Сурагчийн асуулгын хуудаснаас хэмжигдсэн сурагчийн шинж чанар болон танин мэдэхүйн чадварууд нь байгалийн ухаан болон математикийн амжилтад хэрхэн нөлөөлж буйг урьдчилан таамаглахад чухал хувьсагчууд аль нь болох;
- Олон хүчин зүйлийн шугаман регресс, Random Forest, LASSO таамаглах загваруудын тайлбарлах чадварын түвшинг харьцуулах, энэ нь сурагчдын байгалийн ухаан болон математикийн амжилтыг хангалттай нарийвчлалтайгаар таамаглаж чадаж байгаа эсэх;
- Таамаглах загварт тулгуурлан бодит сургалтын орчинд туршилт хийхэд загварын аль хувьсагчуудыг арга зүйн нөлөөллийн үндэс болгох;
- Туршилтаар хэрэгжүүлсэн сургалтын арга зүй нь сурагчдын байгалийн ухаан болон математикийн амжилтад статистикийн хувьд ач холбогдолтой эерэг өөрчлөлт үзүүлж чадаж байгаа эсэх;
- Боловсруулсан загвар нь сурлагын амжилтыг таамаглахаас гадна сурагчдын суралцах хандлага, идэвх оролцоонд нөлөөлж байгаа эсэх;

III. Судалгааны арга зүй

A. Судалгааны ерөнхий загвар

Энэхүү судалгаа нь өгөгдлийн шинжилгээний аргуудыг сурган заах арга зүйн туршилттай хослуулсан холимог арга зүйн загвар, тэр дундаа тайлбарлах зорилготой дараалсан загварт хамаарна. Судалгааны зорилго нь шалтгаан-үр дагаврын хамаарлыг батлах бус, PISA загварын үр дүнд суурилсан арга зүйн нөлөөний боломжит чиглэлийг тандан судлах тул өмнөх-дараах хэмжилтэд тулгуурласан квази-туршилтын загвар зохистой гэж үзэв. Иймд илэрсэн эерэг өөрчлөлтийг туршилтын арга зүйтэй шалтгаан-үр дагаврын хамаарлаар холбохоос татгалзан, нөлөөний чиг хандлагыг тогтоосон эмпирик нотолгоо гэж тайлбарлана. Судалгааг дараах харилцан хамаарал бүхий хоёр үндсэн үе шаттайгаар хийв. Үүнд:

I шат: Таамаглах загварчлал ба хүчин зүйлсийн шинжилгээ

Энэ шатанд, PISA-2022 Монголын сурагчдын өгөгдөлд OLS, RF, LASSO регрессийн шинжилгээ хийж, сурагчдын байгалийн ухаан болон математикийн амжилтад нөлөөлж буй гол хүчин зүйлсийг статистик ач холбогдлын түвшинд эрэмбэлэн тогтоов. Энэхүү шатны үр дүн нь дараагийн шатны сургалтын арга зүйн нөлөөллийн онолын суурь үндэслэл болно.

II шат: Квази-туршилт

Эхний шатаас тодорхойлогдсон хүчтэй нөлөө бүхий хувьсагчуудыг сургалтын арга зүйд тусгаж, бодит сургалтын орчинд туршилтын өмнөх-дараах загварт тулгуурлан туршилтын 3 бүлэгт арга зүйн шинэчлэлийг хэрэгжүүлэв. Статистик шинжилгээний аргуудаар сурагчдын сурлагын амжилт, хандлага, идэвх оролцоонд гарсан өөрчлөлтийг шинжилсэн нь уг судалгааг онолын таамаглалаас хальж, практик нотолгоонд суурилсан үр дүнд хүргэх боломжийг олгосон болно.

B. Судалгааны хамрах хүрээ, түүвэр, өгөгдөл цуглуулах хэрэгсэл

Энэхүү судалгааны хамрах хүрээ болон түүврийг тоон шинжилгээний болон туршилт судалгааны гэсэн хоёр үндсэн үе шатанд хуваан авч үзсэн болно.

I шатны хамрах хүрээ ба түүвэр (PISA-2022 өгөгдөл)

Судалгааны эхний шатанд сурагчдын амжилтад нөлөөлөх хүчин зүйлсийг тодорхойлох зорилгоор PISA-2022 олон улсын үнэлгээний өгөгдлийг ашигласан.

- **Өгөгдлийн эх үүсвэр ба бүтэц:** Судалгаанд PISA-2022 үнэлгээнд хамрагдсан Монгол улсын 4 бүс, 12 давхаргыг төлөөлөх 196 сургуулийн 7,000 гаруй сурагчийн мэдээллийг хоёрдогч эх сурвалж болгон ашиглав. Анхны өгөгдлийн сан нь сурагчийн хувийн шинж чанар, гэр бүлийн нийгэм-эдийн засгийн нөхцөл, сургуулийн орчин, шалгалтын оноо зэрэг 1,278 хувьсагч бүхий олон улсын 613,732 сурагчийн мэдээллийг агуулж байв.
- **Түүврийн арга:** PISA судалгааны олон үе шаттай, давхаргат санамсаргүй түүврийн аргачлалыг баримтлав. Шинжилгээнд, тухайн улсын нийт 15 настай сурагчдын төлөөлөл болгох үүднээс PISA-аас санал болгодог жингийн коэффициент болон

давталтын жингийн коэффициентүүдийг харгалзан үзсэн [10].

- **Өгөгдөл бэлтгэл ба цэвэрлэгээ:** Судалгаанд ашиглах өгөгдлийг олон үе шаттайгаар цэвэрлэж, бэлтгэл ажлыг гүйцэтгэсэн бөгөөд цэвэрлэгээний үр дүнд 195 сургуулийн 4,383 сурагч (2,115 эрэгтэй, 2,268 эмэгтэй)-ийн 205 хувьсагч бүхий 100% бүрэн өгөгдлийн санг бүрдүүлэв. Түүвэрт хамрагдсан сурагчид нь 4 бүс (зүүн-говь, төв, хангай, баруун) болон хот, хөдөөгийн байршлыг хамарсан 12 давхаргад хуваарилагдсан болно.

II шатны хамрах хүрээ ба түүвэр (Туршилт судалгаа)

Загварын үр нөлөөг бодит орчинд шалгах зорилгоор хяналтын бүлэггүй, туршилтын өмнөх-дараах хэмжилтэд тулгуурласан квази-туршилтын загварыг ашигласан бөгөөд туршилт-судалгаанд хамрагдсан 3 бүлгийн түүврийн хэмжээ (n=21, n=24, n=19) нь тоон үзүүлэлтийн хувьд хязгаарлагдмал хэдий ч, боловсролын судалгаанд сургалтын хэвийн байдлыг хадгалах үүднээс анги бүлгийг бүхэлд нь хамруулах зарчмыг баримталсан болно. Cohen (1988)-ын тооцоогоор Paired t-test-д хүчтэй нөлөө ($d \geqslant 0.8$) илрэх тохиолдолд n=15 хангалттай гэж үздэг.

- **Түүврийн бүрэлдэхүүн:** Туршилт судалгаанд Дархан-Уул аймгийн нэг дунд сургуулийн 5-р ангийн 21 сурагч, Сэлэнгэ аймгийн Мандал сумын нэг дунд сургуулийн 6-р ангийн хоёр бүлгийн 43 сурагч, нийт 3 бүлгийн 64 сурагч хамрагдав. Туршилтыг тухайн сургуулийн багш нар өөрсдийн сургалтын үндсэн хөтөлбөрийн хүрээнд хэрэгжүүлсэн болно. Өгөгдлийн шинжилгээгээр хяналтын бүлэгтэй харьцуулах бус, тухайн бүлгийн дотоод ахиц, өөрчлөлтийн чиг хандлагыг тодорхойлоход чиглэсэн.
- **Хэмжилтийн хэрэгсэл ба хувьсагч:** Туршилтын өмнө (T0) болон явцад (T1) сурагчдын амжилтад нөлөөлөх хүчин зүйлсийг олон талт хэмжилтийн аргаар тодорхойлохын тулд дараах 4 ангиллын үзүүлэлтийг ашиглав. Үүнд:
  - **Амжилтын үзүүлэлт:** Өмнөх судалгаа болон I үе шатны таамаглах загварын үр дүнгээр уншиж ойлгох чадвар нь байгалийн ухаан болон математикийн амжилтад хүчтэй нөлөөтэй нь батлагдсан [13] тул сургалтын шууд үр дүнг байгалийн ухаан, математик, хэлний чадварын стандарт шалгалтын оноогоор хэмжив.
  - **Хандлагын үзүүлэлт:** Сурагчдын суралцах хандлагыг Суралцах урам зориг, стратегийг үнэлэх асуулга (Motivated Strategies for Learning Questionnaire-MSLQ)-д тулгуурлан үнэлэв [14]. Уг асуулгын 15 дэд хэсгийн нийт 81 асуултаас судалгааны зорилгод нийцэх 3 дэд хэсгийг сонгож, хэсэг бүрийн хувьд 4 асуулт (нийт 12 асуулт), 5 шатлалт үнэлгээтэйгээр (5–6)-р ангийн сурагчдад ойлгомжтой байдлаар хялбаршуулан боловсруулав. Асуулгын дотоод найдвартай байдлыг шалгахад туршилтын 3 бүлгийн хувьд бүх хичээл дээр Cronbach α утга шаардлага хангасан ($\alpha \geq 0.7$) болно.
  - **Идэвхийн үзүүлэлт:** Сурагчийн идэвхийг сурагчийн өөрийн үнэлгээний асуулга болон багшийн ажиглалт гэсэн хоёр эх сурвалжаар үнэлэв.
    - *Сурагчийн асуулга:* Суралцахуйн оролцооны гурван хэмжээст загвар (Three-Dimensional Model of Engagement) [4] -т тулгуурлан, зөвхөн “зан байдлын идэвх” дэд хэмжээсийн хүрээнд 5 асуулт, 5 шатлалт үнэлгээтэйгээр боловсруулав. Бага насны сурагчдад зориулсан, цөөн асуулттай, шинээр боловсруулсан асуулгын хувьд $\alpha \approx 0.50-0.60$ байхыг зөвшөөрөгдөх доод хязгаар гэж үздэг[3,8]. Асуулгын дотоод найдвартай байдлыг шалгах туршилтаар 3 бүлгийн хувьд бүх хичээл дээр шаардлага хангагдсан ($\alpha \geq 0.5$) болно.
    - *Багшийн ажиглалт:* Сурагчдын чадварын ахицыг чанарын гүнзгий ажиглалтаар дүгнэхийн тулд хичээл бүрийн онцлогт тохируулан чанарын түвшинг илэрхийлсэн 4 шатлалт ажиглалтын үнэлгээг ашиглав.
  - **Богино хугацаан дахь өөрчлөлтийн үзүүлэлт:** Туршилтын арга зүйн нөлөөг тогтвортой хэмжихийн тулд туршилтын хугацаанд давтамжтайгаар шалгалт авах (сурлагын амжилт), давтамжтай асуулга бөглүүлэх (хандлагын өөрчлөлт), болон асуулга, ажиглалтыг хослуулах (идэвхийн өөрчлөлт) хэлбэрээр динамик өөрчлөлтийг тодорхойлов.

C. Өгөгдлийн шинжилгээний үр дүн ба загварчлал

Энэ хэсэгт PISA-2022 өгөгдөлд хийсэн статистик шинжилгээ болон машин сургалтын загваруудын харьцуулсан үр дүнг тусгав.

Корреляцын шинжилгээ

PISA үнэлгээний үндсэн 5 бүлэг хүчин зүйл (сурлага; сургалтын хандлага, сэтгэлгээ; сурах орчин, дэмжлэг; гэр бүлийн нийгэм-эдийн засгийн байдал; цаг ашиглалт, амьдралын чанар)-ийн хүрээнд сурагчдын байгалийн ухаан болон математикийн гүйцэтгэлд нөлөөлөх хүчин зүйлсийг корреляцын шинжилгээгээр үнэлэхэд хоёр судлагдахууны хувьд нөлөөллийн зүй тогтол ижил байгаа нь тогтоогдов. Шинжилгээний үр дүнгээс үзэхэд, сурагчдын танин мэдэхүйн суурь чадварууд сургалтын амжилтад хамгийн хүчтэй эерэг хамааралтай байв. Тодруулбал, математик болон байгалийн ухааны гүйцэтгэл хоорондоо маш хүчтэй эерэг хамааралтай (r=0.845), уншиж ойлгох чадвар нь байгалийн ухаан (r=0.757) болон математикийн (r=0.769) амжилтад хүчтэй эерэг нөлөө үзүүлж байв. Харин нийгэм-эдийн засгийн үзүүлэлт болох гэр бүлийн хөрөнгө (HOMEPOS, r=0.403–0.428) болон нийгэм, соёлын байр суурь (ESCS, r=0.394–0.414) нь сургалтын амжилттай дунд зэргийн эерэг хамааралтай байсан бол, гэр бүлийн дэмжлэг, эцэг эхийн боловсролын түвшин зэрэг бусад хүчин зүйлс дундаас сул хамааралтай байв.

Таамаглах загваруудын харьцуулсан шинжилгээ

Сурагчдын сурлагын амжилтыг таамаглах зорилгоор OLS, RF, LASSO таамаглах загваруудыг туршиж, үр дүнг харьцуулан үнэлэв.

Байгалийн ухааны амжилтыг таамаглах загвар:

- **OLS:** Загварын тайлбарлах чадвар R^2=0.746 байв. Математикийн чадвар (β=0.579, p<0.001) болон уншиж ойлгох чадвар (β=0.255, p<0.001) нь статистик ач холбогдол бүхий хамгийн өндөр эерэг нөлөө үзүүлэв.
- **LASSO:** Анхны 205 хувьсагчаас хамгийн ач холбогдолтой 109 хувьсагчийг сонгосон ба загварын тайлбарлах чадвар R^2=0.769-д хүрсэн нь OLS загвараас илүү нарийвчлалтай байв.
- **Random Forest:** Загварын тайлбарлах чадвар R^2=0.745 байсан бөгөөд хүчин зүйлсийн ач холбогдлын эрэмбэ нь өмнөх загваруудтай ижил үр дүнг үзүүлэв.

Математикийн амжилтыг таамаглах загвар:

- **OLS:** Загварын тайлбарлах чадвар R^2=0.616 байсан ба уншиж ойлгох чадвар (β=0.787, p<0.001) гол тодорхойлогч хүчин зүйл байв. Эхийн боловсролын түвшин сурагчийн сурлагын амжилттай статистик ач холбогдол бүхий эерэг хамааралтай болох нь тогтоогдов.
- **LASSO болон Random Forest:** LASSO загвар нь 41 хувьсагчийг сонгож, тайлбарлах чадварыг R^2=0.850 хүртэл өсгөсөн бол Random Forest загварын тайлбарлах чадвар R^2=0.842 байв. Эдгээр загварууд нь уншиж ойлгох чадвараас гадна сурагчийн танин мэдэхүйн бусад чадварууд математикийн амжилтад тодорхой эерэг нөлөөтэйг илрүүлсэн.
- Таамаглах загваруудын харьцуулан шинжилгээгээр, LASSO загвар нь хувьсагчуудын олонлогт шүүлт хийж, загварын хэт нийцлээс сэргийлсэн тул хамгийн өндөр тайлбарлах чадвартай (R^2≈0.77−0.85) болох нь тогтоогдов.

D. Туршилт хэрэгжүүлэх үйл явц, арга зүй

Туршилтын арга зүй, хэрэгжүүлсэн нөхцөл

Туршилтын хугацаанд сурагчдын уншиж ойлгох чадварыг дэмжих сургалтын арга зүйг бодит сургалтын орчинд системтэйгээр хэрэгжүүлэв. Туршилтын хугацаа 3 сар үргэлжилсэн бөгөөд 5-р ангийн бүлгийн хувьд Монгол хэл, Математик, Хүн ба байгаль хичээлүүдийг, 6-р ангийн бүлгийн хувьд Монгол хэл, Математик, Хими, Мэдээлэл зүй хичээлүүдийг хамруулсан болно. Туршилтын хугацаанд сонгогдсон хичээлүүдийн сургалтын агуулга, шалгалтын бүтэц, үнэлгээний шалгуур өөрчлөгдөөгүй, зөвхөн уншиж ойлгох чадварыг дэмжих дараах нэмэлт арга зүйг тогтмол хэрэгжүүлсэн болно.

5-р ангийн бүлэгт:

- Өглөөний уншлага: Сурагчид өдөр бүр 10–15 минут ном уншиж, гол санаа, сонирхолтой хэсгийг тэмдэглэл хэлбэрээр бичиж тэмдэглэнэ.
- Дэмжих цагийн уншлага: Уншсан номын агуулгыг хичээлийн сэдэвтэй холбон асуулт–хариулт, хэлэлцүүлэг өрнүүлж, ойлгосноо баталгаажуулна.
- Номын тэмдэглэл: Номын үйл явдал, гол санааг тайлбарлах, харьцуулах, дүгнэх, зураг болон уран сэтгэмжээр дүрслэх/илэрхийлэх зэргээр тэмдэглэл хөтөлнө.

6-р ангийн 1-р бүлэгт (Монгол хэлний дэмжлэгтэй)

- Уншлага, хэлэлцүүлэг: Долоо хоног бүрийн хичээлийн нэг цагийн эхний 20 минутад сонгогдсон эхийг чангаар болон чимээгүй унших замаар бүх сурагчийг хамруулна. Сүүлийн 20 минутад эхийг задлан шинжилж, утга санаа, шинэ үг, хэллэгүүдийг тайлбарлан хэлэлцэнэ.

6-р ангийн 2-р бүлэгт (Англи хэлний дэмжлэгтэй)

- Уншлага, хэлэлцүүлэг: Хичээлийн бус цагаар Англи хэл дээрх (Lingors) цахим номын сангаас унших даалгавар өгч, хичээлийн нэг цагийг уншсан номын талаар хэлэлцэх, шинэ үг, хэллэгүүдийг тайлбарлан ярилцах байдлаар сурагчдын гүйцэтгэлийг хянана.

Туршилтын өгөгдөл болон шинжилгээний аргууд

Энэхүү туршилтын судалгаанд хяналтын бүлэг ашиглаагүй бөгөөд туршилтын өмнөх (T0) болон туршилтын явцын (T1) хэмжилтээр цуглуулсан өгөгдөлд тулгуурлан, туршилтын 3 бүлгийн дотоод ахиц, өөрчлөлтийн чиг хандлагыг тодорхойлов. Сурагчдын сурлагын амжилт, хандлага, идэвхийн өөрчлөлтийг үнэлэхэд ашигласан шинжилгээний аргуудыг Хүснэгт 1-д жагсаав.

TABLE I. СУДАЛГААНД АШИГЛАГДСАН ШИНЖИЛГЭЭНИЙ АРГУУД

Шинжилсэн зүйлс	**Хугацаа**
Сурлагын амжилтын өөрчлөлт (дундаж, стандарт хазайлт)	T0 → T1
Хувь сурагчийн түвшин дэх сурлагын амжилтын өөрчлөлт	T0 → T1
Сурлагын амжилтын өсөлтийн давхцал	T0 → T1
Чадварын түвшний өөрчлөлт	T0 → T1
Туршилтын арга зүйн нөлөө, статистикийн ач холбогдол	T0 → T1
Хандлага, идэвхийн өөрчлөлт	T0 → T1
Байгалийн ухаан болон математикийн амжилтад нөлөөлж буй хүчин зүйлсийн хамаарлын шинжилгээ	T1
Богино хугацаан дахь өөрчлөлтийн шинжилгээ	T0 → T1

E. Туршилтын үр дүнгийн шинжилгээ

Сурлагын амжилтын өөрчлөлт

Сурагчдын сурлагын амжилтын өөрчлөлтийг гүйцэтгэлийн дундаж болон стандарт хазайлтаар T0→T1 хугацаанд шинжилсэн үр дүнгээр:

Дархан-Уул аймгийн 5-р анги:

- Хүн ба байгаль хичээлийн хувьд сурагчдын дундаж оноо 64.90→76.05 болон өсөж, стандарт хазайлт 17.04→11.69 болж буурсан нь сурагчдын ахиц харьцангуй жигдэрснийг харуулав. Математикийн хувьд, дундаж оноо 62.64→76.10 болж хамгийн их өсөлт үзүүлсэн бөгөөд стандарт хазайлт 21.15→13.19 болж буурсан нь туршилтын арга зүй энэхүү хичээлд харьцангуй хүчтэй эерэг нөлөө үзүүлсэн байх магадлалтай. Монгол хэлний хувьд дундаж оноо 63.61→75.90 болж, стандарт хазайлт 28.50→12.19 хүртэл буурсан нь сул сурагчдын ахиц өндөр байсныг илтгэв.
- Дүгнэвэл, T0→T1 хугацаанд бүх хичээлийн хувьд сурагчдын дундаж гүйцэтгэл сайжирч, стандарт

хазайлт мэдэгдэхүйц буурсан зүй тогтол ажиглагдав. Туршилтын арга зүй нь сурагчдын сурлагын амжилтыг сайжруулахын зэрэгцээ гүйцэтгэлийн зөрүүг ойртуулж, ахицыг илүү жигд түвшинд хүргэхэд дэмжлэг үзүүлсэн байх боломжтойг харуулж байна.

Сэлэнгэ аймгийн Мандал сумын 6-р анги: Энэ сумын хувьд 6-р ангийн хоёр бүлгийг (6-1, 6-2) харьцуулан шинжилсэн болно.

- Химийн хичээлийн хувьд 6-1 бүлгийн сурагчдын дундаж оноо 59.62→87.79 болж өссөн бол 6-2 бүлгийн хувьд 61.05→86.73 болж өсөв. Энэхүү өсөлт нь туршилтын хугацаанд хоёр бүлгийн аль алинд нь сурагчдын сурлагын амжилт мэдэгдэхүйц ахисныг харуулж байна. Стандарт хазайлт буурсан хандлага ажиглагдсан нь сурагчдын амжилтын ялгаа багасч, гүйцэтгэл илүү жигдэрч байгааг илтгэнэ. Харьцуулбал 6-1 бүлгийн өсөлт (Δ=28.79) нь 6-2 бүлгээс (Δ=25.68) бага зэрэг өндөр байгаа нь ажиглагдав. Математикийн хичээлийн хувьд 6-1 бүлгийн сурагчдын дундаж оноо 71.5→ 82.16 болж өссөн бол 6-2 бүлгийн хувьд 70.73→77.73 болж өсөв. Туршилтын хугацаанд хоёр бүлгийн аль алинд нь сурагчдын сурлагын амжилтад эерэг өөрчлөлт ажиглагдав. Стандарт хазайлт 6-1 бүлгийн хувьд 21.65→13.30 болж буурсан бол 6-2 бүлгийн хувьд 21.78→12.87 болж буурав. Мэдээлэл зүй хичээлийн хувьд, 6-1 бүлгийн гүйцэтгэлийн дундаж 76.20→98.87 болж өссөн бол 6-2 бүлэгт 77.57→99.10 болж өсөв. Стандарт хазайлт буурсан нь сурагчдын гүйцэтгэл жигдэрч, ялгаа багасах хандлага илрэв. Монгол хэлний хувьд, 6-1 бүлгийн дундаж оноо 57.76→78.04 болж өссөн бол 6-2 бүлэгт 65.67→77.00 болж өсөв. Стандарт хазайлт мөн буурсан ажиглагдав.
- Дүгнэвэл, T0→T1 хугацаанд Мандал сумын туршилтад хамрагдсан хоёр бүлгийн хувьд судлагдсан хичээлүүдийн дундаж гүйцэтгэл мэдэгдэхүйц өсөж, гүйцэтгэлийн хэлбэлзэл ойртсон нь ажиглагдав. Энэ нь туршилтад хэрэгжүүлсэн сургалтын арга зүй сурагчдын сурлагын амжилтыг сайжруулж, гүйцэтгэлийг илүү жигд түвшинд хүргэхэд тодорхой хэмжээгээр эерэг нөлөө үзүүлсэн байх боломжтойг илтгэнэ.

Хувь сурагчийн түвшин дэх сурлагын амжилтын өөрчлөлт

Хувь сурагчийн түвшинд сурлагын амжилтын өөрчлөлтийг T0→T1 хугацаанд шинжилсэн үр дүн:

Дархан-Уул аймгийн 5-р анги: Шинжилгээний үр дүнгээс харахад, бүх хичээлийн хувьд эерэг чиг хандлага давамгайлав.

- Хүн ба байгаль хичээлийн хувьд 19 сурагчийн (≈90%) гүйцэтгэл сайжирч, 2 сурагчийн үзүүлэлт буурсан нь уг хичээл дээр хамгийн өндөр хувьтай эерэг өөрчлөлт илрэв. Математикийн хувьд 16 сурагчийн (≈76%) гүйцэтгэл сайжирч, 4 сурагчийнх буурч, нэг сурагчийн үзүүлэлт тогтвортой байсан нь эерэг өөрчлөлт давамгайлж байгаа ч хүн ба байгаль хичээлтэй харьцуулахад ахицын түвшин бага байгааг харуулав. Монгол хэлний хувьд 13 сурагчийн (≈62 хувь) гүйцэтгэл сайжирч, 8 сурагчийн үзүүлэлт буурсан нь эерэг өөрчлөлт ажиглагдсан ч хувь сурагчийн түвшинд хэлбэлзэл харьцангуй их байсныг илтгэж байна.

Сэлэнгэ аймгийн Мандал сумын 6-р анги: Шинжилгээний үр дүнгээс харахад, ихэнх хичээлд эерэг чиг хандлага давамгайлав.

- *6-1 бүлэг*: Бүх хичээл дээр эерэг өөрчлөлт давамгайлав. Монгол хэлний хувьд 23 сурагчийн (≈96%) гүйцэтгэл сайжирч, 1 сурагчийн үзүүлэлт буурав. Хими болон мэдээлэл зүйн хичээл дээр бүх сурагчийн (100%) гүйцэтгэл сайжирч, буурсан болон тогтвортой үзүүлэлт ажиглагдаагүй нь хамгийн өндөр эерэг өөрчлөлт илэрснийг харуулж байна. Харин математикийн хичээлд 16 сурагчийн (≈67%) гүйцэтгэл сайжирч, 4 сурагчийн үзүүлэлт буурч, 4 сурагчийн үзүүлэлт тогтвортой байсан нь эерэг өөрчлөлт давамгайлсан ч бусад хичээлтэй харьцуулахад ахицын түвшин харьцангуй бага байв.
- *6-2 бүлэг*: Ихэнх хичээлийн хувьд эерэг чиг хандлага давамгайлав. Хими болон мэдээлэл зүй хичээлийн хувьд бүх сурагчийн (100%) гүйцэтгэл сайжирч, буурсан болон тогтвортой үзүүлэлт ажиглагдаагүй. Монгол хэлний хичээлд 16 сурагчийн (≈84%) гүйцэтгэл сайжирч, 3 сурагчийн үзүүлэлт буурав. Харин математикийн хичээлд 11 сурагчийн (≈58%) гүйцэтгэл сайжирч, 8 сурагчийн үзүүлэлт буурсан нь бусад хичээлтэй харьцуулахад ахиц харьцангуй бага, хувь сурагчийн түвшинд хэлбэлзэл их байгааг харуулж байна.
- Ерөнхийд нь авч үзвэл, хоёр бүлгийн аль алинд нь хими болон мэдээлэл зүйн хичээл дээр хамгийн өндөр эерэг өөрчлөлт илэрсэн бол математикийн хичээлд ахиц харьцангуй бага, хэлбэлзэл илүү ажиглагдав.

Сурагчдын чадварын түвшний өөрчлөлтийн шинжилгээ

Сурагчдын сурлагын амжилтыг онооны интервалын дагуу тааруу (0–59), дунд (60–79), хангалттай (80–100) гэсэн гурван чадварын түвшинд ангилан, T0→T1 хугацаанд түвшний шилжилтийг шинжлэв.

Дархан-Уул аймгийн 5-р анги: T0→T1 шилжилтийн шинжилгээгээр бүх хичээлийн хувьд эерэг чиг хандлага ажиглагдав.

- Монгол хэлний хувьд, сурагчдын ихэнх нь түвшин хадгалж, дунд болон тааруу түвшний сурагчид түвшин ахисан нь ажиглагдав. Математикийн хувьд, тааруу болон дунд түвшний сурагчид ахисан боловч хангалттай түвшний сурагчдын тодорхой хувь нь дунд түвшинд шилжсэн нь тухайн хичээлийн онцлог болон даалгаврын түвшний нөлөө байж болох талтай. Хүн ба байгаль хичээлийн хувьд, дунд түвшний сурагчдын ахиц хамгийн өндөр хувьтай илэрч, хангалттай түвшний сурагчид бүрэн түвшин хадгалсан нь чадварын түвшний шилжилт харьцангуй тогтвортойг харуулав. Дүгнэвэл, туршилтын хугацаанд сурагчдын чадварын түвшний өөрчлөлт нь “тааруу→дунд”, “дунд→хангалттай” чиглэлд зонхилов.

Сэлэнгэ аймгийн Мандал сумын 6-р анги: T0→T1 хугацаанд 6-1 болон 6-2 бүлгийн сурагчдын гүйцэтгэлийн түвшний шилжилтийг хичээл тус бүрээр харьцуулан шинжилж үзэхэд хоёр бүлэгт нийтлэг эерэг чиг хандлага давамгайлсан ч ахицын түвшин, тогтвортой байдал, хэлбэлзлийн хувьд тодорхой ялгаа ажиглагдав.

- Монгол хэлний хувьд, 6-1 бүлэгт тааруу болон дунд түвшний сурагчдын ихэнх нь дунд болон хангалттай түвшинд шилжиж, нийт 16 сурагчийн түвшин сайжирч, буурсан тохиолдол гарсангүй. Харин 6-2 бүлгийн хувьд мөн эерэг өөрчлөлт ажиглагдсан ч 1 сурагчийн түвшин хангалттай түвшнээс дунд түвшинд буурав. Эндээс, монгол хэлний хичээлд хоёр бүлэгт нийтлэг ахиц илэрсэн ч 6-1 бүлгийн өсөлт илүү тогтвортой, буурсан тохиолдолгүй нь онцлог байв. Математикийн хувьд, 6-1 бүлгийн тааруу түвшний сурагчид бүгд дунд түвшинд, дунд түвшний сурагчдын тодорхой хэсэг нь хангалттай түвшинд шилжиж, түвшин буурсан тохиолдол гараагүй. Харин 6-2 бүлгийн хувьд ахиц ажиглагдсан ч хангалттай түвшний зарим сурагч дунд түшинд буурч, ахиц болон бууралтын аль аль нь илэрсэн, илүү хэлбэлзэлтэй бүтэцтэй байв. Иймээс математикийн хичээлд 6-1 бүлгийн ахиц илүү тогтвортой, харин 6-2 бүлэгт сурагчдын гүйцэтгэл жигд бус байгаа нь ажиглагдав. Хими болон мэдээлэл зүйн хичээлийн хувьд, аль аль бүлэгт тааруу болон дунд түвшний бүх сурагч хангалттай түвшинд шилжиж, хангалттай түвшний сурагчид түвшин хадгалав. Буурсан тохиолдол ажиглагдсангүй.
- Дүгнэвэл, туршилтын хугацаанд сурагчдын чадварын түвшний өөрчлөлт нь “тааруу→дунд”, “дунд→хангалттай” чиглэлд зонхилсон бөгөөд энэ нь сурагчдын сурлагын амжилт нийтлэг байдлаар ахисныг илтгэнэ. Гэсэн хэдий ч 6-1 бүлгийн хувьд ахиц илүү тогтвортой, бууралт ажиглагдаагүй, харин 6-2 бүлэгт ялангуяа математикийн хичээл дээр харьцангуй их хэлбэлзэл ажиглагдав. Хими болон мэдээлэл зүйн хичээлд хоёр бүлэгт адилхан өндөр өсөлт илэрсэн нь хэрэгжүүлсэн сургалтын арга зүй сурагчдын мэдлэг, ойлголтыг үр дүнтэй дэмжсэн байх магадлалтай.

Сурлагын амжилтын өсөлтийн давхцалын шинжилгээ

Дархан-Уул аймгийн 5-р анги: Сурагчдын сурлагын амжилтын өсөлтийн давхцалыг Зураг 1-д үзүүлэв.

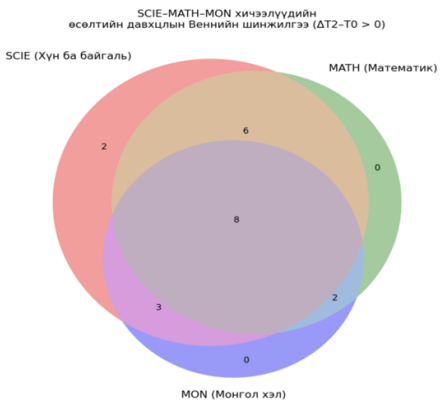


Fig. 1. 5-р ангийн сурагчдын сурлагын амжилтын өсөлтийн давхцал

Шинжилгээний үр дүнгээс харахад, 3 хичээлд зэрэг ахиц гаргасан сурагчид (n=8) нь нийт түүврийн хамгийн том хэсгийг эзэлж байгаа бөгөөд уншиж ойлгох чадварыг дэмжих арга зүй нь сурагчдын танин мэдэхүйд цогц байдлаар, хичээл дамнасан эерэг нөлөө үзүүлэх боломжтойг харуулж байна. Хүн ба байгаль болон математикийн хичээлд зэрэг ахиц үзүүлсэн сурагчдын (n=6) үр дүнгээс харахад, логик сэтгэлгээ, математик чадвар болон байгалийн ухааны ойлголтуудыг холбон хөгжүүлэхэд тодорхой хэмжээний дэмжлэг болсон байх магадлалтай. Хүн ба байгаль болон монгол хэлний хичээлд зэрэг ахиц гаргасан сурагчдын (n=3) үр дүнгээс харахад, хэл шинжлэлд суурилсан уншиж ойлгох чадвар нь байгалийн ухааны агуулгыг ойлгон суралцахад эергээр нөлөөлж болохыг харуулж байна. Математик болон монгол хэлний хичээлд зэрэг ахиц үзүүлсэн сурагчдын (n=2) үр дүн нь уншиж ойлгох чадвар болон математикийн гүйцэтгэлийн хооронд эерэг уялдаа байж болох таамаглалыг дэмжиж байна. Зөвхөн нэг хичээлд ахиц гаргасан сурагчдын (хүн ба байгаль n=2) үр дүнгээс харахад туршилтын арга зүйн нөлөө тухайн хичээлийн агуулгын онцлогтой түлхүү нийцсэн байж болох талтай. Дүгнэвэл, туршилтын арга зүй нь хэдийгээр бага түүвэрт хэрэгжсэн боловч сурагчдын хичээл дамнасан ерөнхий чадвар болон тухайлсан хичээлийн онцлогт нийцсэн эерэг өөрчлөлтүүдийг авчрах чиг хандлагатай байгааг харуулсан үр дүн болж байна.

Сэлэнгэ аймгийн Мандал сумын 6-р анги: Сурагчдын сурлагын амжилтын өсөлтийн давхцалыг Зураг 2-т үзүүлэв.

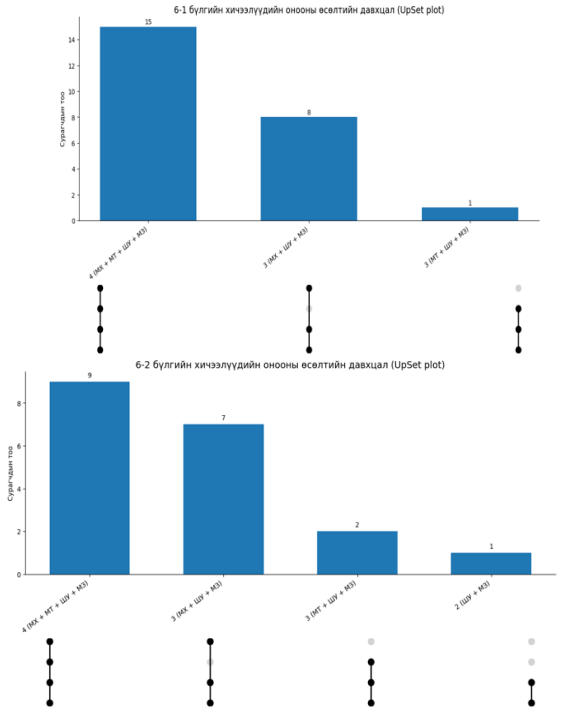


Fig. 2. 6-р ангийн сурагчдын сурлагын амжилтын өсөлтийн давхцал

Үр дүнгээс харахад, хоёр бүлгийн аль алинд нь олон хичээлд зэрэг ахиц илрэх эерэг хандлага ажиглагдсан ч ахицын тархалт, хамрах хүрээний хувьд тодорхой ялгаа илэрч байна. 4 хичээлд зэрэг ахиц гаргасан сурагчдын тоо 6-1 бүлэгт 15, 6-2 бүлэгт 9 байв. Эндээс, 6-1 бүлгийн хувьд, сурагчдын ахиц нь хичээл дамнасан цогц байдлаар

илрэв. 3 хичээлд зэрэг ахиц гаргасан сурагчдын хувьд 6-1 бүлэгт (8+1), 6-2 бүлэгт (7+2) сурагч байв. Хоёр хичээлд ахиц гаргасан тохиолдол 6-2 бүлгийн 1 сурагчид ажиглагдав. Дүгнэвэл, хоёр бүлгийн аль алинд сурагчдын сурлагын амжилтын өсөлт нь хичээл дамнасан байдлаар ажиглагдсан ч 6-1 бүлгийн хувьд 4 хичээлд зэрэг ахиц гаргасан сурагчдын тоо олон байгаа нь сурлагын амжилтын өсөлт илүү өргөн хүрээтэй, цогц байдлаар илэрснийг харуулж байна. Харин 6-2 бүлгийн хувьд ахиц мөн ажиглагдсан ч зарим тохиолдолд тодорхой хичээлүүдийн хослолд төвлөрөх шинж илрэв. Ялангуяа хими болон мэдээлэл зүйн хичээл нь бусад хичээлтэй хамт ахиц илрэх давтамж өндөр байгаа нь эдгээр хичээлийн суралцах үйл явцад уншиж ойлгох чадварын нөлөө тодорхой хэмжээнд илэрч байж болохыг харуулж байна. Иймээс хэрэгжүүлсэн сургалтын арга зүй нь сурагчдын суралцах амжилтад хичээл дамнасан байдлаар эерэг нөлөө үзүүлэх магадлалтайг харуулав.

Туршилтын арга зүйн нөлөө, статистикийн ач холбогдол

Туршилтын арга зүйн нөлөө, статистикийн ач холбогдлыг шинжлэхийн тулд хичээл тус бүрийн хувьд туршилтын өмнөх болон дараах шалгалтын оноонд Paired T-test болон Cohen's D-ын нөлөөллийг тус тус шинжлэв.

Дархан-Уул аймгийн 5-р анги: T0→T1 хугацаанд бүх хичээлийн хувьд сурагчийн дундаж оноо статистикийн хувьд ач холбогдолтойгоор өссөн нь ажиглагдав ($p<0.05$). Нөлөөний хэмжээгээр авч үзвэл хүн ба байгаль ($d=0.994$) болон математик ($d=0.894$)-т хүчтэй нөлөө илэрсэн бол монгол хэлний хувьд ($d=0.600$) дунд–хүчтэй нөлөө илэрсэн нь туршилтын арга зүй нь бодит практик ач холбогдолтойг харуулж байх талтай.

Сэлэнгэ аймгийн Мандал сумын 6-р анги: T0→T1 хугацаанд туршилтад хамрагдсан хоёр бүлгийн хувьд сурагчдын дундаж оноо ихэнх хичээлд статистикийн хувьд ач холбогдолтойгоор өссөн нь ажиглагдав. Нөлөөний хэмжээгээр хими болон мэдээлэл зүйн хичээлд маш хүчтэй нөлөө илрэв. Химийн хичээлийн хувьд 6-1 бүлэгт $d=2.56$, 6-2 бүлэгт $d=2.31$; мэдээлэл зүйн хувьд 6-1 бүлэгт $d=3.56$, 6-2 бүлэгт $d=3.82$ маш хүчтэй нөлөө ажиглагдсан нь түүврийн хэмжээ бага, шалгалтын дээд хязгаарт хүрэх нөлөөтэй холбоотой байж болзошгүй тул энэ утгыг болгоомжтой тайлбарлах шаардлагатай. 6-1 бүлгийн хувьд монгол хэлний хичээлд $d=1.48$, математикт $d=0.82$ гэсэн хүчтэй болон дунд–хүчтэй нөлөө илрэв. Харин 6-2 бүлгийн математикийн хичээлд статистикийн ач холбогдолгүй ($p = 0.111$) байв. Дүгнэвэл, туршилтын арга зүй нь хоёр бүлгийн сурагчдад статистикийн ач холбогдолтой эерэг өөрчлөлт үзүүлсэн ч харьцуулсан шинжилгээгээр 6-1 бүлэгт үзүүлэх нөлөө илүү хүчтэй, тогтвортой байгаа нь ажиглагдав. Уг үр дүн нь сурагчдын уншиж ойлгох суурь чадварыг гадаад хэлний дэмжлэгтэйгээр хөгжүүлэхээс илүүтэй эх хэлний дэмжлэгт суурилсан сургалтын арга зүйг хэрэгжүүлэх нь танин мэдэхүйн хувьд илүү үр өгөөжтэй байх магадлалтайг харуулж байна.

Суралцах хандлага, идэвхийн өөрчлөлтийн шинжилгээ

Туршилтын өмнөх ба дараах (T0 → T1) хэмжилтээр сурагчдын суралцах хандлага, идэвх оролцоонд гарсан өөрчлөлтийг асуулга болон багшийн ажиглалтын үнэлгээгээр шинжилсэн үр дүнг дор нэгтгэв.

- *Дархан-Уул аймгийн 5-р анги:* Сурагчдын суралцах хандлага болон идэвхийн дундаж үзүүлэлт бүх хичээл дээр өссөн зүй тогтол ажиглагдав. Стандарт хазайлтын утга бага түвшинд хадгалагдсан нь энэхүү өсөлт нийтийг хамарсан бөгөөд тогтвортой шинжтэй болохыг илтгэв. Багшийн ажиглалтаар сурагчдын асуулт асуух, санал бодлоо тайлбарлах, багаар ажиллах зэрэг бодит үйлдлүүд тогтмол сайжирсан нь туршилтын арга зүй нь танин мэдэхүйн сонирхлыг өдөөж, анхаарал төвлөрлийг нэмэгдүүлэхэд ач холбогдолтой нөлөө үзүүлсэн байх талтай.
- *Сэлэнгэ аймгийн Мандал сумын 6-р анги:*
  - 6-1 бүлэг: Монгол хэл, хими, мэдээлэл зүйн хичээлд хандах хандлага, идэвх оролцоо жигд өссөн ажиглагдав. Математикийн хичээлд хандах хандлага ялимгүй буурсан (-0.07) ч суралцах идэвх (+0.06) нэмэгдсэн нь оролцоо хадгалагдаж буйг харуулж байна.
  - 6-2 бүлэг: Суралцах хандлагын өөрчлөлт монгол хэл, хими, мэдээлэл зүйн хичээлд өссөн ажиглагдав. Математикийн хичээлийн хандлага ялимгүй (-0.06) буурсан ч идэвх (+0.03) өссөн үзүүлэлттэй байв. Хими (-0.05), болон мэдээлэл зүйн (-0.03) идэвх бага зэрэг буурсан нь хувь сурагчийн түвшинд хичээлийн идэвх ялгаатай байж болох юм.
  - Дээрх хэмжилтээс үзэхэд, T0→T1 хугацаанд хоёр бүлгийн сурагчдын суралцах хандлага, идэвхийн дундаж үзүүлэлтүүд ерөнхийдөө өссөн нь туршилтаар хэрэгжүүлсэн уншиж ойлгох чадварыг дэмжих арга зүй нь сурагчдын суралцах сонирхол, идэвхтэй оролцоог нэмэгдүүлэхэд дэмжлэг болж байх магадлалтай.

Хамаарлын шинжилгээ

Дархан-Уул аймгийн 5-р анги:

- Хүн ба байгаль хичээлийн хувьд, T1 үеийн хэмжилтээр тус хичээлийн амжилт нь математикийн шалгалтын оноо ($r\approx0.866$), монгол хэлний шалгалтын оноо ($r\approx0.848$), хүн ба байгаль хичээлийн идэвх ($r\approx0.793$), мөн багшийн ажиглалтаар үнэлэгдсэн бүх хичээлийн үзүүлэлтүүд ($r\approx0.706$–0.757)-тэй хүчтэй эерэг хамааралтай гарав. Үүнээс гадна монгол хэл, хүн ба байгаль, математикийн хичээлд хандах хандлага болон суралцах идэвх нь дундаас дээш хүчтэй эерэг хамаарал ($r\approx0.542$–0.666) үзүүлсэн нь сурагчдын сэтгэл зүй, хандлагын хүчин зүйлс танин мэдэхүйн гүйцэтгэлтэй нягт уялдаатай байж болохыг харуулж байна.
- Математикийн хувьд, T1 үед энэ хичээлийн амжилт нь монгол хэл ($r\approx0.879$) болон хүн ба байгаль ($r\approx0.866$) хичээлийн шалгалтын оноо, хүн ба байгаль хичээлийн идэвх ($r\approx0.857$), багшийн ажиглалтын үзүүлэлтүүд ($r\approx0.760$–0.827)-ээс хүчтэй эерэг хамааралтай болох нь ажиглагдлав. Сонирхолтой нь, математикийн хичээлд хандах шууд хандлага, идэвхээс (сул-дунд хамааралтай) илүүтэйгээр, монгол хэл болон хүн ба байгаль хичээлд хандах хандлага нь математикийн

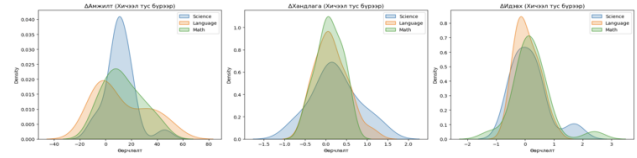
амжилттай илүү хүчтэй эерэг хамаарал (r≈0.678–0.697) үзүүлэв. Энэ нь уншиж ойлгох чадварыг дэмжих туршилтын арга зүй нь хичээл дамнасан эерэг хамаарлыг үүсгэж, математикийн өгөгдөлтэй ажиллах чадварыг давхар дэмжсэн байх талтай.

Сэлэнгэ аймгийн Мандал сумын 6-р анги:

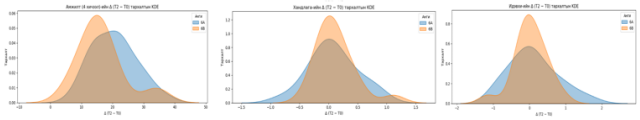
- Химийн хичээлийн хувьд, 6-1 бүлэгт хийсэн T1 үеийн хамаарлын шинжилгээний үр дүнгээс харахад, химийн хичээлийн амжилт нь математик (r≈0.445) болон монгол хэлний чадвар (r≈0.424)-аас дунд, мэдээлэл зүйн шалгалтын оноо, сурагчдын суралцах хандлага, идэвх (r≈0.05–0.3) болон бусад үзүүлэлтээс сул хамаарал үзүүлсэн байна. Харин 6-2 бүлгийн үр дүнгээс харахад, энэ хичээлийн амжилт нь уншиж ойлгох чадвар (r≈0.675), математикийн чадвар (r≈0.507)-аас хүчтэй, бусад үзүүлэлтүүдээс дунд болон сул хамааралтай байв.
- Математикийн хичээлийн хувьд, 6-1 бүлэгт хийсэн хамаарлын шинжилгээгээр монгол хэлний чадвар (r≈0.505)-аас хүчтэй хамааралтай, бусад үзүүлэлтээс дунд болон сул хамаарал үзүүлэв. Харин 6-2 бүлгийн үр дүнгээс харахад, энэ хичээлийн амжилт нь монгол хэлний чадвараас (r≈0.842) маш хүчтэй, бусад үзүүлэлтүүдээс дунд болон хүчтэй хамаарал үзүүлэв. Энэ нь математикийн чадварт уншиж ойлгох чадвар, ялангуяа даалгаврын агуулгыг ойлгох хэлний чадвар чухал үүрэгтэй байж болохыг харуулж байна.

Богино хугацаан дахь өөрчлөлтийн шинжилгээ

Богино хугацаан дахь өөрчлөлтийн чиг хандлагыг өсөлт/өөрчлөлтийн үнэлгээний шинжилгээний үр дүнд тулгуурлан Kernel density estimate (KDE) графикаар дүрслэн харуулав (3-р зураг).



а. Дархан-Уул аймгийн 5-р анги



б. Сэлэнгэ аймгийн Мандал сумын 6-р анги

Fig. 3. Өсөлт/өөрчлөлтийн үнэлгээний шинжилгээ

Дархан-Уул аймгийн 5-р анги: Сурлагын амжилтын өөрчлөлтийн тархалтаас харахад хүн ба байгаль болон математикийн хичээлд эерэг өсөлт давамгайлсан бол монгол хэлний хувьд тархалт бага зэрэг сөрөг чиглэлд шилжсэн байв. Суралцах хандлагын өөрчлөлт ихэнх сурагчдын хувьд тэг орчим болон эерэг утгын мужид төвлөрч, сайжирсан байдал ажиглагдав. Суралцах идэвхийн тархалт харьцангуй тогтвортой байсан ч зарим сурагчдын хувьд эерэг болон сөрөг өөрчлөлт зэрэгцэн илэрсэн. Дүгнэвэл, ихэнх сурагчдын хувьд өсөх хандлага давамгайлсан ч хувь сурагчийн түвшинд ялгаатай өөрчлөлт байгааг тодотгож байна.

Сэлэнгэ аймгийн Мандал сумын 6-р анги: Сурлагын амжилтын өөрчлөлтийн тархалтаас харахад 6-1 болон 6-2 бүлгийн аль алинд нь эерэг өсөлт давамгайлсан ч 6-1 бүлгийн ахиц нь 6-2 бүлэгтэй харьцуулахад илүү байв. Суралцах идэвх болон хандлагын өөрчлөлтийн тархалтыг ажиглахад ихэнх сурагчдын үзүүлэлт тэг орчим болон бага зэрэг эерэг утгын мужид төвлөрч, сайжирсан байдал ажиглагдав. Гэсэн ч зарим сурагчийн хувьд эерэг, сөрөг өөрчлөлтүүд зэрэгцэн илэрч байгаа нь суралцах хандлага, идэвхийн өсөлт хувь сурагчийн түвшинд харилцан адилгүй байгааг харуулав. Дүгнэвэл, 6-1 бүлгийн сурагчдын сурлагын амжилт болон хандлага, идэвх оролцоо нь 6-2 бүлгээс илүү хүчтэй, эерэг чиглэлд өөрчлөгдсөн ажиглагдав. 6-2 бүлгийн хувьд сурлагын амжилт нь тодорхой хэмжээнд өссөн ч хандлага болон идэвх нь илүү тогтвортой шинжтэй байв.

IV. Хэлэлцүүлэг

Энэхүү судалгааны үр дүнгээс үзэхэд уншиж ойлгох чадварыг дэмжих арга зүй нь сурагчдын сурлагын амжилтад зөвхөн хэл шинжлэлийн хүрээнд бус, STEM чиглэлийн хичээлүүдэд илүү хүчтэй нөлөө үзүүлж буй нь ажиглагдав. Ялангуяа математик болон байгалийн ухааны хичээлийн амжилт нь монгол хэлний чадвартай өндөр хамааралтай (r≈0.84–0.88) байгаа нь өгүүлбэртэй бодлого болон шинжлэх ухааны ухагдахууныг ойлгоход текстгэй ажиллах чадвар нэн чухал үүрэгтэйг харуулж байна. Энэ нь олон улсын судлаачдын үр дүнтэй нийцэж байна [1,5,7,13]. Тус судлаачид PISA болон TIMSS өгөгдөлд тулгуурлан уншиж ойлгох чадвар нь байгалийн ухаан болон математикийн амжилтыг урьдчилан таамаглахад хүчтэй нөлөөтэй болохыг баталсан. Мөн 6-1 (монгол хэл) болон 6-2 (гадаад хэл) бүлгийн харьцуулсан үр дүнгээс харахад монгол хэлний дэмжлэгтэй бүлгийн ахиц илүү тогтвортой байгаа нь сурагчид өөрийн эх хэл дээр мэдээллийг гүнзгий боловсруулснаар танин мэдэхүйн үйл явц илүү үр дүнтэй явагддаг болохыг харуулж байна. Эх хэлний дэмжлэгт суурилсан 6-1 бүлгийн үр дүн илүү тогтвортой байгаа нь Cummins-ын харилцан хамаарлын таамаглалтай нийцэж байна[2]. Тус онолын дагуу эх хэлний гүн боловсруулалт нь дараагийн хэлний академик чадварыг дэмждэг гэж үздэг. Хэдийгээр туршилтыг харьцангуй богино хугацаанд (3 сар) хэрэгжүүлсэн боловч статистикийн ач холбогдолтой нөлөө илэрсэн нь туршилтын арга зүйг ерөнхий боловсролын сургуулийн сургалтын хөтөлбөрт нэвтрүүлэх бодит боломж байгааг харуулав. Гэвч энэхүү судалгаа нь хяналтын бүлэггүйгээр, зөвхөн гурван туршилтын бүлгийн хүрээнд туршилтын хугацаанд цуглуулсан өгөгдөлд тулгуурласан тул илэрсэн эерэг өөрчлөлтүүд болон туршилтын арга зүйн хооронд шалтгаан-үр дагаврын шууд хамаарал оршин байна гэж батлахад статистикийн хувьд хязгаарлагдмал бөгөөд үүнд нэн болгоомжтой хандах нь зүйтэй.

V. Дүгнэлт

Энэхүү судалгаанд PISA-2022 судалгааны Монгол улсын өгөгдөлд тулгуурлан сурагчдын байгалийн ухаан болон математикийн амжилтад нөлөөлөх хүчин зүйлсийг таамаглах загвараар тодорхойлж, уг үр дүнд суурилсан сургалтын арга зүйг бодит сургалтын орчинд туршиж, үнэлэн судалгааны үр дүнг дараах хоёр хэсгээр дүгнэв.

I. PISA-2022 өгөгдөлд хийсэн таамаглах загварын шинжилгээний дүгнэлт

1. PISA-2022 судалгааны Монгол улсын өгөгдөлд хийсэн шинжилгээний үр дүнгээр сурагчдын байгалийн ухаан болон математикийн амжилтад уншиж ойлгох чадвар болон танин мэдэхүйн чадварууд хамгийн хүчтэй нөлөөтэй хүчин зүйлс болох нь тогтоогдов.
2. Олон хүчин зүйлийн шугаман регресс, Random Forest, LASSO таамаглах загваруудын харьцуулсан шинжилгээгээр, LASSO загвар нь илүү өндөр тайлбарлах чадвартай байж, сурагчдын амжилтад нөлөөлөх чухал хүчин зүйлсийг илүү оновчтой тодорхойлов.
3. Таамаглах загваруудын нийтлэг үр дүнгээр уншиж ойлгох чадвар нь хэлний чадвар төдийгүй математик болон байгалийн ухааны суралцах үйл явцад чухал үүрэгтэй суурь чадвар болох нь батлагдав.
4. Иймээс боловсролын өгөгдөлд суурилсан шинжилгээний үр дүн нь сурагчдын амжилтад нөлөөлөх хүчин зүйлсийг тодорхойлохоос гадна сургалтын арга зүйг нотолгоонд суурилан боловсруулах боломжийг бүрдүүлж байна.

II. Бодит сургалтын орчинд хэрэгжүүлсэн туршилт судалгааны дүгнэлт

Сурагчдын уншиж ойлгох чадварыг дэмжих сургалтын арга зүйг бодит орчинд хэрэгжүүлсэн туршилт судалгааны үр дүнгээс дараах дүгнэлтүүдийг нэгтгэн гаргав. Энэхүү дүгнэлт нь зөвхөн туршилтын бүлгээс туршилтын хугацаанд цуглуулсан өгөгдөлд тулгуурласан болно. Үүнд:

1. Уншиж ойлгох чадварыг дэмжих сургалтын арга зүйг 5 болон 6-р ангийн сурагчдын сургалтын орчинд хэрэгжүүлэхэд сурагчдын сурлагын амжилт ихэнх хичээлд мэдэгдэхүйц сайжирсан бөгөөд дундаж оноо өсөж, гүйцэтгэлийн ялгаа буурах хандлага ажиглагдав.
2. 5-р ангийн туршилтын бүлгийн хувьд монгол хэл, математик, хүн ба байгаль хичээлүүдийн сурлагын амжилт өсөж, сурагчдын чадварын түвшний шилжилт ихэвчлэн “тааруу→дунд”, “дунд→хангалттай” чиглэлд илэрсэн нь бага ангийн сурагчдын суурь танин мэдэхүйн чадварыг хөгжүүлэхэд уншиж ойлгох чадвар чухал нөлөөтэй байж болохыг харуулж байна.
3. 6-р ангийн туршилтын бүлгүүдийн хувьд монгол хэл, математик, хими, мэдээлэл зүй хичээлүүдийн амжилтад эерэг өөрчлөлт ажиглагдсан бөгөөд ялангуяа хими болон мэдээлэл зүйн хичээлд сурлагын амжилтын өсөлт өндөр илрэв. Энэ нь уншиж ойлгох чадвар нь агуулгын ойлголт, мэдээлэл боловсруулах чадварт суурилсан хичээлүүдтэй нягт холбоотой байж болохыг харуулж байна.
4. Сурагчдын сурлагын амжилтын өсөлтийн давхцалын шинжилгээгээр олон хичээлд зэрэг ахиц илрэх хандлага 5 болон 6-р ангийн аль алинд ажиглагдсан нь уншиж ойлгох чадвар нь хичээл дамнасан байдлаар суралцах ерөнхий чадварт нөлөөлж болохыг илтгэж байна.
5. Мөн сурагчдын суралцах хандлага болон идэвхийн шинжилгээгээр туршилтын хугацаанд 5 болон 6-р ангийн аль алинд эерэг өөрчлөлт илэрч, сурагчдын хичээлд оролцох идэвх, сонирхол нэмэгдэх хандлага ажиглагдав.
6. Ерөнхийд нь дүгнэвэл, PISA-2022 өгөгдөлд хийсэн таамаглах загварын шинжилгээний үр дүн болон бодит сургалтын орчинд хэрэгжүүлсэн туршилт судалгааны үр дүн хоорондоо уялдаатай байж, уншиж ойлгох чадвар нь сурагчдын танин мэдэхүйн чадвар болон хичээл дамнасан суралцах чадварыг дэмжих чухал хүчин зүйл байж болохыг харуулав. Иймээс уншиж ойлгох чадварыг дэмжих сургалтын арга зүйг сургалтын үйл ажиллагаанд системтэй хэрэгжүүлэх нь сурагчдын сурлагын амжилтыг сайжруулахад тодорхой хэмжээгээр эерэг нөлөө үзүүлэх магадлалтай. Гэсэн хэдий ч түүврийн хэмжээ харьцангуй бага тул энэхүү үр дүнг макро түвшинд буюу нийт сурагчдын төлөөлөл болгон өргөн хүрээнд ерөнхийлөн дүгнэх боломжгүй. Тиймээс энэхүү судалгааг тухайн арга зүйн сургалтын үйл явцад үзүүлэх боломжит эерэг нөлөөг тандан судалсан эмпирик нотолгоо хэмээн үзэх нь илүү зохистой бөгөөд цаашид хяналтын бүлэгтэй, түүврийн хэмжээг өргөтгөсөн харьцуулсан судалгаагаар дахин баталгаажуулах зайлшгүй шаардлагатай.

Ном зүй

[1] Caponera E., Sestito P., & Russo P.M. (2016). *The influence of reading literacy on mathematics and science achievement*. The Journal of Educational Research, 109(2), 197-204. https://doi.org/10.1080/00220671.2014.936998

[2] Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. Working Papers on Bilingualism, 19, 121–129.

[3] DeVellis, R. F. (2017). *Scale development: Theory and applications (4th ed.)*. Thousand Oaks, CA: Sage Publications

[4] Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). *School engagement: Potential of the concept, state of the evidence*. Review of Educational Research, 74(1), 59–109.

[5] Hanushek E. & Woessmann L. (2015). *Universal Basic Skills: What Countries Stand to Gain*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264234833-en

[6] Koyuncu İ. & Fırat T. (2020). *Investigating reading literacy in PISA 2018 assessment*. International Electronic Journal of Elementary Education, 13(2), 263-275. DOI: 10.26822/iejee.2021.189 https://iejee.com/index.php/IEJEE/article/view/1137/499

[7] Larwin K.H. (2010). *Reading is fundamental in predicting math achievement in 10th graders?* International Electronic Journal of Mathematics Education, 5(3), 131-145. https://doi.org/10.29333/iejme/254

[8] Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.). New York, NY: McGraw-Hill.

[9] OECD (2015). PISA Data Results: Excellence and Equity in Education, https://www.oecd.org/en/publications/pisa-2015-results-volume-i_9789264266490-en.html

[10] OECD (2022). PISA Data Results. Learning During and From Disruption (2022),

[11] https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/12/pisa-2022-results-volume-ii_222a5ef6/a97db61c-en.pdf

[12] OECD (2023). PISA 2022 Technical Report. OECD Publishing. https://www.oecd.org/pisa/data/pisa2022technicalreport/

[13] OECD Education GPS (2022). Mongolian Students Performance, https://gpseducation.oecd.org

[14] Ömer Cem Karacaoğlu & Yusuf Kasap (2023). *The Effect of Reading Comprehension Skills on Mathematics and Science According PISA Data*. https://doi.org/10.24331/ijere.1246885

[15] Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. Ann Arbor: University of Michigan, National Center for Research to Improve Postsecondary Teaching and Learning.

[16] Yergeshova Nafisa Shuxrat qizi (2025). *The Importance of education in today's world*, PEDAGOGS, 84(1), 308-309. https://scientific-jl.com/ped/article/view/23861

САЛБАР ХУРАЛДААН - VI

Дохио, медиа боловсруулалт, автоматжуулалт

№	**УБХИС-ийн хичээлийн I байр, 201 тоот, Эрдэм танхим** *Хуралдаан даргалагч: А.Одгэрэл /ШУТИС/, С.Баттуяа /ХААИС/*
1	“Дөрвөн сэнст төхөөрөмжийн PD удирдлагын коэффициентуудыг хиймэл оюун ухаан ашиглан оновчлох нь” *Namuunaa Dashtseren and Tengis Tserendondog*
2	“Speech-Conditioned Image Generation via Direct Projection into a Diffusion Decoder” *Magvan-Erdene Gantulga, Zoljargal Munkhnasan and Orgil Jargalsaikhan*
3	“Монгол гэрийн хаалганы текстур үүсгэх арга” *Orgil Ulzii-Undral and Amarsanaa Ganbold*
4	“Assessment of Vibration Energy Harvesting Potential on Mongolian Freight Trains for IoT Applications” *Batdalai Sukh, Nyamsuren Purevsuren, Myeirambyek Tashkyen, Battur Ganbat and Ulzii-Orshikh Manjaa*
5	“Дээд мөчний нөхөн сэргээх эмчилгээг дэмжих VR тоглоомын хөгжүүлэлт” *Oyunbat Otgonbayar and Oyundolgor Khorloo*
6	“Modeling ecological stability via combined phenotypic plasticity and genetic evolution: a unity-based agent simulation.” *Bilguun Baigalmaa*

Дөрвөн сэнст төхөөрөмжийн PD удирдлагын коэффициентуудыг хиймэл оюун ухаан ашиглан оновчлох нь

1st Дашцэрэн Намуунаа
Электроникийн тэнхим
Шинжлэх Ухаан Технологийн Их Сургууль
Улаанбаатар, Монгол
Namuunaa@must.edu.mn

2nd Цэрэндондог Тэнгис
Электроникийн тэнхим
Шинжлэх Ухаан Технологийн Их Сургууль
Улаанбаатар, Монгол
Tengis@must.edu.mn

***Хураангуй*—Орчин үеийн нисгэгчгүй нисэх төхөөрөмжүүдэд (UAV) тогтвортой удирдлага нь хамгийн чухал асуудлын нэг юм. Энэхүү судалгааны ажилд дөрвөн сэнст нисэх төхөөрөмжийн тогтвортой байдлыг сайжруулах зорилгоор бататган суралцах (Reinforcement Learning, RL)-д суурилсан PD параметр автоматаар тохируулах аргыг боловсруулсан. TD3 (Twin Delayed Deep Deterministic Policy Gradient) алгоритмыг ашиглан Kp, Kd утгуудыг динамикаар тохируулж, уламжлалт PD-тэй харьцуулахаар зорьж байна. Туршилтад `STATE_DIM = 8`, `ACTION_DIM = 4` бүхий загвар ашиглаж, коэффициентуудыг санамсаргүй утгуудаар эхлүүлэн сургасан. Үр дүнгээс харахад шагналын утга сургалтын явцад сайжирч, өнцгийн хэлбэлзэл ±20° орчмоос ±2–5° хүртэл буурсан бөгөөд Kp, Kd коэффициентууд тогтвортой мужид ойртсон нь RL-д суурилсан арга нь уламжлалт PD тохируулгатай харьцуулахад бага алдаа, илүү тогтвортой хариу үзүүлэх боломжтойг харуулж байна.**

***Түлхүүр үгс*: дөрвөн сэнст, PID Controller, бататган сурах арга, TD3, Control System**

I. Удиртгал

Сүүлийн жилүүдэд нисгэгчгүй нисэх төхөөрөмж, дөрвөн сэнст төхөөрөмжийн судалгаа, үйлдвэрлэл, хяналт-шинжилгээ, хүргэлт зэрэг олон салбарт өргөн хэрэглэгдэх болсон. Эдгээр системүүд нь өндөр динамик, шугаман бус шинж чанартай тул тогтвортой удирдлага боловсруулах нь төвөгтэй асуудалд тооцогддог. Ялангуяа орчны нөлөөлөл (салхи, ачаалал, мэдрэгчийн алдаа) болон параметрийн тодорхой бус байдал нь удирдлагын гүйцэтгэлд хүчтэй нөлөөлдөг. Уламжлалт удирдлагын аргуудаас PD (Proportional–Differential) удирдлага нь энгийн бүтэц, хэрэгжүүлэхэд хялбар байдлаараа өргөн хэрэглэгддэг. Гэсэн хэдий ч PD коэффициентуудыг (Kp, Kd) зөв тохируулах нь туршилт, туршлага дээр ихээхэн тулгуурладаг бөгөөд системийн нөхцөл өөрчлөгдөхөд гүйцэтгэл буурах сул талтай. Ялангуяа дөрвөн сэнст төхөөрөмжийн хувьд олон тэнхлэгт, харилцан хамааралтай динамик бүхий системд статик PD тохиргоо хангалтгүй байж болно [1].

Сүүлийн үед хиймэл оюун ухаан болон бататгах сургалт (Reinforcement Learning, RL) нь удирдлагын системд дасан зохицох чадварыг нэмэгдүүлэх шинэ боломжуудыг нээж байна. RL нь агент орчинтой харилцан үйлчилж, шагналын функц дээр суурилан оновчтой бодлогыг өөрөө сурах чадвартай. Үүнийг ашиглан PD коэффициентыг динамикаар тохируулах буюу Коэффициентын тохируулга хийх нь уламжлалт аргуудын сул талыг даван туулах боломжийг олгодог.

Энэхүү судалгааны ажлын зорилго нь RL алгоритмыг ашиглан дөрвөн сэнст төхөөрөмжийн удирдлагад PD параметрүүдийг автоматаар тохируулах арга боловсруулах явдал юм. Тодруулбал, Deep Q-Network (DQN) эсвэл Q-Learning аргачлал ашиглан системийн төлөвт тулгуурлан тохиромжтой Kp, Kd утгуудыг сонгох замаар тогтвортой байдал, хурд, алдааны бууралтыг сайжруулахыг зорьж байна. Энэхүү ажлын хүрээнд дараах асуудлуудыг авч үзнэ. Үүнд:

- Дөрвөн сэнст төхөөрөмжийн системийн динамик загварчлал
- RL орчин, төлөв, үйлдэл, шагналын функцийн тодорхойлолт
- Arduino болон Python хоорондын харилцан ажиллагаанд суурилсан бодит туршилт
- RL болон PD аргын гүйцэтгэлийн харьцуулалт

Энэхүү судалгаа нь RL-д суурилсан удирдлагын аргыг бодит системд хэрэгжүүлэх боломжийг харуулах бөгөөд ирээдүйд өөрөө дасан зохицох, ухаалаг удирдлагын систем боловсруулахад хувь нэмэр оруулах ач холбогдолтой юм [1] - [4].

II. Судлагдсан байдлын судалгаа

Сүүлийн жилүүдэд дөрвөн сэнст нисэх төхөөрөмжийн удирдлагын судалгаа нь уламжлалт загварт суурилсан удирдлагын аргуудаас өгөгдөлд суурилсан, суралцдаг удирдлагын арга руу эрчимтэй шилжиж байна. Дөрвөн сэнст төхөөрөмж нь шугаман бус, олон хувьсагчтай, тэнхлэгүүдийн хоорондын харилцан нөлөөлөл ихтэй динамик систем тул бүх ажлын нөхцөлд нэг тогтмол параметртэй удирдлагын өндөр чанартай удирдлага хэрэгжүүлэхэд хүндрэлтэй байдаг. Иймээс тогтворжилт, хурдан хариу, бага хэтрэлт, гаднын нөлөөлөлд тэсвэртэй байдал зэрэг шаардлагыг зэрэг хангах удирдлагын арга боловсруулах асуудал нь өнөөг хүртэл идэвхтэй судлагдаж байна.

Уламжлалт аргуудаас PD, PID, LQ болон тэдгээрийн төрөлжсөн хувилбарууд дөрвөн сэнст төхөөрөмжийн удирдлагад хамгийн өргөн хэрэглэгддэг. Luukkonen-ийн ажилд дөрвөн сэнст төхөөрөмжийн динамик загварчлал, удирдлагын суурь тэгшитгэлүүд, тогтворжилтын үндсэн зарчмуудыг системтэй авч үзсэн нь энэ чиглэлийн онолын чухал суурь болсон [5]. Bouabdallah *et al.* micro дөрвөн сэнст төхөөрөмж дээр PID ба LQ аргуудыг харьцуулан судалж, уламжлалт удирдлагын практик давуу болон сул талыг тодорхой харуулсан [6]. Мөн Cooper *et al.* классик PID удирдлагаар дөрвөн сэнст төхөөрөмжийн тогтворжилтыг үнэлж, уламжлалт аргуудыг суурь харьцуулалтын загвар болгон ашиглах боломжтойг харуулсан [7].

Гэсэн хэдий ч эдгээр аргуудын гол сул тал нь коэффициентуудыг гараар, эсвэл туршилт-алдааны аргаар оновчлох шаардлагатай байдагт оршино. Системийн параметр, ачаалал, гаднын орчны нөлөөлөл, мэдрэгчийн шуугиан өөрчлөгдөхөд өмнө нь тохируулсан коэффициентуудын гүйцэтгэл буурах магадлалтай. Дөрвөн сэнст төхөөрөмжийн хувьд roll, pitch, yaw тэнхлэгүүдийн хоорондын харилцан холбоос их тул нэг тэнхлэгийн параметрийг сайжруулахад нөгөө тэнхлэгт сөрөг нөлөө үүсэх тохиолдол элбэг байдаг. Иймээс уламжлалт удирдлагын бүтцийг хадгалсан хэвээр коэффициентуудыг автоматаар шинэчлэх ухаалаг арга руу судалгааны чиглэл шилжиж байна.

Энэ хүрээнд бататган суралцах (Reinforcement Learning, RL) аргууд дөрвөн сэнст төхөөрөмжийн удирдлагад ирээдүйтэй шийдэл болж гарч ирсэн. Koch *et al.* UAV attitude control асуудалд PPO, TRPO, DDPG зэрэг RL алгоритмуудыг харьцуулан судалж, RL аргыг зөвхөн дээд түвшний навигацид бус, дотоод attitude control давталтад ашиглах боломжтойг харуулсан [1]. Wu *et al.* дрон hovering удирдлагад supervised reinforcement learning хэрэглэж, сургалтын үр ашгийг нэмэгдүүлэн, hover тогтворжилтын чанарыг сайжруулсан [2]. Эдгээр ажлууд нь RL арга нь шугаман бус, тодорхой бус орчинд дасан зохицох чадвар өндөртэй болохыг харуулдаг.

Сүүлийн үеийн судалгаануудад deep reinforcement learning аргыг дөрвөн сэнст төхөөрөмжийн UAV-ийн удирдлагад бодит хугацаанд хэрэгжүүлэх асуудалд илүү их анхаарч байна. Trad *et al.* deep reinforcement learning суурьтай дөрвөн сэнст төхөөрөмжийн control system-ийг бодит хугацаанд хэрэгжүүлэх боломжийг авч үзсэн [3]. Мөн Sönmez *et al.* дөрвөн сэнст төхөөрөмж UAV-ийн PD controller gains prediction асуудалд RL-д суурилсан арга санал болгож, нислэгийн явцад PD коэффициентуудыг онлайнаар нарийн тохируулах боломжийг харуулсан [4]. Энэ нь controller-ийн бүтцийг бүрэн солихын оронд, уламжлалт PD удирдлагыг RL алгоритмаар дэмжин сайжруулах hybrid чиглэлийн судалгааны чухал жишээ болж байна. RL-д суурилсан коэффициентуудыг гараар тохируулах хамаарлыг бууруулах, өөрчлөгдөж буй орчинд илүү уян хатан хариу үзүүлэх давуу талтай.

Гэсэн хэдий ч RL аргыг бодит дөрвөн сэнстийн системд хэрэгжүүлэхэд хэд хэдэн сорилт байсаар байна.

- Симуляцийн орчинд сурсан бодлогыг бодит төхөөрөмж дээр шууд ашиглахад simulation-to-real зөрүү үүсдэг.
- бодит системд мэдрэгчийн шуугиан, сериал холболтын саатал, механик чичиргээ, тоон квантчилал зэрэг хүчин зүйлс сургалтын тогтвортой байдалд сөргөөр нөлөөлж болно.
- Шагналын функцийг оновчгүй тодорхойлбол агент буруу бодлого сурах, эсвэл тогтворжилт сайжрахын оронд хэт савлалттай зан төлөв үзүүлэх эрсдэлтэй.

Иймээс RL аргыг бодит системд нэвтрүүлэхдээ аюулгүй ажиллагаа, сургалтын тогтвортой байдал, өгөгдлийн чанар зэрэг асуудлыг хамтад нь авч үзэх шаардлагатай.

Ийм нөхцөлд орчин үеийн судалгааны гол чиг хандлага нь уламжлалт удирдлагын найдвартай бүтцийг хадгалсан хэвээр, RL алгоритмаар gain-уудыг автоматаар шинэчилдэг хосолмол архитектур боловсруулахад чиглэж байна. Энэхүү ажлын хүрээнд TD3 алгоритмыг ашиглан дөрвөн сэнст төхөөрөмжийн PD коэффициентуудыг бодит системийн төлөвөөс хамааран динамикаар шинэчлэх нь дээрх судалгааны чиглэлтэй нийцэж байгаа бөгөөд шугаман бус динамик, мэдрэгчийн шуугиан, параметрийн өөрчлөлттэй нөхцөлд илүү уян хатан, дасан зохицох чадвартай удирдлагын шийдэл боловсруулахад чиглэж байна.

III. TD3 алгоритм

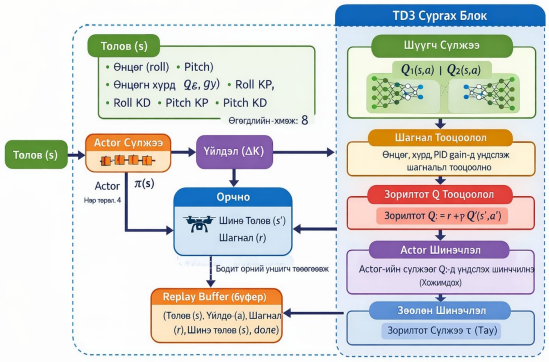
TD3 гэдэг нь Twin Delayed Deep Deterministic Policy Gradient гэсэн нэртэй, үргэлжилсэн action space-тай бодлогын асуудалд зориулагдсан off-policy actor-critic арга юм. Энэ нь DDPG дээр суурилдаг ч DDPG-ийн нийтлэг сул тал болох Q-утгын хэт өндөр үнэлгээ, critic алдаатай байхад actor- буруу чиглэлд сурах, мөн Deterministic policy нарийн “peak” дээр хэт таарах асуудлыг тусгайлан засах 3 өөр механизм нэмж өгдөг.

TD3 алгоритм нь DDPG шиг ганц critic хэрэглэвэл critic зарим action-ыг бодитоос нь өндөр үнэлэх үед actor тэр алдаатай өндөр үнэлгээг “ашиглаад” улам буруу action руу чиглэдэг сул талтай. TD3 энэ асуудлыг бууруулахын тулд хоёр critic-ийн target-оос бага утгыг нь авдаг.

TD3-ийн гол онцлог нь хоёр critic буюу шүүгч сүлжээ $Q_1(s,a)$ болон $Q_2(s,a)$-г зэрэг ашигладаг. Эдгээр critic сүлжээ нь өгөгдсөн төлөв ба үйлдлийн хослолын чанарыг тус тус үнэлж, target Q утгыг тооцохдоо энэ хоёрын бага утгыг сонгодог. Ингэснээр DDPG алгоритмд түгээмэл ажиглагддаг Q-утгын хэт өндөр үнэлгээний асуудлыг бууруулдаг. Мөн actor сүлжээ critic-ээс илүү удаан шинэчлэгддэг бөгөөд үүнийг Delayed policy update гэж нэрлэдэг. Энэ механизм нь actor-ийг тогтворгүй critic үнэлгээн дээр тулгуурлан буруу сурахаас сэргийлнэ.

Сургалтын явцад target Q утга нь шагнал болон дараагийн төлөвийн үнэлгээнд үндэслэн тооцогдоно. Үүний дараа critic сүлжээнүүдийн алдагдлыг бууруулан шинэчилж, тодорхой алхам тутамд actor сүлжээг шинэчилнэ. Мөн үндсэн ба зорилтот сүлжээнүүдийн хооронд зөөлөн шинэчлэл τ коэффициентоор хийгддэг. Ийм байдлаар TD3 алгоритм нь дөрвөн сэнст төхөөрөмжийн roll болон pitch

тэнхлэгийн төлөвөөс хамааран коэффициентуудыг ааж-маар оновчлон сурч, илүү тогтвортой, бага савлалттай удирдлагын бодлого үүсгэдэг.



Зураг 1. TD3 алгоритмын урсгал

A. Шагналын функц

RL агентын суралцах чанар нь шагналын функцийн зөв тодорхойлтоос ихээхэн шалтгаалдаг . Энэхүү ажилд шагналын функцийг дөрвөн сэнст нисэх төхөөрөмжийн өнцгийн алдаа болон өнцгийн хурдны хэмжээг багасгах зарчимд тулгуурлан тодорхойлов. Сургалтын явцад төхөө-рөмж тохируулсан утгад ойртож, савлалт багасах тусам өндөр шагнал авах, харин огцом хазайлт, их хурд, тогтвор-гүй байдал үүсэх тусам сөрөг шагнал авах юм. Шагналын функцийг дараах ерөнхий хэлбэрээр илэрхийлж болно.

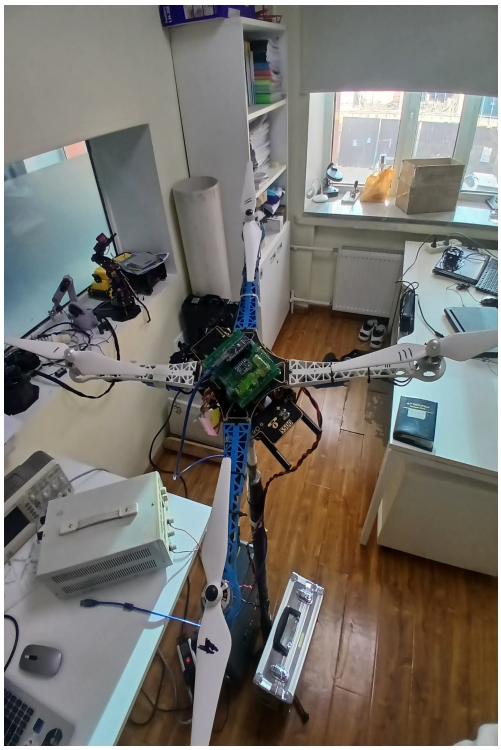
$$r_t = -(w_1|\phi_t| + w_2|\theta_t| + w_3|\dot{\phi}_t| + w_4|\dot{\theta}_t|) \qquad (1)$$

Энд w_1,w_2,w_3,w_4 нь коэффициентууд бөгөөд өнцгийн ал-даа, өнцгийн хурдны аль нь илүү чухал болохыг илэр-хийлнэ. Хэрэв систем тохируулсан утга орчимд бага хурдтай тогтвортой болсон бол нэмэлт эерэг бонус өгч болно. Үүний эсрэгээр хэт их налалт үүсвэл эпизодыг дуусгаж, том хэмжээний сөрөг шагнал өгч болно. Ийм бүтэцтэй шагналын функц нь агент зөвхөн эцсийн алдааг багасгахаас гадна шилжилтийн процессын чанарыг давхар сайжруулахад чиглүүлдэг. Өөрөөр хэлбэл хэтрэлт багатай, хэлбэлзэл багатай, илүү сайн хариу үүсгэхэд тусалдаг

IV. Системийн хэрэгжилт ба туршилтын орчин

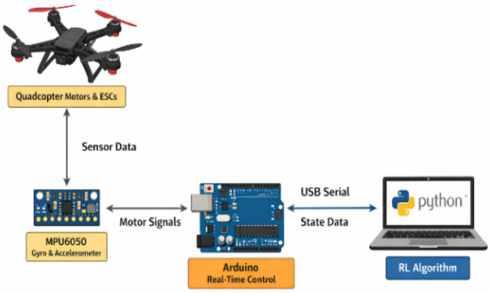
Энэхүү судалгаанд боловсруулсан систем нь хоёр түвшний архитектураас бүрдэх бөгөөд доод түвшинд Arduino микроконтроллёр нь мэдрэгчийн өгөгдөл унших, өнцөг тооцоолох, моторын үндсэн удирдлагыг хэрэгжүүлнэ. Дээд түвшинд Python орчинд TD3 алгоритм ажиллаж, системийн тухайн үеийн төлөвөөс хамааран коэффициентуудыг шинэчилнэ. Ийм модульчлагдсан бүтэц нь бодит хугацааны удирдлага болон тооцооллын хувьд хүнд RL сургалтыг тусгаарлан хэрэгжүүлэх боломж олгосон.

Arduino талд MPU6050 мэдрэгчээс дөрвөн сэнст төхөөрөмжийн өнцөг болон өнцгийн хурдтай холбоотой



Зураг 2. Туршилтад ашиглаж буй дөрвөн сэнст

мэдээллийг уншиж, roll, pitch, gx, gy зэрэг хувьсагчийг Python- тал руу сериалаар дамжуулан илгээдэг. MPU6050-ийн DMP (Digital Motion Processor) функцийг ашиглан тодорхой хурдтайгаар байрлалын мэдээллийг авна. Энэхүү мэдээллээр pitch-roll өнцгүүдийг бодож авна.



Зураг 3. Системийн ерөнхий бүтэц

MPU6050 мэдрэгчийн DMP функцийг ашигласнаар төхөөрөмжийн чиглэлийн quaternion өгөгдлийг бодит хуга-цаанд авна. Quaternion (q_0, q_1, q_2, q_3)-оос roll болон pitch өнцгийг дараах байдлаар тооцоолно:

$$\text{roll} = \tan^{-1}\left(\frac{2\left(q_0 q_1 + q_2 q_3\right)}{1 - 2\left(q_1^2 + q_2^2\right)}\right) \qquad (2)$$

$$\text{pitch} = \sin^{-1}\left(2\left(q_0 q_2 - q_3 q_1\right)\right) \qquad (3)$$

Эдгээрийг градусаар илэрхийлэхдээ:

$$\text{rollDeg} = \text{degrees}\left(\tan^{-1}\left(\frac{2\left(q_0 q_1 + q_2 q_3\right)}{1 - 2\left(q_1^2 + q_2^2\right)}\right)\right) \tag{4}$$

$$\text{pitchDeg} = \text{degrees}\left(\sin^{-1}\left(2\left(q_0 q_2 - q_3 q_1\right)\right)\right) \tag{5}$$

Мөн өнцгийн хурдны мэдээллийг гироскопоос авч,

$$g_x = \text{gyroX}, \qquad g_y = \text{gyroY}$$

гэж тэмдэглэн roll болон pitch тэнхлэгийн өнцгийн хурд болгон ашиглана. Ийнхүү roll, pitch g_x, g_y хувьсагчдыг Python тал руу сериалаар дамжуулан илгээж, RL агентын төлөвийн оролт болгон хэрэглэдэг. Python кодонд ч roll, pitch, gx, gy утгуудыг state болон reward тооцоонд ашиглаж байна.

Үүний зэрэгцээ Python талаас шинэчлэгдсэн $K_{(p,r)}$, $K_{(d,r)}$, $K_{(p,p)}$, $K_{(d,p)}$ коэффициентуудыг хүлээн авч моторын удирдлагад ашиглана. Өөрөөр хэлбэл, Arduino нь бодит объекттой харилцах интерфейс ба локал хяналтын давхарга болж, Python- нь ухаалаг шийдвэр гаргах агентын үүргийг гүйцэтгэнэ. Ийм хуваарилалт нь бодит туршилтын үед системийн модульчлал, өргөтгөх боломж, алдаа оношлох чадварыг сайжруулдаг.

PD удирдлагад зөвхөн тухайн үеийн алдаа бус, алдаа хэр хурдан өөрчлөгдөж байгааг харгалзан үздэг. Үүнийг дараах байдлаар бичнэ.

$$\begin{aligned} \dot{e}_\phi(t) &= \frac{de_\phi(t)}{dt} \\ \dot{e}_\theta(t) &= \frac{de_\theta(t)}{dt} \end{aligned} \tag{6}$$

Үүнд:

- $e_\phi(t)$ - roll тэнхлэгийн алдаа,
- $e_\theta(t)$ - pitch тэнхлэгийн алдаа,
- $\dot{e}_\phi(t)$, $\dot{e}_\theta(t)$ - эдгээр алдааны өөрчлөгдөх хурд.

Алдаа нь дөрвөн сэнстийн хэлбэлзэл, савлалтын шинж чанарыг бууруулахад чухал үүрэгтэй. Практикт энэхүү хэмжигдэхүүн нь гироскопын мэдрэгчийн өгөгдөлтэй холбоотойгоор ойролцоолон үнэлэгдэж болно.

Python орчинд хэрэгжүүлсэн RL агент нь TD3 алгоритм дээр суурилсан. Энэхүү алгоритмыг сонгосон шалтгаан нь тасралтгүй үйлдлийн орон зайд тогтвортой суралцах чадвар сайтай, DDPG төрлийн аргуудад түгээмэл ажиглагддаг Q-утгын хэт үнэлгээний асуудлыг хоёр critic сүлжээ ашиглан бууруулдагт оршино. Тухайн системийн төлөв нь нийт найман хэмжээст векторооос бүрдэх бөгөөд үүнд roll өнцөг, pitch өнцөг, тухайн тэнхлэгүүдийн өнцгийн хурд, мөн тухайн мөчид хэрэглэгдэж буй дөрвөн коэффициент орсон.

RL агентын оролтын төлөвийг дараах вектороор илэрхийлж болно.

$$s_t = [\phi_t,\ \theta_t,\ g_{x,t},\ g_{y,t},\ K_{p,\phi,t},\ K_{d,\phi,t},\ K_{p,\theta,t},\ K_{d,\theta,t}] \tag{7}$$

Энд

- ϕ_t, θ_t — тухайн мөчийн roll болон pitch өнцөг,
- $g_{x,t}$, $g_{y,t}$ — өнцгийн хурдтай холбоотой мэдрэгчийн хэмжилт,
- $K_{p,\phi,t}$, $K_{d,\phi,t}$, $K_{p,\theta,t}$, $K_{d,\theta,t}$ — тухайн мөчид хэрэглэж буй коэффициентууд болно.

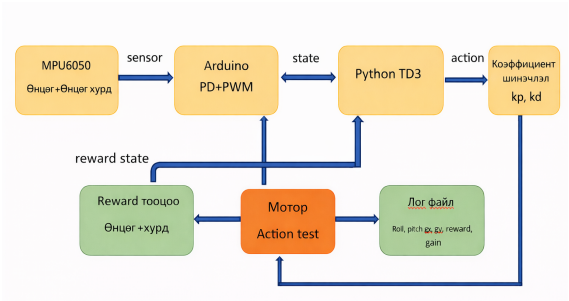
Иймт төлөв сонгосноор агент зөвхөн механик төлөвийг ажиглаад зогсохгүй, өөрийн өмнөх шийдвэрийн нөлөөллийг хамтад нь харгалзан үзэж, дараагийн коэффициентийн өөрчлөлтийг илүү үндэслэлтэйгээр гаргах боломжтой болно. Кодын түвшинд энэхүү төлөвийн хэмжээг `STATE_DIM = 8`, үйлдлийн хэмжээг `ACTION_DIM = 4` гэж тодорхойлсон.

RL агентын үйлдлийг ашиглан коэффициентуудыг дараах байдлаар шинэчилнэ:

$$\begin{aligned} K_{p,\phi,t+1} &= \text{clip}(K_{p,\phi,t} + \Delta K_{p,\phi,t},\ K_{p,\min},\ K_{p,\max}), \\ K_{d,\phi,t+1} &= \text{clip}(K_{d,\phi,t} + \Delta K_{d,\phi,t},\ K_{d,\min},\ K_{d,\max}), \\ K_{p,\theta,t+1} &= \text{clip}(K_{p,\theta,t} + \Delta K_{p,\theta,t},\ K_{p,\min},\ K_{p,\max}), \\ K_{d,\theta,t+1} &= \text{clip}(K_{d,\theta,t} + \Delta K_{d,\theta,t},\ K_{d,\min},\ K_{d,\max}) \end{aligned} \tag{8}$$

Энд $\text{clip}(\cdot)$ функц нь коэффициентуудын утгыг урьдчилан тогтоосон доод болон дээд хязгаарт барина.

Сургалтын үр дүнд гарсан эдгээр өнцгийн утгуудаар дроны байрлал, хазайлтыг үнэлэх боломжтой. Мөн эдгээр утгууд нь дроны моторуудын PWM утгыг удирдахад ашиглагдана. Туршилтын явцад коэффициентуудыг эхлээд



Зураг 4. RL-д суурилсан дөрвөн сэнстийн PD tuning системийн блок-схем

санамсаргүй мужид (ойролцоогоор 0–3) эхлүүлж, дараа нь эпизод бүрээр TD3 агентын бодлого сайжрах тусам шинэчлэгдсэн.

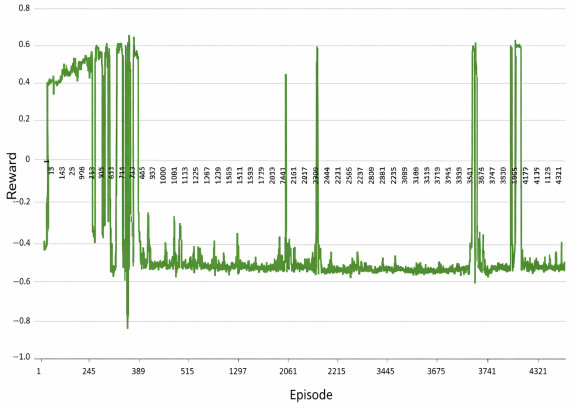
Энэхүү судалгааны туршилт нь симуляцийн орчинд бус, бодит дөрвөн сэнст төхөөрөмж дээр хэрэгжсэн болно. Туршилтад нийт 1кг-ийн масстай дөрвөн сэнст ашигласан. 8x4.5 inch -ийн урттай дөрвөн сэнс , 1000kv -ийн дөрвөн мотор, Arduino NaNo -г контроллер, тэнхлэг хэмжилтэд MPU6050 мэдрэгчийг тус тус сонгон ашигласан.

V. Туршилтын үр дүн

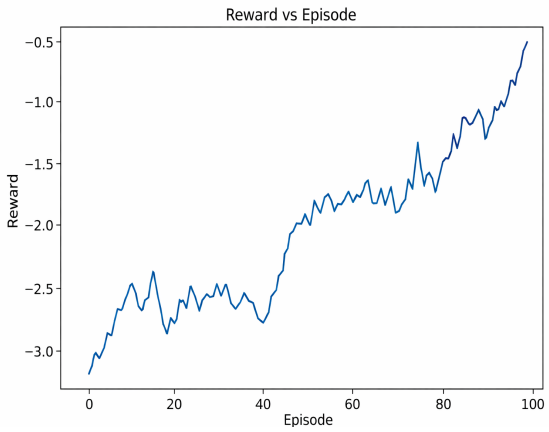
Туршилтын явцад RL агентын суралцах явцыг шагнал, өнцгийн хэлбэлзэл, коэффициентуудын өөрчлөлт гэсэн үндсэн үзүүлэлтүүдээр үнэлж үзлээ. Шагналын графикаас харахад сургалтын эхний хэсэгт агентын үйлдлүүд тогтворгүй, санамсаргүй шинжтэй байснаас шалтгаалан reward харьцангуй бага буюу сөрөг утгатай байсан. Харин

total_step	kp_roll	kd_roll	kp_pitch	kd_pitch	roll	pitch	gx	gy	reward
13746	3	0.90953	3	0.03377	-1.883	3.25433	-0.44767	-0.08167	-0.44571
13747	3	0.89535	3	0.02931	-2.09033	3.496	-0.54733	-0.48633	-0.47711
13748	3	0.88684	3	0.01117	-2.333	3.722	-0.68433	-0.089	-0.51712
13749	3	0.88295	3	0	-2.36167	3.69367	0.27733	0.42	-0.50531
13750	3	0.8739	3	0	-2.18933	3.57567	-0.135	0.11933	-0.49398
13751	2.98827	1.38471	3	0	-2.11167	3.45667	0.34833	0.173	-0.48063
13752	2.98895	1.36765	3	0	-1.954	3.46267	-0.15267	0.26967	-0.46794
13753	2.99135	1.34409	3	0	-1.92233	3.27	-0.02567	0.229	-0.45536
13754	2.99397	1.32473	3	0	-1.87167	3.29533	-0.295	-0.107	-0.45517
13755	2.99422	1.3113	3	0	-2.04	3.46933	-0.229	-0.17567	-0.47422
13756	2.99615	1.30286	2.99997	0	-2.27033	3.66433	-0.794	-0.23167	-0.50363
13757	2.99665	1.28261	3	0	-2.52067	3.792	-0.28267	0.16533	-0.52904
13758	3	1.28835	3	0	-2.30033	3.72733	0.237	0.32067	-0.51629
13759	3	1.26215	3	0	-2.09567	3.38067	0.90833	-0.018	-0.4821
13760	3	1.24518	3	0	-1.972	3.33633	0.22133	-0.074	-0.45612
13761	3	1.23386	3	0	-1.864	3.42767	-0.407	0.31567	-0.46012
13762	3	1.23372	3	0	-2.07267	3.43167	-0.36667	-0.30767	-0.47137
13763	2.99955	1.22766	2.99986	0	-2.184	3.66833	0.00233	-0.318	-0.49861
13764	3	1.23272	3	0	-2.40367	3.742	-1.46567	0.63333	-0.52038

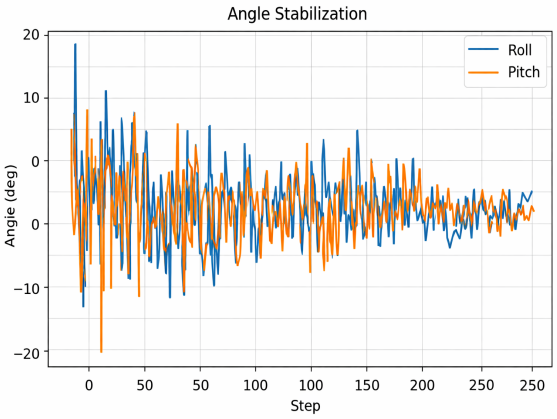
Зураг 5. Сериалаар дамжуулан бүртгэгдсэн төлөв ба шагналын хэсэг өгөгдөл



Зураг 6. Туршилтын явцад сурч буй дата



Зураг 7. Шагналын өсөлт



Зураг 8. Өнцгийн тогтворжилт

туршилтын тоо нэмэгдэхийн хэрээр шагнал аажмаар өссөн нь агент илүү оновчтой коэффициент сонгож эхэлснийг илтгэнэ.

Энэ ажлын хүрээнд RL агентын суралцах явцыг үндсэндээ reward, өнцгийн хэлбэлзэл, коэффициентын өөрчлөлтөөр үнэлсэн. Гэвч удирдлагын системийн гүйцэтгэлийг илүү бүрэн үнэлэхийн тулд overshoot, setpoint time, steady-state error, RMSE зэрэг сонгодог control metrics-ийг хамтатган тооцох шаардлагатай. Шагналын өсөлт нь RL агентын сургалтын дотоод шалгуур болохоос удирдлагын чанарыг дангаар бүрэн илэрхийлэхгүй. Иймээс цаашдын судалгаанд уламжлалт PD болон RL-д суурилсан PD аргуудыг дээрх үзүүлэлтүүдээр нэг ижил туршилтын нөхцөлд харьцуулан үнэлэх нь зүйтэй.

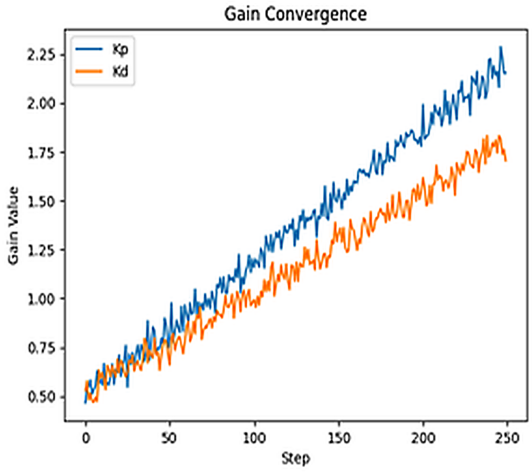
Шагнал болон эпизодын графикаас харахад сургалтын эхний шатанд утга нь ойролцоогоор −3 орчим байсан бол дараагийн шагналуудын явцад аажмаар өсөж −0.5 орчимд хүрсэн байна. Шагналын өсөлт жигд биш, тодорхой хэсгүүдэд хэлбэлзэлтэй байгаа нь бодит системийн шуугиан, механик чичиргээ, мэдрэгчийн алдаа, сериал холбооны саатал, off-policy сургалтын онцлогтой холбоотой байж болно Зураг 7.

Сургалтын эхэнд roll болон pitch тэнхлэгүүд ±20° орчим хүчтэй савлаж байсан бол сургалтын сүүл үед хэлбэлзлийн амплитуд мэдэгдэхүйц багасаж ±2–5° орчимд тогтворжсон байна. Энэ нь RL зөвхөн тохируулсан утгад ойртуулах биш, мөн системийн савлалт, огцом хариуг багасгах чиглэлд коэффициентуудыг тохируулж чадсаныг харуулж байна. дөрвөн сэнстийн механик чичиргээ, мэдрэгчийн шуугиан, сэнсний микро саатлыг тооцвол энэ үр дүн нь бодит туршилтын хувьд сайн үзүүлэлт юм Зураг 8.

Kp болон Kd утгууд сургалтын эхэнд санамсаргүй утгуудаас эхэлсэн боловч эпизодын явцад тодорхой муж руу дөхөн тогтворжиж байгааг харж болно. Энэ нь RL -ийн бодит системийн динамикт нийцсэн ашигтай коэффициентуудыг олж чадсаныг илтгэж байна Зураг 9.

А. Хэлэлцүүлэг

Энэхүү судалгааны үр дүнгээс харахад RL-д суурилсан PD tuning арга нь бодит дөрвөн сэнст нисэх төхөөрөмж дээр хэрэгжих боломжтой бөгөөд уламжлалт гараар тохируулдаг аргыг тодорхой хэмжээнд орлох чадвартай байна. Ялангуяа системийн параметрүүд, ачаалал, механик орчин, мэдрэгчийн шуугиан өөрчлөгдөх үед статик коэффициентуудтай хянагчаас илүү уян хатан шийдэл болж чадаж байна.



Зураг 9. Kp ба Kd коэффициентын тогтворжилт

Гэсэн хэдий ч энэхүү аргын үр дүн нь хэд хэдэн нөхцөлөөс ихээхэн хамаарна. Үүнд:

1) Шагналын функцийг оновчтой загварчлах нь маш чухал. Хэрэв шагналын функц зохистой сонгогдоогүй бол агент буруу чиглэлд суралцах, эсвэл хэт савлалтыг зөвшөөрөх сул талтай.
2) Бодит систем дээр сериалаар өгөгдөл дамжуулах үед үүсэх саатал, мэдрэгчийн шуугиан, тоон квантчилал зэрэг хүчин зүйлс нь RL сургалтын тогтвортой байдалд сөргөөр нөлөөлж болно.
3) Бодит төхөөрөмж дээр туршилт хийх явцад аюулгүй ажиллагааны асуудал маш чухал байр суурь эзэлнэ.

ДҮГНЭЛТ

Энэхүү судалгаагаар дөрвөн сэнст төхөөрөмжийн roll болон pitch тэнхлэгийн тогтворжилтыг сайжруулах зорилгоор TD3 бататган суралцах алгоритмд суурилсан коэффициентуудыг автоматаар тохируулах аргыг бодит систем дээр туршив. Сургалтын явцад шагнал аажмаар багасаж байна. харин өнцгийн алдаа ±20° орчим хүчтэй савлаж байсан бол сургалтын сүүл үед хэлбэлзлийн амплитуд мэдэгдэхүйц багасаж ±2–5° орчимд тогтворжсон. Kp болон Kd утгууд сургалтын эхэнд санамсаргүй утгуудаас эхэлсэн боловч туршилтын эпизодуудын явцад тодорхой муж руу дөхөж Kp=2.1, kd=1.75 орчимд тогтворжсон нь RL-ийн бодит хяналтын параметрүүдийг амжилттай суралцаж буйг харуулж байна. Судалгааны гол практик ач холбогдол нь дөрвөн сэнст төхөөрөмжийн шугаман бус, олон хувьсагчтай системд гар аргаар коэффициент тохируулах шаардлагыг бууруулж, нөхцөл өөрчлөгдөхөд дасан зохицох хяналтын суурь шийдэл санал болгосонд оршино. Туршилтын хэрэгжилт нь одоогоор PD коэффициентод төвлөрсөн боловч бүрэн PID сургах боломж бүрдсэн гэж үзэж байна.

Ашигласан материал

[1] W. Koch, R. Mancuso, R. West, and A. Bestavros, “Reinforcement Learning for UAV Attitude Control,” *arXiv preprint arXiv:1804.04154*, Apr. 2018.

[2] J. Wu, Z. Yang, H. Zhuo, C. Xu, C. Zhang, N. He, L. Liao, and Z. Wang, “A Supervised Reinforcement Learning Algorithm for Controlling Drone Hovering,” *Drones*, vol. 8, no. 3, Art. no. 69, 2024, doi: 10.3390/drones8030069.

[3] T. Y. Trad, K. Choutri, M. Lagha, S. Meshoul, F. Khenfri, R. Fareh, and H. Shaiba, “Real-Time Implementation of Quadrotor UAV Control System Based on a Deep Reinforcement Learning Approach,” *Computers, Materials & Continua*, vol. 81, no. 3, pp. 4757–4786, 2024, doi: 10.32604/cmc.2024.055634.

[4] S. Sönmez, L. Montecchio, S. Martini, M. J. Rutherford, A. Rizzo, M. Stefanovic, and K. P. Valavanis, “Reinforcement Learning-Based PD Controller Gains Prediction for Quadrotor UAVs,” *Drones*, vol. 9, no. 8, Art. no. 581, 2025, doi: 10.3390/drones9080581.

[5] T. Luukkonen, *Modelling and Control of Quadcopter*. Espoo, Finland: Aalto University School of Science, Aug. 2011.

[6] S. Bouabdallah, A. Noth, and R. Siegwart, “PID vs LQ Control Techniques Applied to an Indoor Micro Quadrotor,” in *Proc. 2004 IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, 2004, pp. 2451–2456, doi: 10.1109/IROS.2004.1389776.

[7] Y. N. Cooper, R. K. Ganesh Ram, V. Kalaichelvi, and V. Bhatia, “Stabilization and Control of an Autonomous Quadcopter,” *Applied Mechanics and Materials*, vol. 666, pp. 161–165, 2014, doi: 10.4028/www.scientific.net/AMM.666.161.

[8] J. Hwangbo, I. Sa, R. Siegwart, and M. Hutter, “Control of a Quadrotor with Reinforcement Learning,” *IEEE Robotics and Automation Letters*, 2017. [Online]. Available: arXiv:1707.05110.

[9] C.-H. Pi, Y.-W. Dai, K.-C. Hu, and S. Cheng, “General Purpose Low-Level Reinforcement Learning Control for Multi-Axis Rotor Aerial Vehicles,” *Sensors*, vol. 21, no. 13, Art. no. 4560, 2021, doi: 10.3390/s21134560.

[10] Z. Zhang, X. Zhang, Y. Chen, and H. Li, “Control of UAV Quadrotor Using Reinforcement Learning and Robust Compensator,” *IET Cyber-Systems and Robotics*, 2023, doi: 10.1049/cth2.12496.

[11] P. C. Yuste, J. A. Iglesias Martínez, and M. A. Sanchis de Miguel, “Simulation-Based Evaluation of Model-Free Reinforcement Learning Algorithms for Quadcopter Attitude Control and Trajectory Tracking,” *Neurocomputing*, vol. 608, Art. no. 128362, 2024, doi: 10.1016/j.neucom.2024.128362.

[12] Y. Ren, F. Zhu, S. Sui, Z. Yi, and K. Chen, “Enhancing Quadrotor Control Robustness with Multi-Proportional–Integral–Derivative Self-Attention-Guided Deep Reinforcement Learning,” *Drones*, vol. 8, no. 7, Art. no. 315, 2024, doi: 10.3390/drones8070315.

[13] Y. Yan, H. Cao, B. Zhang, W. Ni, B. Wang, and X. Ma, “Large-Scale Solar-Powered UAV Attitude Control Using Deep Reinforcement Learning in Hardware-in-Loop Verification,” *Drones*, vol. 8, no. 9, Art. no. 428, 2024, doi: 10.3390/drones8090428.

[14] D. Wada, S. A. Araujo-Estrada, and S. Windsor, “Unmanned Aerial Vehicle Pitch Control Using Deep Reinforcement Learning with Discrete Actions in Wind Tunnel Test,” *Aerospace*, vol. 8, no. 1, Art. no. 18, 2021, doi: 10.3390/aerospace8010018.

[15] A. T. Azar, A. Koubaa, N. Ali Mohamed, H. A. Ibrahim, Z. F. Ibrahim, M. Kazim, A. Ammar, B. Benjdira, A. M. Khamis, I. A. Hameed, and A. Al-Darraji, “Drone Deep Reinforcement Learning: A Review,” *Electronics*, vol. 10, no. 9, Art. no. 999, 2021, doi: 10.3390/electronics10090999.

[16] E. Kaufmann, L. Bauersfeld, A. Loquercio, M. Müller, V. Koltun, and D. Scaramuzza, “Champion-Level Drone Racing Using Deep Reinforcement Learning,” *Nature*, vol. 620, pp. 982–987, 2023, doi: 10.1038/s41586-023-06419-4.

Speech-Conditioned Image Generation via Direct Projection into a Diffusion Decoder

Magvan-Erdene G.
SICT, MUST
Ulaanbaatar, Mongolia
magvan-erdene@must.edu.mn

Zoljargal M.
SICT, MUST
Ulaanbaatar, Mongolia
zzoloo2727@gmail.com

Orgil J.
SICT, MUST
Ulaanbaatar, Mongolia
orgil@must.edu.mn

***Abstract*—Text-to-image generation systems require a text prompt as input, discarding the richer acoustic signal present in human speech—prosody, emotional tone, speaking rate, and speaker identity. We present a minimalist framework for speech-conditioned image generation that eliminates the text bottleneck entirely. A frozen Qwen2AudioEncoder processes raw audio into a sequence of 1280-dimensional hidden states. A single linear projection—the only trainable component in the system—maps these into the 4096-dimensional T5 conditioning slot of a frozen Stable Diffusion 3.5 Medium MM-DiT decoder. Both CLIP timestep-embedding slots are zeroed throughout training and inference. Training proceeds in two stages on a synthetic voice-cloned audio–image dataset built on top of MS-COCO: a projection warmup stage using a cosine similarity auxiliary loss, followed by a joint denoising and auxiliary loss stage. Our system achieves FID 122.4 on a held-out split—competitive with a Whisper-based ASR bridge (FID 116.8) that discards all acoustic-affective content—while substantially outperforming all baselines on emotional resonance in human evaluation (3.8 vs. 2.9). One linear layer suffices to bridge frozen speech and image generation backbones.**

***Index Terms*—speech-to-image generation, multimodal learning, diffusion transformers, audio-visual learning, emotion-aware generation**

I. Introduction

The dominant paradigm for conditional image generation requires a text prompt. From DALL-E 3 [1] to Stable Diffusion 3.5 [2], conditioning on natural language has become the de facto interface for creative image synthesis. Yet human communication is fundamentally multimodal: spoken language simultaneously encodes lexical content, prosody, emotion, speaking rate, and speaker identity—none of which survive the transcription step that converts speech to text. A prompt for "a dog" is indistinguishable whether the speaker was joyful or frightened, yet those differences carry genuine communicative intent that could meaningfully shape the mood and style of a generated image.

Speech-to-image generation has attracted growing interest as a way to bypass the text bottleneck [3], [4]. Prior approaches fall into two main categories. The first—and currently dominant—category uses automatic speech recognition (ASR) as a preprocessing bridge: audio is transcribed by a model such as Whisper [5], and the resulting text feeds a standard text-to-image pipeline. This recovers lexical semantic content efficiently but discards all suprasegmental acoustic features that encode affect, emphasis, and expressiveness. The second category fine-tunes components of the generative backbone—often cross-attention layers or CLIP-aligned adapters—to accept audio embeddings directly. This increases training cost and constrains the choice of generation model.

We propose a third approach: *direct linear projection* from a frozen speech encoder into the conditioning token slot of a frozen image decoder, with no intermediate text, no adapter, and no backbone fine-tuning. The insight is architectural: large-scale speech encoders such as Qwen2AudioEncoder [6] produce 1280-dimensional hidden state sequences that are richly organized by audio-language alignment pretraining; the T5 conditioning slot of SD 3.5 Medium MM-DiT expects 4096-dimensional token sequences. A single $1280 \rightarrow 4096$ linear layer bridges these two spaces. Both backbones remain fully frozen throughout training; only the linear projection—approximately 5.2M parameters—is updated.

This hypothesis is motivated by the success of linear-probe cross-modal projection in vision-language models. LLaVA [7] demonstrated that a single linear projection is sufficient to align frozen CLIP visual features with a frozen large language model, enabling rich multimodal dialogue with minimal training. We apply the same principle across speech and image generation modalities.

We train on a synthetic voice-cloned audio–image dataset built on top of MS-COCO (Section IV-A); each COCO image is paired with multiple voice-cloned audio variants spanning five emotional registers synthesized by an open-source TTS using RAVDESS reference speakers. Training proceeds in two stages: a projection warmup phase (Stage 1) using a cosine similarity auxiliary loss between mean-pooled audio embeddings and CLIP ViT-bigG image embeddings, and a joint training phase (Stage 2) combining the denoising objective with the auxiliary loss weighted at λ=0.1. Inference uses classifier-free guidance at scale 7.0 with no text input.

Our contributions are as follows:

- The first demonstration that a single linear layer can directly condition a frozen MM-DiT diffusion decoder on raw speech audio via the T5 token slot, requiring neither an adapter, a Q-Former, nor any fine-tuning of either backbone. Unlike adapter-based methods (e.g., AudioToken [3]) that require per-input optimization at inference, our projection is a single static forward pass. Unlike ASR

bridges, it preserves all suprasegmental acoustic-affective content discarded during transcription.
- A text-free, emotion-sensitive image generation system achieving FID 122.4 on a held-out audio–image split—competitive with a strong ASR-bridge baseline—while significantly outperforming it on human-rated emotional resonance (3.8 vs. 2.9 on a 5-point Likert scale).
- An empirical validation of two-stage training with an audio-image cosine auxiliary loss, showing that the Stage 1 warmup reduces final FID by 10.3 points compared to cold-start joint training.

II. Related Work

A. Text-to-Image Diffusion Models

Latent diffusion models [8] decouple image generation from pixel-space optimization by encoding images through a variational autoencoder (VAE) and applying the diffusion process in a compact latent space. Subsequent transformer-backbone variants demonstrated improved quality with model scale; the MM-DiT (Multimodal Diffusion Transformer) architecture introduced in Stable Diffusion 3 [2] applies bidirectional joint attention over image patch tokens and text conditioning tokens within the same transformer blocks, enabling tight semantic coupling between the two modalities during denoising. SD 3.5 Medium provides three independent conditioning streams: a T5-XXL token sequence (4096d per token), and two pooled CLIP embeddings (CLIP ViT-L at 768d and CLIP ViT-bigG at 1280d) added to the timestep embedding. DALL-E 3 [1] demonstrates that rich, dense semantic alignment in the conditioning pathway—achieved through highly descriptive synthetic captions—drives large gains in text-image fidelity. Our work investigates whether an equivalently rich conditioning signal can be derived from raw speech without any text.

B. Audio and Speech Encoders for Vision

Prior work has shown that speech and image representations can be jointly aligned without text supervision, enabling audio-image retrieval directly from raw speech. AudioToken [3] adapts text-conditioned diffusion models to environmental audio by learning a token embedding from CLIP-aligned audio features using Textual Inversion; however, it requires audio-specific optimization per input at inference. Lee et al. [4] use CLIP-based sound guidance to manipulate images in latent space, demonstrating semantic audio-visual correspondence without text. The ASR bridge remains the dominant approach for speech-to-image generation, leveraging Whisper [5] for high-quality transcription; its key limitation is the discarding of all suprasegmental, affective content.

Qwen2-Audio [6] trains a 7B audio-language model whose encoder (Qwen2AudioEncoder) processes 128-channel log-mel spectrograms into 1280-dimensional hidden states per frame. Trained on diverse audio—speech, sound events, and music—with large-scale audio-language alignment objectives, the encoder's hidden states capture both lexical meaning and the acoustic-affective variation between utterances, making it an ideal backbone for our approach.

C. Cross-Modal Projection Without Fine-Tuning

LLaVA [7] established the precedent that a single linear projection layer is sufficient to align a frozen CLIP image encoder with a frozen large language model, enabling multimodal instruction following with minimal trainable parameters. This "linear probe" approach preserves the linear geometry of the source encoder's representation space: semantic directions and relative distances in the encoder space are carried intact into the target space under linear transformation, rather than being warped by nonlinear adapters. CLIP [9] itself demonstrates that visual and language representations can be brought into alignment via contrastive training on image-text pairs, producing linearly structured embedding spaces. CLAP [10] extends this contrastive paradigm to the audio domain with a large-scale audio-language pretrained model. Our auxiliary loss exploits the dimension match between the Qwen2AudioEncoder mean-pooled output and the CLIP ViT-bigG image embedding (both 1280d) to provide a text-free alignment signal during Stage 1 warmup.

D. Speech Emotion in Latent Spaces

Human speech encodes emotional state through multiple acoustic channels: fundamental frequency (F_0) contours encode pitch variation and intonation; intensity envelopes reflect emphasis and energy; spectral features including formant trajectories and speaking rate encode articulation style. These features are directly represented in the 128-channel log-mel spectrogram consumed by Qwen2AudioEncoder. Large-scale audio encoders trained with audio-language alignment have been shown to organize their latent spaces such that the same words spoken in different emotional registers occupy distinct regions, with acoustic-affective variation represented as directional structure in the hidden state space [3], [6]. Because a linear projection preserves linear structure—directions and relative distances in the source space map proportionally into the target space—acoustic-affective variation in the encoder output translates to proportional variation in the projected 4096-dimensional token sequence. This geometric preservation is precisely what MM-DiT's joint attention can exploit: different projected audio sequences for the same semantic content but different emotional registers will produce perceptibly different visual outputs, as we confirm empirically.

E. Positioning Relative to Prior Work

Our method occupies a distinct position in the design space that prior work does not fill. *ASR-bridge* approaches (Whisper $\rightarrow$ SD 3.5) maximize lexical semantic fidelity but irrecoverably discard prosodic and affective content during transcription; our system matches their FID (122.4 vs. 116.8) while substantially outperforming them on human-rated emotional resonance (3.8 vs. 2.9). *Adapter and fine-tuning methods* such as AudioToken [3] learn to represent audio by running a per-input optimization loop at inference, whereas CLIP-based sound guidance methods [4] require fine-tuning backbone layers; both impose significant computational cost or model coupling. Our approach instead freezes both backbones

entirely and trains only a single $1280 \to 4096$ linear layer ($\approx$5.2M parameters), reducing training to under three days on one A100 and making inference a single deterministic forward pass. The key enabling insight—that linear projection preserves the geometry of a pretrained encoder's latent space, allowing affective directions to survive intact into the decoder's conditioning space—is the principled mechanism that distinguishes our method from prior adaptation strategies. Table I summarizes how our approach differs from each prior strategy on the dimensions that matter for deployment.

III. Method

A. System Overview

Fig. 1 illustrates the complete pipeline. Raw speech audio is preprocessed into a log-mel spectrogram by Qwen2AudioProcessor, encoded into a variable-length hidden state sequence by the frozen Qwen2AudioEncoder, and then projected by the single trainable linear layer into the T5 conditioning slot of the frozen SD 3.5 Medium MM-DiT. Both CLIP timestep-embedding slots receive zero vectors. The MM-DiT denoises the image latent over 28 steps, and the predicted latent is decoded to pixels by the frozen VAE decoder. The only component that receives gradient updates during training is the linear projection layer.

B. Qwen2AudioEncoder as Frozen Speech Backbone

Qwen2AudioEncoder is the audio encoder component of the Qwen2-Audio 7B model [6]. Because it is bundled within the Qwen2-Audio-7B-Instruct checkpoint, extraction proceeds as follows: the complete 7B model is downloaded once ($\approx$15 GB), the `model.audio_tower` component (a `Qwen2AudioEncoder` instance) is saved with `save_pretrained` to a standalone checkpoint ($\approx$1.5 GB), and the LLM component is discarded entirely.

During preprocessing, raw audio at any sample rate is resampled to 16 kHz by `Qwen2AudioProcessor`, which then computes a 128-channel log-mel spectrogram with the windowing and normalization expected by the encoder. The processor is not trainable.

The encoder accepts the mel spectrogram and produces hidden states of dimension 1280 per temporal frame. For a batch of B clips with stride-2 temporal pooling, the output shape is $\mathbf{H} \in \mathbb{R}^{B \times T \times 1280}$ where T varies with clip duration. The encoder is trained on large-scale audio-language alignment across speech, environmental sound, and music, yielding semantically rich, linguistically organized embeddings that simultaneously encode lexical content and acoustic-affective variation between utterances. The encoder is frozen throughout all training stages; no gradients flow through it.

C. Linear Projection: The Sole Trainable Component

The sole trainable component is a linear layer without bias normalization:

$$\mathbf{Z} = \mathbf{H}\mathbf{W}^\top + \mathbf{b}, \tag{1}$$

where $\mathbf{H} \in \mathbb{R}^{B \times T \times 1280}$ is the encoder output, $\mathbf{W} \in \mathbb{R}^{4096 \times 1280}$ is the weight matrix, $\mathbf{b} \in \mathbb{R}^{4096}$ is the bias, and $\mathbf{Z} \in \mathbb{R}^{B \times T \times 4096}$ is the projected conditioning sequence. Total trainable parameters: $1280 \times 4096 + 4096 = 5{,}246{,}976$ ($\approx$5.2M).

A single linear layer is preferred to a multi-layer adapter because it preserves the linear geometry of the encoder output space: semantic concepts encoded as directions in $\mathbb{R}^{1280}$ are carried intact into $\mathbb{R}^{4096}$ rather than being warped by nonlinear transformations [7], while simultaneously minimizing trainable parameters and keeping the projection interpretable as a single change-of-basis between encoder and decoder spaces.

D. SD 3.5 Medium MM-DiT Conditioning

Stable Diffusion 3.5 Medium [2] uses an MM-DiT backbone that applies bidirectional joint attention over image patch tokens and text conditioning tokens within the same transformer blocks. Its conditioning pipeline comprises three streams:

1) **T5-XXL tokens** (4096d): A variable-length sequence of dense token embeddings produced by a T5-XXL encoder. This is the primary semantic conditioning stream and is where our projected audio tokens are injected.
2) **CLIP ViT-L pooled** (768d): A single pooled embedding added to the timestep embedding, providing global semantic grounding.
3) **CLIP ViT-bigG pooled** (1280d): A single pooled embedding also added to the timestep embedding, providing complementary global semantic context.

In our system, the projected audio sequence $\mathbf{Z} \in \mathbb{R}^{B \times T \times 4096}$ replaces the T5-XXL token sequence in stream (1). Both CLIP pooled slots (streams 2 and 3) receive zero vectors of the appropriate dimension during both training and inference. Zeroing the CLIP slots removes the global semantic grounding that these embeddings ordinarily provide to the timestep embedding, and this constitutes the primary quality ceiling of the current architecture. The auxiliary cosine similarity loss (Section III-E) partially compensates by pulling the mean-pooled audio representation toward the CLIP image embedding space, providing an implicit surrogate for the missing CLIP signal.

E. Training Objective

Training proceeds in two stages.

Stage 1: Projection Warmup. The linear projection layer alone is trained with an auxiliary cosine similarity loss:

$$\mathcal{L}_{\text{aux}} = 1 - \cos\big(\bar{\mathbf{h}},\ \mathbf{e}_{\text{img}}\big), \tag{2}$$

where $\bar{\mathbf{h}} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{h}_t \in \mathbb{R}^{1280}$ is the mean-pooled Qwen2AudioEncoder output *before* the linear projection, and $\mathbf{e}_{\text{img}} \in \mathbb{R}^{1280}$ is the CLIP ViT-bigG embedding of the corresponding target image. These dimensions match natively, requiring no additional projection. The loss encourages the encoder's representation to align with the visual semantic space using only audio and image supervision—no text and no transcription. The voice-cloned synthetic clips per image (Section IV-A) span 5 reference speakers and 5 emotional registers, providing diverse audio inputs that should all map near the same image embedding and giving the loss a natural

TABLE I: Comparison of speech/audio-to-image conditioning strategies. Trainable params: parameters updated during training. Inf. cost: forward passes per generation. Text req.: whether transcription is required. BB. FT: whether either backbone is fine-tuned. Prosody: whether suprasegmental acoustic content reaches the decoder.

Method	Trainable params	Inf. cost	Text required	Backbone fine-tune	Prosody preserved
ASR-bridge (Whisper→SD)	0	1×+ASR	yes	no	no
AudioToken [3]	adapter	per-input opt.	no	no	partial
CLAP-cond. [10]	linear	1×	no	yes	weak
Ours	**5.2M**	**1×**	**no**	**no**	**yes**

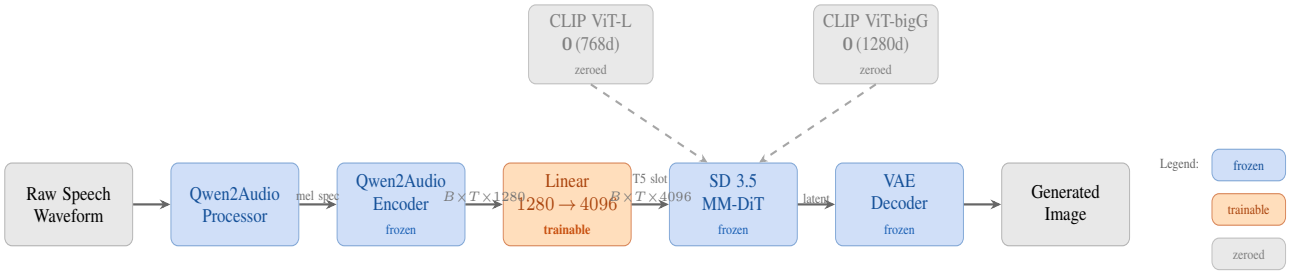


Fig. 1: System architecture. The frozen Qwen2AudioEncoder produces a sequence of 1280-dimensional hidden states from the 128-channel log-mel spectrogram. A single linear layer (the *only* trainable component, ≈5.2M parameters) projects these into the 4096-dimensional T5 conditioning slot of the frozen SD 3.5 Medium MM-DiT. Both CLIP timestep slots receive zero vectors throughout training and inference. The VAE decoder remains frozen. Only the shaded orange component receives gradient updates.

contrastive structure across both speaker timbre and acoustic register.

Stage 1 runs for 5 epochs with batch size 32, AdamW optimizer (β_1=0.9, β_2=0.999, weight decay 0.01), learning rate 10^{-3} with cosine decay and a 500-step linear warmup, and gradient clipping at 1.0. All training uses bf16 precision on a single NVIDIA A100 80 GB GPU.

Stage 2: Joint Training. The denoising and auxiliary objectives are combined:

$$\mathcal{L} = \mathcal{L}_{\text{denoise}} + \lambda \cdot \mathcal{L}_{\text{aux}}, \tag{3}$$

where $\mathcal{L}_{\text{denoise}}$ is the rectified flow denoising objective [2]: the target image is encoded to a latent $\mathbf{z}$ by the frozen VAE encoder, a noise schedule samples a timestep t, the frozen MM-DiT predicts the denoised latent conditioned on $\mathbf{Z}$, and the loss is the mean squared error between the predicted and actual velocity field. We set λ=0.1, which stabilizes the balance between the two loss terms throughout Stage 2 training.

Stage 2 runs for 20 epochs with batch size 16, learning rate 10^{-4} with cosine decay and a 500-step linear warmup. During Stage 2, the audio conditioning tokens are dropped with probability p=0.1 (replaced with zero sequences) to enable classifier-free guidance (CFG) at inference. At inference, CFG is applied with guidance scale 7.0. Stage 1 completes in approximately 18 hours; Stage 2 in approximately 52 hours on the same A100.

F. Sequence Length Budget

Synthesized TTS clips in our pipeline range from approximately 5 to 15 seconds in duration. With stride-2 temporal pooling in Qwen2AudioEncoder, each clip yields approximately 125–375 tokens; SD 3.5 Medium is designed for T5 conditioning sequences up to 256 tokens. We clip audio at 15 seconds and front-truncate any resulting token sequence longer than 256 tokens, since speech front-loads principal nouns and key descriptors. At 1024 × 1024 resolution the MM-DiT processes approximately 4096 image patch tokens; an additional 256 audio conditioning tokens raise the attention sequence length by only ≈6%, comfortably within the A100's 80 GB memory budget.

IV. Experiments

A. Dataset

We construct a synthetic emotional audio–image dataset on top of MS-COCO. For each COCO image, the original caption is first expanded by a large language model (Claude) into a richer descriptive paragraph. For example, the COCO-style caption *"A lake/pond with an orange item floating on it"* is expanded to *"A tranquil lake stretches across the scene, its calm, glassy surface mirroring the muted tones of the surrounding landscape, while a single vibrant orange object drifts gently upon the water, drawing the eye instantly with its bold color as it floats in peaceful solitude, swaying ever so slightly with the subtle, unhurried rhythm of the ripples that radiate outward around it, making it the quiet yet undeniable focal point of an otherwise still and serene natural setting"*; this expanded paragraph is what the TTS reads aloud. Emotional voice references are drawn from the RAVDESS [11] corpus: 5 actors × 5 emotion classes (neutral, happy, sad, angry, fearful) × 2 reference clips per (actor, emotion), yielding 50 voice-clone references. An open-source voice-cloning TTS, OmniVoice [12], then synthesizes the expanded caption text

in each cloned voice, producing speech that carries the lexical content of the expanded caption with the speaker timbre and emotional prosody of the corresponding RAVDESS reference. Each synthesized clip is paired with its source COCO image, yielding approximately 10,000 (audio, image) training pairs and 2,000 held-out evaluation pairs spanning 5 reference speakers and 5 emotional registers per image. Inference uses the same pipeline with no transcription at any point.

We partition the dataset at the image level (90% train, 10% held-out evaluation). All audio clips synthesized from a given image are assigned to the same partition to prevent leakage.

B. Baselines

We compare against four baselines that span the space of conditioning strategies:

(a) Null-audio. The linear projection output is set to zero (all-zero T5 conditioning tokens), with zero CLIP slots. This is equivalent to unconditional SD 3.5 generation and establishes the frozen decoder's quality floor with no conditioning signal whatsoever.

(b) Mean-pooled audio. The Qwen2AudioEncoder output is mean-pooled to a single 1280d vector, projected by a learned $1280 \to 4096$ linear layer, and used as a single conditioning token in the T5 slot. This isolates the contribution of sequential temporal structure in the conditioning.

(c) CLAP-conditioned. The CLAP audio encoder [10], pretrained on large-scale contrastive audio-language alignment, encodes each audio clip to a global embedding. A learned linear layer maps this to the 4096d T5 slot as a single token. CLAP provides audio-language-aligned global conditioning but discards all temporal structure.

(d) ASR-bridge (Whisper $\to$ SD 3.5). Audio is transcribed by Whisper-large-v3 [5] and the transcript is fed to SD 3.5 Medium through its standard T5-XXL text encoder with populated CLIP slots. This is the current best-practice approach and a strong topline for semantic fidelity; however, it discards all suprasegmental acoustic information during transcription.

C. Evaluation Metrics

FID. We compute FID between 10,000 images generated from held-out evaluation audio clips and their paired ground-truth COCO images using InceptionV3 features. Lower is better. Absolute FID values above 100 are typical for non-photorealistic outputs on COCO-scale evaluations and are most meaningful for relative ordering between methods evaluated on the same reference set, as in Table II.

Audio-image retrieval (R@1, R@5). For each audio clip in the evaluation set we generate one image. We then compute cosine similarity between the mean-pooled Qwen2AudioEncoder embedding of the source audio (1280d) and the CLIP ViT-bigG embedding of the generated image, ranking the target generated image against 99 randomly sampled generated images from other audio clips (100 candidates total). R@1 (R@5) is the fraction of queries for which the target generated image ranks first (in the top five); higher is better. This text-free metric measures whether generated images are semantically aligned with their source audio, collapsing to random chance (1%, 5%) for an unconditional generator and approaching 100% as audio–image semantic alignment becomes perfect.

Human preference. We recruit 50 annotators in a controlled evaluation. Each annotator rates a randomly assigned subset of 200 generated images on two dimensions: *semantic match* (does the image depict what was described?) and *emotional resonance* (does the image's mood or style reflect the emotional tone of the speech?) on a 5-point Likert scale, where 1 indicates "not at all" and 5 indicates "very strongly." This provides a human-centric assessment of the dimensions where speech- and text-based conditioning differ most. Statistical significance for between-method comparisons is reported via paired t-tests over per-rater per-image scores.

D. Results

Because each speech prompt in our pipeline is an LLM-expanded paragraph rather than the original COCO caption (Section IV-A), the generated image is not expected to be a pixel-level reconstruction of any single COCO photo; quantitative comparison is therefore reported as distributional FID over 10,000 held-out pairs, plus audio–image retrieval and human ratings, rather than per-image side-by-side fidelity.

Table II presents the quantitative results. Our full system (Stage 2, combined loss) achieves FID 122.4, R@1 of 27.6%, and R@5 of 60.8%. The ASR-bridge achieves slightly better FID 116.8 and R@1 29.4%, reflecting the advantage of an explicit transcription with populated CLIP slots. Nonetheless, our method substantially narrows the gap relative to all other baselines, while operating with no text input at any stage.

The mean-pooled audio baseline (FID 153.7) demonstrates that sequential temporal structure in the conditioning tokens is important: discarding this structure in favor of a single mean-pooled vector raises FID by 31.3 points. The CLAP-conditioned baseline (FID 136.4) benefits from audio-language pretraining but is limited by its compressed global representation; our full model improves over CLAP by 14.0 FID points.

The ablation rows in Table II isolate the contribution of each training stage. Stage 1 alone (cosine loss only, no denoising) achieves FID 142.1 and R@1 18.4%—already competitive with mean-pooled audio, confirming that the cosine auxiliary loss provides useful conditioning signal on its own. Removing the auxiliary loss entirely (denoising from Stage 1 initialization, no $\mathcal{L}_{\text{aux}}$) yields FID 132.7; adding the auxiliary loss in Stage 2 improves this to FID 122.4, a gain of 10.3 points.

Table III presents the human preference study. Our method scores 3.7 on semantic match (vs. 3.9 for the ASR-bridge), a small gap consistent with the modest FID difference. On emotional resonance, our method achieves 3.8 versus 2.9 for the ASR-bridge (paired t-test, $p<0.001$) and 2.6 for CLAP. This confirms the central claim: the speech signal carries genuine affective content that survives linear projection into the diffusion conditioning space and perceptibly modulates the visual output.

TABLE II: Quantitative Results on the Held-Out Audio–Image Split (10%)

Method	FID↓	R@1↑	R@5↑
Null-audio (unconditional)	191.3	1.0	5.0
Mean-pooled audio	153.7	15.2	40.1
CLAP-conditioned [10]	136.4	23.8	54.3
ASR-bridge (Whisper [5]→SD 3.5)	**116.8**	**29.4**	**63.2**
Ours – Stage 1 only (ablation)	142.1	18.4	46.7
Ours – no auxiliary loss (ablation)	132.7	22.1	53.5
Ours – Stage 2, full	122.4	27.6	60.8

TABLE III: Human Preference Evaluation (Likert 1–5, N=50 Raters, 200 Samples Each)

Method	Semantic Match↑	Emotional Resonance↑
CLAP-conditioned	3.2	2.6
ASR-bridge	**3.9**	2.9
Ours (Stage 2, full)	3.7	**3.8****

$^{**}p < 0.001$ vs. ASR-bridge (paired t-test).

E. Qualitative Examples

Fig. 2 shows generated images for a single COCO scene—"a lake with a vivid orange object floating on its surface"—under four contrasting emotional deliveries. Lexical content is held constant across all four panels; only the acoustic register differs, and no text is fed to the decoder at any stage.

Neutral delivery (Fig. 2a). Balanced midday lighting, natural greens and blues, the orange object centered in clear, calm water. Acts as the baseline reference register.

Happy delivery (Fig. 2b). Warmer color temperature, increased saturation, golden-hour highlights on the water surface, and a more open, inviting composition.

Sad delivery (Fig. 2c). Cooler palette, lowered exposure, overcast atmosphere; the orange object retains its identity but reads as muted and isolated against a desaturated environment.

Fearful delivery (Fig. 2d). Darker, higher-contrast scene with more agitated water, narrower depth of field, and ominous lighting; the orange object is preserved but the mood is turbulent rather than serene.

Across all four panels the subject identity and the spatial layout are preserved; what changes is global chromatic choice, lighting, and surface detail, supporting the central claim that affective signal in the audio survives linear projection into MM-DiT's conditioning space and modulates the visual output in coherent, perceptible ways.

V. Analysis and Discussion

A. Stage 1 Warmup vs. Cold-Start Training

As shown in Table II, Stage 1-only training (cosine loss, no denoising) achieves FID 142.1. Cold-start Stage 2 (denoising + auxiliary loss, without Stage 1 warmup) achieves FID 132.7. Our full two-stage pipeline achieves FID 122.4, a 10.3-point improvement over cold-start.

During Stage 1 the validation cosine similarity between the mean-pooled Qwen2AudioEncoder output and the CLIP ViT-bigG image embedding rises from $\approx$0.05 at random initialization to $\approx$0.41 by epoch 4, plateauing thereafter. This confirms that the warmup establishes a stable, semantically meaningful initialization in the decoder's conditioning space before the noisier denoising gradient takes over, significantly reducing the number of Stage 2 steps required to reach a given FID.

B. CLIP Slot Zeroing

Setting both CLIP timestep slots to zero is the primary architectural compromise: the MM-DiT must infer global image semantics entirely from the projected audio sequence in the T5 slot, a conditioning mode the decoder was not pretrained on. An oracle evaluation that supplies the *ground-truth* CLIP ViT-bigG image embedding while retaining our projected audio tokens achieves FID 97.3, quantifying the $\approx$25-point FID ceiling imposed by zeroing. A natural extension—a second linear projection from the mean-pooled 1280d audio embedding into the natively dimension-matched CLIP bigG slot—should recover a significant fraction of this gap without adding any new component types.

C. Emotional Modulation

To quantify emotional modulation beyond the human evaluation, we construct a paired image-level experiment. For each of 500 held-out images we generate one positive-valence (happy) and one negative-valence (sad or fearful) audio variant and measure variance between the resulting image pairs in three photometric features: mean brightness, mean saturation, and dominant color temperature. Our full model shows statistically significant variance (σ^2=18.4, 22.7, 31.1 respectively; paired t-test, p<0.01 for all three), while the ASR-bridge baseline receives identical transcribed text for the two valences and shows near-zero variance (σ^2<1.2 in all three features). This confirms that acoustic-affective variation in Qwen2AudioEncoder's hidden states propagates through the linear projection and influences the MM-DiT's output distribution in perceptible, systematic ways.

D. Limitations

CLIP slot zeroing quality ceiling. As shown by the oracle experiment, zeroing both CLIP slots leaves approximately 25 FID points on the table. This is the most impactful known limitation and is directly addressable in a straightforward extension.

Synthetic-TTS training distribution. Audio is produced by a voice-cloning TTS conditioned on RAVDESS reference clips (Section IV-A) rather than by recording fresh human speech. The acoustic distribution of voice-cloned speech differs from natural emotional speech in several known ways—reduced microprosody, narrower speaker variability beyond the 5 reference actors, and discrete emotion-label conditioning rather than continuous affective gradation. Generalization to unconstrained recorded human emotional speech remains untested. Expanding to recorded human-speech corpora and to a larger reference-speaker pool would address this.

No explicit emotion supervision. Modulation arises implicitly from acoustic variation between voice-clone reference



(a) Neutral delivery (b) Happy delivery (c) Sad delivery (d) Fearful delivery

Fig. 2: Emotion-conditioned generation for a single scene synthesized from speech reading the LLM-expanded form of the COCO caption *"A lake/pond with an orange item floating on it"* (full expansion in Section IV-A). The four panels share identical lexical content; only the acoustic delivery differs. Subject identity (the body of water and the orange object) and overall composition are preserved across all four panels; visual mood, palette, and lighting shift systematically with the emotional register of the speaker.

speakers and per-emotion RAVDESS conditioning rather than a designed objective; the system can reflect affect present in the input audio but cannot be directed to increase or decrease emotional intensity without explicit emotion-labeled fine-tuning.

VI. Conclusion

We have demonstrated that a single linear layer is sufficient to condition a frozen state-of-the-art diffusion image decoder on raw speech audio. Projecting the output of a frozen Qwen2AudioEncoder into the T5 conditioning slot of a frozen SD 3.5 Medium MM-DiT achieves FID 122.4 on a held-out audio–image split—competitive with a Whisper-based ASR bridge (FID 116.8) that requires transcription and discards all acoustic-affective content—while significantly outperforming all baselines on human-rated emotional resonance (3.8 vs. 2.9 on a 5-point Likert scale). The total trainable parameter count is approximately 5.2M, trained on a single NVIDIA A100 80 GB GPU in under three days across both stages.

The results validate three central hypotheses: a large-scale speech encoder's hidden states are linearly alignable with a diffusion decoder's conditioning space; a cosine auxiliary warmup yields meaningfully better final FID than cold-start joint training; and acoustic-affective features survive linear projection to produce perceptible, systematic variation in generated image style and mood. Promising extensions include filling the CLIP timestep slots via a second linear projection of the mean-pooled audio embedding, joint training with ASR decoding, and fine-tuning on emotion-labeled speech corpora.

Acknowledgment

The authors thank the MS-COCO team for the image dataset, the RAVDESS authors for the emotional speech corpus, the k2-fsa team for releasing OmniVoice, and the Qwen team at Alibaba Group for releasing the Qwen2-Audio model weights.

References

[1] J. Betker, G. Goh, L. Jing, T. Brooks, J. Wang, L. Li, L. Ouyang, J. Zhuang, J. Lee, Y. Guo, W. Manassra, P. Dhariwal, C. Chu, Y. Jiao, and A. Ramesh, "Improving image generation with better captions," OpenAI, Tech. Rep., 2023.

[2] P. Esser, S. Kulal, A. Blattmann, R. Entezari, J. Müller, H. Saini, Y. Levi, D. Lorenz, A. Sauer, F. Boesel, D. Podell, T. Dockhorn, Z. English, K. Lacey, A. Goodwin, Y. Marek, and R. Rombach, "Scaling rectified flow transformers for high-resolution image synthesis," in *Proceedings of the 41st International Conference on Machine Learning*, 2024.

[3] G. Yariv, I. Gat, L. Wolf, Y. Adi, and I. Schwartz, "AudioToken: Adaptation of text-conditioned diffusion models for audio-to-image generation," in *Proc. Interspeech 2023*, 2023, pp. 5446–5450.

[4] S. H. Lee, W. Roh, W. Byeon, S. H. Yoon, C. Y. Kim, J. Kim, and S. Kim, "Sound-guided semantic image manipulation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 3377–3386.

[5] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust speech recognition via large-scale weak supervision," in *Proceedings of the 40th International Conference on Machine Learning*, vol. 202. PMLR, 2023, pp. 28 492–28 518.

[6] Y. Chu, J. Xu, Q. Yang, H. Wei, X. Wei, Z. Guo, Y. Leng, Y. Lv, J. He, J. Lin, C. Zhou, and J. Zhou, "Qwen2-Audio technical report," *arXiv preprint arXiv:2407.10759*, 2024.

[7] H. Liu, C. Li, Q. Wu, and Y. J. Lee, "Visual instruction tuning," in *Advances in Neural Information Processing Systems*, vol. 36, 2023.

[8] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 10 684–10 695.

[9] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever, "Learning transferable visual models from natural language supervision," in *Proceedings of the 38th International Conference on Machine Learning*. PMLR, 2021.

[10] Y. Wu, K. Chen, T. Zhang, Y. Hui, M. Nezhurina, T. Berg-Kirkpatrick, and S. Dubnov, "Large-scale contrastive language-audio pretraining with feature fusion and keyword-to-caption augmentation," in *ICASSP 2023 – IEEE International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2023, pp. 1–5.

[11] S. R. Livingstone and F. A. Russo, "The Ryerson audio-visual database of emotional speech and song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English," *PLoS ONE*, vol. 13, no. 5, p. e0196391, 2018.

[12] k2-fsa, "OmniVoice: Open-source voice-cloning TTS," https://huggingface.co/k2-fsa/OmniVoice, 2024, accessed 2026-04-28.

Монгол гэрийн хаалганы текстур үүсгэх арга

Өлзий-Ундралын Оргил
Машин оюуны лаборатори
Мэдээлэл, компьютерын ухааны тэнхим
МТЭС, МУИС
25d1num0110@stud.num.edu.mn

Ганболдын Амарсанаа
Машин оюуны лаборатори
Мэдээлэл, компьютерын ухааны тэнхим
МТЭС, МУИС
amarsanaag@num.edu.mn

***Хураангуй*—Монгол гэрийн архитектурт хаалга нь хийцлэлийн хувьд босоо болон хөндлөн хээний нарийн зүй тогтол, үндэсний бэлгэдэл бүхий хээ угалзын нэгдэл байдаг. Энэхүү судалгааны ажилд Монгол гэрийн хаалганы олон төрлийн архитектурын хувилбаруудыг модулиар параметрчилсэн процедур аргаар бодит цаг хугацаанд үүсгэх алгоритмыг санал болгож байна. Уламжлалт статик текстур зураг ашиглах арга нь санах ойн ачааллыг нэмэгдүүлж, бүтцийн хувьд өөрчлөх боломжгүй байдаг бол бидний боловсруулсан шийдэл нь хаалганы хээний зохиомжийг дөрвөн дүрмийн системд шилжүүлж, хаалганы хатавчны зохиомжийг гурван дүрмээс хамаарч бүтцийг динамикаар тооцоолдог.**

Алгоритм нь үндэсний ”Өлзий” болон ”Булан” хээг математик координатын жагсаалтаар тодорхойлж, пикселийн түвшинд алгоритмаар зурдаг бөгөөд нэгж сегментийн t параметрээр дамжуулан текстурын нарийвчлалыг удирдах боломжийг бүрдүүлсэн. Туршилтын үр дүнд уг арга нь хаалганы олон төрлийн зохиомжийг дунджаар 100 мс-ийн дотор үүсгэж байгаа бөгөөд санах ойн ачааллыг эрс багасган, ямар ч хэмжээтэй текстурыг чанарын алдагдалгүй гаргаж авах боломжтойг баталсан. Энэхүү шийдэл нь соёлын биет өвийг дижитал хэлбэрт хадгалах, виртуал орчинд Монгол ахуйг архитектурын дүрмийн дагуу уян хатан дүрслэхэд техникийн болон соёлын стратегийн ач холбогдолтой юм.

***Түлхүүр үгс*—Процедур арга, Монгол гэр, Хаалга, Дижитал өв**

I. Удиртгал

Монгол гэрийн хаалга нь нүүдэлчдийн орон сууцны архитектурын гол цөм, соёл гоо зүйн үнэ цэнийг илэрхийлэгч элемент юм. Уламжлалт Монгол гэрийн хаалга нь зөвхөн хээ угалзаар чимэглэгдээд зогсохгүй, хийцлэлийн хувьд босоо болон хөндлөн хээний тодорхой зүй тогтол, пропорцын дагуу урлагддаг [1]. Монгол гэрийн хаалганы бэлгэдэл болон хийцлэлийн уламжлалт хэв шинж нь орон зайн онцлог болон гэр бүлийн статусаас хамааран өөр өөр байдаг нь нүүдэлчин ард түмний материаллаг соёлын нэгэн өвөрмөц илэрхийлэл болдог [2]. Орчин үеийн виртуал орчин, компьютер тоглоом болон дижитал музейн салбарт соёлын өвийг дүрслэхдээ хаалганы энэхүү олон янз байдлыг гаргах шаардлага тулгардаг боловч үүнийг уламжлалт аргаар шийдвэрлэх нь нөөц бололцооны хувьд хүндрэлтэй байдаг [3].

Уламжлалт 3D загварчлалд хаалганы дүр төрхийг бий болгохдоо бэлэн текстур зураг ашигладаг нь хэд хэдэн хязгаарлалтыг бий болгодог. Нэгдүгээрт, хаалганы бүтцийн өөр өөр хувилбаруудыг дүрслэхийн тулд олон тооны зураг бэлтгэх шаардлагатай болдог нь хадгалах сан болон видео санах ойд ихээхэн ачаалал өгдөг [4]. Хоёрдугаарт, бэлэн зургууд нь нарийвчлалын хувьд тогтмол байдаг тул моделийг ойроос харах эсвэл масштабыг өөрчлөхөд пикселийн суналт буюу чанарын алдагдал гардаг [5]. Гуравдугаарт, Монгол гэрийн хаалганы архитектурын ”дүрэм” буюу хээний харьцаа, хээний байрлал зэрэг нь кодоор бус статик байдлаар хадгалагдаж байна.

Энэхүү судалгааны ажлаар дээрх асуудлуудыг шийдвэрлэх үүднээс Монгол гэрийн хаалганы текстурыг ”модулиар параметрчилсэн” процедур аргын тусламжтайгаар үүсгэх алгоритмыг санал болгож байна. Бидний боловсруулсан шийдлийн гол онцлог, шинэлэг тал нь:

- **Архитектурын дүрэмд суурилсан бүтэц:** Хаалганы хээний зохиомжийг дөрвөн төрлийн дүрмийн системд шилжүүлж, мөр болон баганы хээний тооцооллыг динамик болгосон. Энэ нь Мюллерийн санал болгосон *CGA shape grammar* алгоритмын зарчмыг Монгол гэрийн хаалганы онцлогт тохируулан хөгжүүлсэн хэрэг юм [6].
- **Модулиар хаалганы хатавч:** Хаалганы үндсэн хэсгээс гадна хатавчны бүтцийг нэг хээтэй, хоёр хээтэй эсвэл хээгүй бүтэцтэй байхаар алгоритмчилж, хаалганы нийт өргөнийг сонголтоос хамааран агшин зуурт тооцоолдог.
- **Нарийвчлалын уян хатан байдал:** Нэгж сегментийн хэмжээг тодорхойлогч t параметр дээр суурилан текстурын нягтралыг хэрэглэгч өөрөө удирдах боломжтой болсон [7].
- **Математик хээ дүрслэл:** Үндэсний ”Өлзий” болон ”Булан” хээг цэгэн координатын жагсаалт хэлбэрээр хадгалж, түүнийг алгоритмын аргаар масштаблах, эргүүлэх замаар текстур дээр байрлуулдаг. Энэ нь вектор графикийн зарчмыг процедур аргаар үүсгэх нэгтгэсэн шинэлэг хандлага юм [8].

Ингэснээр Монгол гэрийн хаалганы олон зуун жилийн турш боловсронгуй болсон хийцлэлийн дүрмийг кодын хэлэнд шилжүүлж, өндөр гүйцэтгэлтэй, уян хатан дижитал хувилбарыг үүсгэх техникийн шийдлийг санал болгож байна. Үйлдвэрлэлийн салбарт (CNC/CAD) ашиглагддаг

статик загварчлалын аргуудаас ялгаатай нь бидний санал болгож буй алгоритм нь бодит цагийн график хөдөлгүүр дээр маш бага санах ой ашиглан, хязгааргүй олон хувилбарыг шууд үүсгэх зориулалттай. Энэ нь зөвхөн нэг загвар гаргах бус, виртуал орчинд соёлын өвийг уян хатан байдлаар дүрслэхэд чиглэгдсэнээрээ шинэлэг юм.

II. Холбогдох судалгаанууд

Монгол гэрийн хаалганы текстурыг процедур аргаар үүсгэх асуудал нь архитектурын дүрэмд суурилсан загварчлал, соёлын өвийн дижитал хадгалалт болон алгоритмын аргаар текстур үүсгэх гэсэн үндсэн чиглэлүүдэд хамаарч байна.

A. Архитектурын дүрэмд суурилсан загварчлал

Компьютер графикийн салбарт барилга байгууламжийг алгоритмын аргаар үүсгэх судалгаанууд *Shape Grammar* буюу дүрмийн систем дээр үндэслэн хөгжиж ирсэн. Париш болон Мюллер нар хотын гудамж, барилгын бүтцийг шаталсан дүрмээр үүсгэх суурь боломжийг бүрдүүлсэн [6]. Энэхүү хандлага нь архитектурын элементүүдийг статик бус, харин математик логикоор удирдах боломжийг олгосон юм. Бидний санал болгож буй арга нь энэхүү дүрмийн загварчлалын зарчмыг Монгол гэрийн хаалганы хээний зохиомжид тусгайлан зориулж хөгжүүлсэн шийдэл юм.

B. Соёлын өвийн дижитал сэргээн босголт

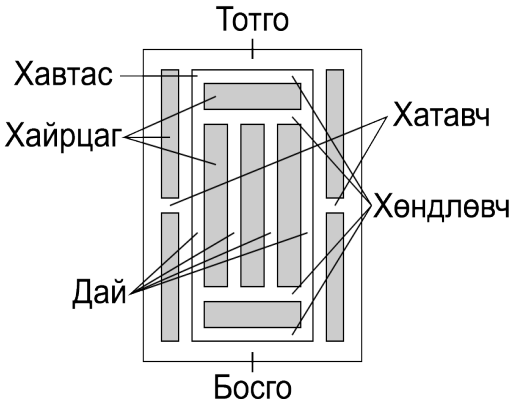
Түүхэн дурсгалт газрууд болон соёлын биет өвийг 3D хэлбэрт оруулахдаа ихэвчлэн фотограмметри болон лазер сканнерын аргыг ашигладаг [3]. Хэдийгээр эдгээр аргууд нь бодит байдлыг өндөр нарийвчлалтай дүрсэлдэг ч моделийн бүтцийн логикгүй байдаг тул параметрээр удирдах боломжгүй статик шинжтэй байдаг [5]. Монгол гэрийн архитектурыг судлаач Майдар болон Хамфри нарын бүтээлд гэрийн бүтэц нь геометрийн нарийн хамаарал бүхий систем болохыг онцолсон байдаг [1], [2]. Иймд уламжлалт архитектурыг процедур аргаар загварчлах нь соёлын өвийг зөвхөн үзүүлэн төдий бус, бүтцийн логиктой нь хамт дижитал хэлбэрт хадгалах хамгийн оновчтой арга юм.

C. Алгоритмын аргаар текстур үүсгэх

3D моделийн бодит байдлыг дүрслэхэд текстур чухал үүрэгтэй боловч системийн нөөц хязгаарлагдмал үед өндөр нарийвчлалтай зураг ашиглах нь хүндрэлтэй байдаг [4]. Процедур аргаар үүсгэх нь математик функц ашиглан пикселийг тооцоолж, санах ойн ачааллыг оновчилдог [7]. Сүүлийн үеийн судалгаанууд текстурын нарийвчлалыг масштабаас үл хамааран чанартай хадгалах асуудалд төвлөрч байна [8]. Бидний ажилд ашигласан ”Өлзий” болон ”Булан” хээг координатаар дүрслэн зурах арга нь вектор графикийн уян хатан байдлыг процедур аргатай нэгтгэж, системийн ачааллыг хамгийн бага түвшинд барьж байгаагаараа онцлог юм.

III. Арга зүй

Монгол гэрийн хаалганы текстурыг процедур аргаар үүсгэх нь параметрчилсэн геометрийн тооцоолол, дүрмийн системд суурилсан зохиомж болон вектор хээний хувиргалт гэсэн шат дараалаар хэрэгжинэ. Зураг 1-ээс хаалганы бүрдэл хэсгүүдийн нэрсийг харж болно.



Зураг 1: Хаалганы бүрдэл хэсгүүд ба тэдгээрийн нэрс.

A. Текстур атласын хэмжээ ба динамик өргөн

Алгоритмын үндсэн масштаб нь нэгж сегментийн хэмжээг тодорхойлогч t параметр байна. Хаалганы хавтас буюу онгойх хэсгийг Зураг 2-ын дотор талын дөрвөлжингүүдээр харуулсан. Үүнд өргөнийг $W_{base} = 10 \cdot t$, өндрийг $H_{base} = 20 \cdot t$ хэмжээгээр тогтоосон бөгөөд нийт текстурын хэмжээг (W, H) дараах байдлаар тодорхойлно:

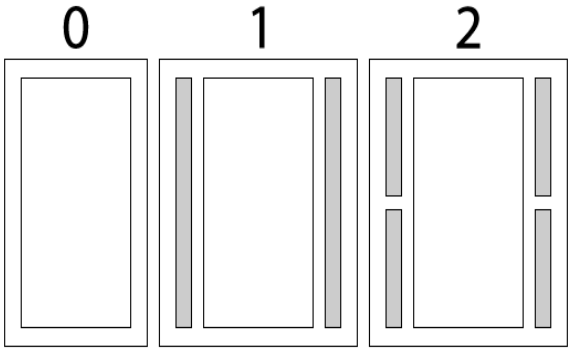
$$H = H_{base} + 2 \cdot H_{border} \tag{1}$$

$$W = W_{base} + 2 \cdot W_{side} \tag{2}$$

Үүнд хаалганы босго тотгоны өндрийг $H_{border} = 1.5 \cdot t$ байхаар харин хатавчны төрлөөс хамаарах өргөнийг W_{side}-аар илэрхийлсэн бөгөөд дараах нөхцөлт функцээр илэрхийлэгдэнэ:

$$W_{side} = \begin{cases} W_{border} & \text{if } side = 0 \\ W_{border} + W_{pattern} + W_{borderIn} & \text{if } side \in 1,\, 2 \end{cases} \tag{3}$$

Хаалганы хатавчны хувьд гадна захын дайны өргөнийг $W_{border} = 1.5 \cdot t$, дотор захын дайны өргөнийг $W_{borderIn} = t$, хээ зурагдах хэсэг буюу хайрцгийн өргөнийг $W_{pattern} = 1.5 \cdot t$ байхаар сонгосон. Хаалганы хатавчны төрлийг $side$ нь 0 үед хайрцаггүй харин 1 болон 2 үед харгалзах тооны хайрцагтай зурж байгааг Зураг 2-оос саарал өнгөөр будагдсан байгаа хэсгээс харж болно.

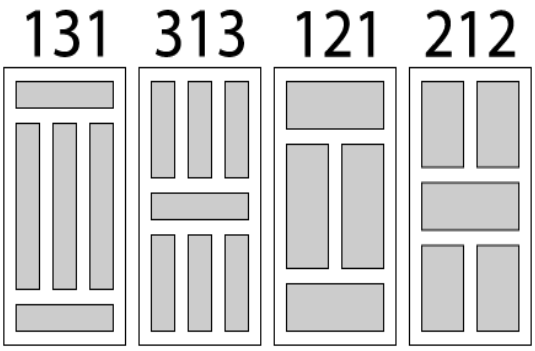


Зураг 2: Хаалганы хатавчны хайрцгийн төрлүүд нь нэг талдаа 0 үед хайрцаггүй, 1 үед нэг том хайрцагтай, 2 үед хоёр хуваагдсан жижиг хайрцагтай байна.

B. Зохиомжийн дүрмийн систем

Хаалганы хавтас дээрх хайрцгийн байрлал болон харьцааг R олонлогоор тодорхойлно. Энэхүү олонлог нь хаалганы төрөл бүрд харгалзах тэгш өнцөгтүүдийн (x, y, w, h) координатуудыг агуулна:

$$R_{type} = r_i \mid r_i = (x_i, y_i, w_i, h_i), i = 1, \ldots, n \tag{4}$$



Зураг 3: Хаалганы хавтасны хайрцгийн төрлүүд.

Зураг 3-аас хаалганы төрөл бүрд харгалзах хайрцгийн байрлалуудыг саарал өнгөөр дүрслэн байгааг харж болно. Хаалганы хавтас нь зохиомжийн хувьд хайрцгийн өгөгдөл нь дараах тогтмол утгуудтай байна:

$$R_{131} = \{(1, 1, 8, 2), (1, 4, 2, 12), (4, 4, 2, 12), (7, 4, 2, 12), (1, 17, 8, 2)\} \tag{5}$$

$$R_{313} = \{(1, 1, 2, 7), (4, 1, 2, 7), (7, 1, 2, 7), (1, 9, 8, 2), (1, 12, 2, 7), (4, 12, 2, 7), (7, 12, 2, 7)\} \tag{6}$$

$$R_{121} = \{(1, 1, 8, 3.5), (1, 5.5, 3.5, 9), (5.5, 5.5, 3.5, 9), (1, 15.5, 8, 3.5)\} \tag{7}$$

$$R_{212} = \{(1, 1, 3.5, 6.25), (5.5, 1, 3.5, 6.25), (1, 8.25, 8, 3.5), (1, 12.75, 3.5, 6.25), (5.5, 12.75, 3.5, 6.25)\} \tag{8}$$

Дээрх хайрцгуудын байрлал болон харьцааг t параметрээр үржүүлж эхлэх цэгийн координатыг нэмснээр текстур дээрх жинхэнэ байрлал гарна. Хаалганы хавтасны хайрцгийн хоорондох зай буюу хөндлөвч болон дайны хэмжээг нь 1 нэгж буюу t параметр зайтай. Цаашдаа хаалганы хавтасны хөндлөвч болон дай нь ижил хэмжээтэй учраас нэг нэрээр дуудах зорилгоор хүрээ гэж нэрлэе. Нэг хаалган доторх хайрцгуудын хувьд бүгдийг босоо болговол өргөний хувьд ижил хэмжээтэй, харин өндөр нь өөр өөр байж болно. Жишээ нь 131, 313 төрлүүдийн хувьд бүх босоо байвал хайрцгийн өргөн нь 2 байна. Харин 121, 212 төрлүүдийн хувьд өргөн нь 3.5 байна. Өргөнөөс хамааруулаад өндрийг нь тооцоолж болно.

Хайрцаг бүрийн төв цэг $C_i(x_c, y_c)$-ийг өлзий хээ байрлуулах зорилгоор дараах байдлаар тооцоолно:

$$C_i = \left(x_i + \frac{w_i}{2}, y_i + \frac{h_i}{2}\right) \cdot t \tag{9}$$

Харин булангийн хээг байрлуулахдаа хайрцаг нь босоо эсвэл хэвтээ байхаас хамааран босоо бол дээр болон доор нь харин хэвтээ бол зүүн болон баруун талд C_i цэгийг өөр өөрөөр тооцоолно.

C. Үндэсний хээний вектор загварчлал ба хувиргалт

”Өлзий” болон ”Булан” хээг P цэгэн координатын олонлогоор тодорхойлно. Энэхүү олонлог нь хээний төрөл бүрд харгалзах (x, y) координатуудыг агуулна:

$$P_{type} = (x_i, y_i) \mid i = 1, \ldots, n \tag{10}$$

Өлзий хээг босоо хэлбэртэй төвд нь координатын эх байхаар загварчилбал координатын цэгийн олонлог нь дараах утгуудтай байна:

$$P_{ulzii} = \{(0, -4), (-1, -3), (3, 1), (2, 2), (-2, -2), (-3, -1), (1, 3), (0, 4), (-1, 3), (3, -1), (2, -2), (-2, 2), (-3, 1), (1, -3), (0, -4)\} \tag{11}$$

Булангийн хээг хэвтээ хэлбэртэй бөгөөд хэвтээ тэнхлэгийн хувьд төвд, босоо тэнхлэгийн хувьд доор нь координатын эх байхаар загварчилбал координатын цэгийн олонлог нь дараах утгуудтай байна:

$$P_{bulan} = \{(-4, 6), (-4, 2), (-1, 2), (-1, 1), (-4, 1), (-4, 0), (-3, 0), (-3, 3), (-2, 3), (-2, 0), (2, 0), (2, 3), (3, 3), (3, 0), (4, 0), (4, 1), (1, 1), (1, 2), (4, 2), (4, 6)\} \tag{12}$$

Өлзий хээг зурах функцэд эдгээр суурь цэгүүдийг бодит пикселийн координатад шилжүүлэхдээ масштабын

коэффициент (s_x, s_y) болон t параметрийг ашиглан шугаман хувиргалт хийнэ:

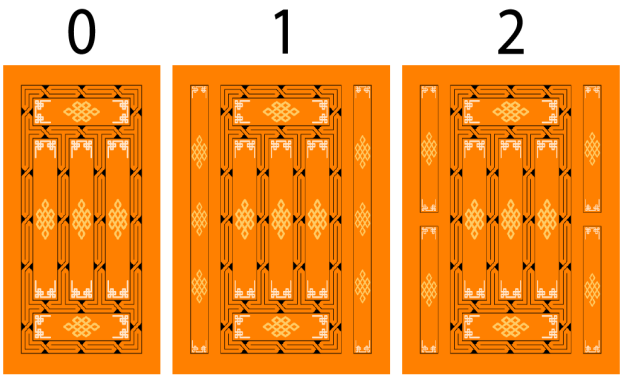
$$\begin{bmatrix} x_i' \\ y_i' \end{bmatrix} = \begin{bmatrix} x_c \\ y_c \end{bmatrix} + \begin{bmatrix} s_x \cdot t & 0 \\ 0 & s_y \cdot t \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix} \quad (13)$$

Үүнд s_x, s_y нь хайрцгийн харьцаанаас хамаарч хээг эргүүлэх эсвэл масштабыг тааруулах утгуудыг авна. Харин булангийн хээний хувьд x_c, y_c цэгийн утгыг харгалзах утгаар сольж тооцоолно.

D. Текстур үүсгэх алгоритм

Текстурыг бодит цаг хугацаанд үүсгэх процесс нь өмнөх бүлгүүдэд тодорхойлсон геометрийн өгөгдлүүдийг пикселийн түвшинд давхарлан зурах замаар хэрэгжинэ.

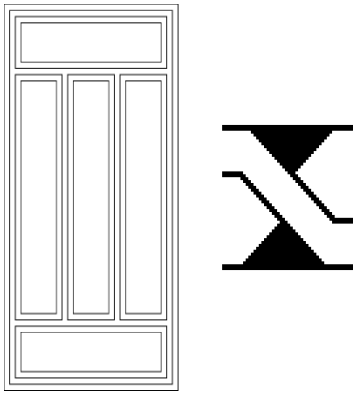
1) Хээ зурах болон хаалганы хатавч: Эхлээд тооцоолсон $W \times H$ хэмжээтэй текстурыг суурь өнгө буюу улбар шар өнгөөр дүүргэнэ. Хээ зурахдаа P олонлогийн хувиргалт хийсэн цэгүүдийг ашиглаж хөрш хоёр цэгүүдийг холбож өргөн шулуун зурах байдлаар зурна. Шулуун зурахдаа шулуун зурдаг ямар ч алгоритм ашиглаж болно. Жишээ нь P_{ulzii} олонлогийн бүх хөрш цэгүүдийг холбож зурвал Зураг 4 дээрх байдлаар өлзий хээнүүд зурагдана. Хайрцаг нь хөндлөн эсвэл босоо байхаас хамааран хээний x, y байрлалыг y, x болгож сольж зурж болно. Хайрцаг болгоны дотор талд өлзий хээ болон хоёр булангийн хээ зурна. Хаалганы хатавчны хайрцгийн хувьд төрөл нь 1 бол нэг том хайрцаг байх учраас гурван жижиг өлзий хээ харин 2 бол нэг жижиг өлзий хээ зурна. Зураг 4 дээр хаалганы хатавчны хайрцгийг тойрч хар зураас зурсан нь тэнд хайрцаг байгаа гэдгийг харуулах зорилгоор зурсан гэхдээ 3D модель дээр ашиглахдаа зураас зурах шаардлагагүй учир нь модель өөрөө хайрцгийн хэлбэрээр цааш нүхлэгдсэн байна.



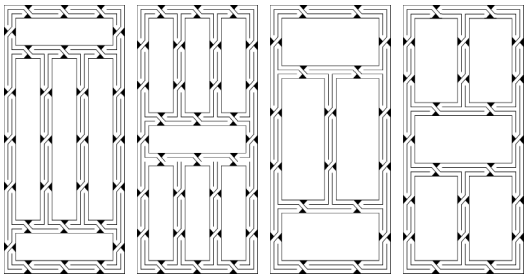
Зураг 4: 131 суурь төрөлтэй хаалганы хатавчны төрлүүд.

2) Хаалганы хавтасны хүрээний хэв зурах: Зураг 4-өөс хаалганы хавтас дээр хар хүрээтэй хэв зурагдсан байгааг харж болно. Энэ хэв хаалганы хавтасны хайрцгаас бусад хэсэг буюу хүрээнд зурагдана. Хүрээн дотор дөрвөн шулуун зурж Зураг 5-ын зүүн талын зурагтай адил болгоно. Тэр зургийг ажиглавал хайрцаг болгоныг тойрсон хоёр хар дөрвөлжин байгааг ажиглаж болно. Хамгийн сүүлд хаалганы хавтасны дотор талд хоёр хар дөрвөлжин зурснаар зүүн талын зурагтай ижил үр дүнд хүрнэ.

Дараа нь Зураг 5-ын баруун талын зурагт байгаа холбоосыг үүсгэж Зураг 6 дээр байгаа шиг хаалганы төрөлдөө харгалзах цэгүүд дээр холбоосны зургийг зурна. Жишээ нь 131 болон 313 төрөлтэй хаалган дээр босоо гурван баганатай учраас хэвтээ хүрээн дээр байрлаж байгаа холбоос бүр гурван ширхэг, харин 121 болон 212 баганатай бол хоёр ширхэг холбоос байна. Харин хайрцаг хэвтээ хэлбэртэй байвал 1, босоо хэлбэртэй байвал 131 төрөлд 3 бусад төрөлд 2 холбоос хажууд нь байна.



Зураг 5: Хаалганы хавтасны хүрээ болон холбоос.



Зураг 6: Хаалганы төрлөөс хамаарсан хүрээний хээнүүд.

3) Хүрээний холбоос үүсгэх: Зураг 5 баруун талд байгаа хүрээний холбоосыг үүсгэхдээ Алгоритм 1-ийн дагуу үүсгэнэ. Энэ алгоритмд ω нь хар шулууны өндөр, δ нь хар шулуун хоорондох зай, λ нь диагональ дээрх хэвтээ шулууны урт, x нь дээд талд зурах шулууны эхлэх координат, w нь шулууны уртыг тус бүр илэрхийлнэ. Зураг 5-ын баруун талыг харвал дээд талд доошоо харсан λ хэлбэрийн хэвтээ шулуунууд доод талд нь урвуу зурсан шулуунууд харагдана.

IV. Хэрэгжүүлэлт ба Үр дүн

Энэхүү бүлэгт санал болгосон процедур аргыг *Unity* график хөдөлгүүр дээр хэрэгжүүлсэн явц болон түүний гүйцэтгэлийн үр дүнг танилцуулна.

Алгоритм 1 Хүрээний холбоос үүсгэх алгоритм

Require: $t, \omega = \max(0.04t, 1), \delta = \frac{t-4\omega}{3}, \lambda = \sqrt{2\omega^2}$

1: $n \leftarrow t - \delta - \omega$
2: **for** $i = 0$ **to** $n - 1$ **do**
3: $x \leftarrow \begin{cases} 0 & i < \omega \\ \frac{t-2\delta}{2} + i - \omega & \text{бусад үед} \end{cases}$
4: $w \leftarrow \begin{cases} t & i < \omega \\ 2\delta + \lambda - 2(i-\omega) & i < \omega + \delta \\ \lambda & i < n - \omega \\ t - x & \text{бусад үед} \end{cases}$
5: хөндлөнШулуун$(x, t-1-i, w, border)$
6: хөндлөнШулуун$(t-w-x, i, w, border)$
7: **end for**

A. Хэрэгжүүлэлтийн орчин

Процедур аргаар үүсгэх алгоритмыг *Unity 6* (6000.3.5f2 LTS) график хөдөлгүүр дээр *C#* програмчлалын хэл ашиглан хөгжүүлсэн. Туршилтыг дараах техникийн үзүүлэлт бүхий систем дээр гүйцэтгэв:

- **Процессор:** Intel(R) Core(TM) i7-8850H CPU @ 2.60GHz (6 Cores, 12 Threads)
- **График карт:** Intel(R) UHD Graphics 630 (Integrated GPU)
- **Санах ой:** 16GB DDR4 @ 2667 MHz

B. Визуал үр дүн ба хувилбарууд

Бидний боловсруулсан алгоритм нь хаалганы төрөл болон хаалганы хатавчны төрлийн хослолоор Монгол гэрийн хаалганы архитектурын үндсэн хувилбаруудыг бодит цаг хугацаанд үүсгэж байна. Зураг 7-оос харахад алгоритмаар үүссэн текстур нь уламжлалт хаалганы пропорцыг бүрэн хадгалж, үндэсний хээ угалзыг геометрийн алдаагүй байршуулсан байна.

Нэгж сегментийн хэмжээг тодорхойлогч t параметрийг өөрчлөхөд текстурын нарийвчлал дараах байдлаар өөрчлөгдөж байна. $t = 25$ үед бага нарийвчлалтай атлас үүсэх ба хол зайнаас харагдах объектод тохиромжтой бол, $t = 100$ үед хаалганы ирмэг болон хээний зураас маш цэвэрхэн харагдаж байна. Энэ нь бидний алгоритм масштабаас үл хамаарах шинж чанартайг баталж байна.

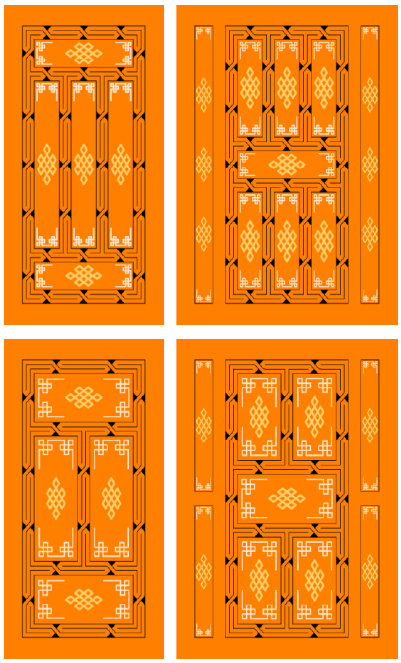
C. Гүйцэтгэлийн шинжилгээ

Алгоритмын үр ашгийг тодорхойлохын тулд текстур үүсгэх хугацаа болон санах ойн ашиглалтыг хэмжив.

1) Үүсгэх хурд: Текстурын бүх пикселийг тооцоолж, *Texture2D* объект болгон үүсгэхэд зарцуулсан дундаж хугацааг Хүснэгт I-д харуулав.

Хүснэгт I: Текстур үүсгэх хугацааны шинжилгээ (t параметрээс хамааран)

Нарийвчлал (t)	10	25	50	75	100
Хугацаа (мс)	1 мс	4 мс	16 мс	64 мс	100 мс



Зураг 7: Алгоритмаар үүсгэсэн хаалганы текстурын төрлүүдийн болон хатавчны төрлийн хувилбарууд.

Туршилтаас үзэхэд дунджаар 37 мс зарцуулж байгаа нь уг алгоритмыг аппликейшн ажиллаж байх явцад ямар ч гацалтгүйгээр ашиглах боломжтойг харуулж байна.

2) Санах ойн оновчлол: Уламжлалт аргаар хаалганы архитектурын 12 төрлийн хувилбарыг (4 төрлийн үндсэн хайрцаг $\times$ 3 төрлийн хатавч) 1800×2300 пикселийн нарийвчлалтайгаар, PNG форматаар хадгалсан текстурын сан (нийт 1.8 MB) ашиглахтай харьцуулахад бидний санал болгож буй алгоритмын эх код ердөө 10 КБ зай эзэлж байна. Энэ нь санах ойн ашиглалтыг 99.9%-аар бууруулж байгаа үзүүлэлт юм.

D. Хүрээний холбоосны чанар

Алгоритм 1-ийн дагуу үүсгэсэн 45°-ийн налуу холбоос нь текстурын ирмэг бүрд тэгш хэмтэй Зураг 8-аас харагдаж байна. Энэ нь гар аргаар зурсан текстурт тохиолддог пикселийн суналт болон ирмэгийн бүрсийлтийг бүрэн арилгаж, архитектурын хийцийг инженерийн нарийвчлалтай дүрсэлж чадсан.

V. Дүгнэлт

Энэхүү судалгааны ажлаар Монгол гэрийн архитектурын гол төлөөлөл болох хаалганы текстурыг компьютер графикийн модулиар параметрчилсэн процедур аргаар бодит цаг хугацаанд нийлэгжүүлэх алгоритмыг боловсруулж, амжилттай туршлаа. Судалгааны үр дүнд дараах үндсэн дүгнэлтүүдийг гаргаж байна:

1) **Архитектурын уян хатан зохиомж:** Хаалганы хавтасны бүтцийг хаалганы төрөл (131, 313, 121, 212) хэмээх дүрмийн системд шилжүүлснээр уламжлалт статик текстурын хязгаарлагдмал



Зураг 8: Хүрээний холбоос хэсгийг ойртуулсан байдал: Налуу холбоосны пикселийн тэгш хэм.

байдлыг шийдвэрлэсэн. Энэхүү модулиар хандлага нь хаалганы үндсэн хэсэг болон хатавчны олон төрлийн хувилбаруудыг нэг алгоритмаар динамикаар үүсгэх нөхцөлийг бүрдүүлэв.

2) **Математик хээ дүрслэл ба чанар:** Үндэсний ”Өлзий” болон ”Булан” хээг цэгэн координатын олонлогоор загварчилж, шугаман хувиргалтын матрицын тусламжтайгаар зурсан нь текстурыг масштабаас үл хамаарах шинж чанартай болгосон. t параметрийг өөрчлөх замаар текстурын нарийвчлалыг ямар ч чанарын алдагдалгүйгээр удирдах боломжтойг туршилтаар баталлаа.
3) **Системийн гүйцэтгэл ба хэмнэлт:** Санал болгосон процедур арга нь уламжлалт зургийн файл ашиглах аргатай харьцуулахад видео санах ойн ачааллыг 99.9%-аар бууруулж байна. Текстур үүсгэх хугацаа нь нарийвчлалаас хамааран 1–100 мс байгаа нь уг шийдлийг бодит цагийн тоглоом хөгжүүлэлт, виртуал болон нэмэлт бодит байдлын (*VR/AR*) салбарт өндөр гүйцэтгэлтэйгээр ашиглах боломжтойг харуулж байна.
4) **Соёлын өвийг дижиталчлах шинэ хандлага:** Монгол хаалганы хийцлэлийн дүрмийг математик логик болон кодын хэлэнд шилжүүлсэн нь соёлын биет өвийг зөвхөн дүрс төдий бус, инженерийн нарийн өгөгдөл хэлбэрээр дижитал орчинд өвлүүлэн үлдээх стратегийн ач холбогдолтой юм.

Цаашид энэхүү судалгааг өргөжүүлэн Монгол гэрийн доторх тавилга, дотоод засал болон үндэсний нарийн хээ угалзыг процедур аргаар илүү нарийвчлалтай алгоритмчилах, мөн олон төрлийн материалын шинж чанарыг алгоритмдаа нэгтгэхээр төлөвлөж байна.

Зүүлт

[1] D. Maidar, *Монголын архитектур ба хот байгуулалт*. Улаанбаатар: Улсын хэвлэлийн газар, 1972.

[2] C. Humphrey, “Inside a mongolian tent,” *Architectural Design*, vol. 44, no. 5, pp. 273–275, 1974.

[3] F. Remondino, “Heritage recording and 3d modeling with photogrammetry and 3d scanning,” *Remote sensing*, vol. 3, no. 6, pp. 1104–1138, 2011.

[4] D. S. Ebert, F. K. Musgrave, D. Peachey, K. Perlin, and S. Worley, *Texturing and modeling: a procedural approach*. Morgan Kaufmann, 2003.

[5] G. Pavlidis, A. Koutsoudis, F. Arnaoutoglou, V. Tsioukas, and C. Chamzas, “Methods for 3d digitization of cultural heritage,” *Journal of cultural heritage*, vol. 8, no. 1, pp. 93–98, 2007.

[6] P. Mueller, P. Wonka, S. Haegler, A. Ulmer, and L. Van Gool, “Procedural modeling of buildings,” *ACM Transactions on Graphics (TOG)*, vol. 25, no. 3, pp. 601–614, 2006.

[7] L.-Y. Wei, S. Lefebvre, V. Kwatra, and G. Turk, “State of the art in example-based texture synthesis,” *Eurographics 2009-State of the Art Reports*, 2009.

[8] S. Lefebvre and F. Neyret, “Stochastic procedural textures,” in *ACM SIGGRAPH 2005 Courses*, 2005, p. 6.

Assessment of Vibration Energy Harvesting Potential on Mongolian Freight Trains for IoT Applications

Batdalai Sukh[1], Ulzii-Orshikh Manjaa[1], Nyamsuren Purevsuren[1*], Myeirambyek Tashkyen[2], Battur Ganbat[2]
[1]Mongolian University of Science and Technology, Ulaanbaatar, Mongolia
[2]Government Implementing Agency Railway Authority of Mongolia, Ulaanbaatar, Mongolia
*Corresponding author: Nyamsuren@must.edu.mn

Abstract—This paper presents field measurement and spectral analysis of freight train vibrations on the Mongolian railway network to assess the feasibility of piezoelectric vibration energy harvesting (PVEH) for powering Internet of Things (IoT) monitoring nodes. A custom data acquisition system based on an ESP32-S3 microcontroller and an MPU9250 inertial measurement unit was developed and deployed on a freight train bogie to record triaxial vibrations at 500 Hz along a 400 km section under both loaded (80 t) and unloaded conditions. Power spectral density analysis using Welch's method identified two distinct, speed-invariant dominant frequencies: approximately 138 Hz for the unloaded wagon and 198 Hz for the loaded wagon. These load-dependent resonance characteristics, combined with measured acceleration amplitudes, establish the vibrational input parameters critical for designing dual-resonance piezoelectric energy harvesters tailored to Mongolian freight operations. The results demonstrate that the identified vibration signatures are sufficiently stable and energetic to support PVEH-based self-powering of low-energy IoT sensor nodes in remote railway environments.

Index Terms—vibration energy harvesting, railway monitoring, data acquisition, power spectral density, IoT, freight train

I. Introduction

Self-powered sensor networks are increasingly recognized as a transformative technology for railway management, enabling real-time condition assessment and predictive maintenance without reliance on conventional power infrastructure [1]. Freight trains operating across the Mongolian railway network traverse remote, sparsely populated regions where grid power is unavailable and frequent battery replacement is impractical. This creates a critical need for autonomous energy sources capable of sustaining wireless IoT monitoring nodes deployed on rolling stock.

Piezoelectric vibration energy harvesting (PVEH) offers a promising solution by converting ambient mechanical vibrations directly into electrical energy [2]. Unlike thermoelectric, solar, or electromagnetic alternatives, piezoelectric devices are compact, robust, and well-suited to the harsh operational environment of railway bogies [3]. However, the effectiveness of PVEH is strongly dependent on matching the harvester's resonant frequency to the dominant excitation frequencies present in the operating environment [4]. Therefore, accurate characterization of the vibrational spectrum is an essential prerequisite for harvester design.

Various approaches have been proposed to harvest mechanical energy in railway environments, including inductive voice-coil devices, thermoelectric generators, and piezoelectric transducers [6]. Among these, piezoelectric devices offer advantages in compactness, robustness, and direct strain-to-electricity conversion, making them well-suited for harsh railway environments [3]. Prior studies have demonstrated PZT harvester outputs of 81–1300 μW from sleeper-mounted configurations and approximately 1 mW from beam-mounted designs [5]. However, harvesting efficiency is strongly dependent on matching the harvester's resonant frequency to the ambient vibration spectrum [4].

While prior research has validated the potential of PVEH for railway monitoring, there is limited work on tailoring harvester design to the specific vibration spectrum of freight trains in Mongolia. This study addresses this gap by developing a data acquisition system, collecting field vibration data, and performing comprehensive spectral analysis. The primary contributions are: (1) the design and deployment of a custom microcontroller-based data acquisition system for characterizing railway vibrations; (2) comprehensive spectral analysis identifying the dominant vibration frequencies under loaded and unloaded conditions at multiple operating speeds; and (3) assessment of the suitability of these vibration characteristics for dual-frequency PVEH targeting IoT applications.

The remainder of this paper is organized as follows. Section II describes the vibration data acquisition methodology, including the system design and field deployment. Section III presents the data processing and spectral analysis approach. Section IV reports the measurement results. Section V discusses the implications for energy harvesting design. Section VI concludes with key findings and future directions.

II. Data Acquisition System and Deployment

A. System Design

A dedicated vibration data acquisition (DAQ) system was developed to capture high-resolution triaxial acceleration data from an in-service freight train. The system was built on the Rakwireless WisBlock platform, integrating the following components:

- Microcontroller: ESP32-S3 for onboard data processing and storage, programmed in C++ with hardware timer interrupts for precise sampling control.
- Inertial Measurement Unit: MPU9250, configured with a $\pm 16\,\mathrm{g}$ accelerometer range and $\pm 2000°/\mathrm{s}$ gyroscope range, communicating via $\mathrm{I^2C}$ at 400 kHz fast mode with a 500 Hz sampling rate.
- GPS Module: u-blox receiver recording latitude, longitude, and ground speed via UART at 9600 baud, with geographic coordinates scaled to microdegrees for enhanced storage precision.
- Storage: MicroSD card operating at 80 MHz SPI, with a double-buffering scheme and automatic file segmentation at approximately 50 MB.

A digital low-pass filter (DLPF) with a cutoff of 184 Hz was applied within the MPU9250 to mitigate high-frequency noise while preserving the relevant vibration bandwidth. Static calibration was performed at system initialization using the sensor's built-in auto-offset function to remove the initial DC bias. Each recorded entry comprised a timestamp, sampling index, triaxial accelerometer readings (in g), triaxial gyroscope readings (in °/s), latitude, longitude, and ground speed. Blue and green LED indicators provided visual feedback on write and sampling activity, respectively, enabling system health monitoring.

All components were integrated into a compact weatherproof IP65-rated ABS enclosure, as shown in Fig. 1, forming a self-contained portable unit suitable for harsh railway environments.

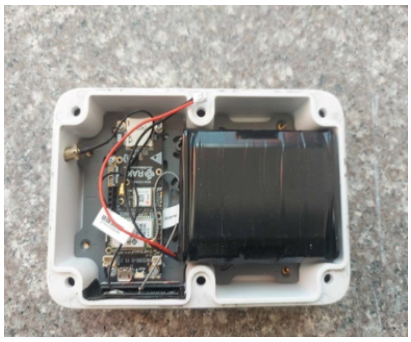


Fig. 1. Internal view of the custom data acquisition unit: ESP32-S3 WisBlock module and rechargeable battery in a weatherproof enclosure.

B. Software Architecture

Fig. 2 illustrates the software architecture of the data logger. The firmware implements a concurrent dual-context design: a hardware timer interrupt service routine (ISR) handles time-critical IMU sampling at a fixed 2 ms interval, while the main program loop manages non-time-critical tasks including GPS serial data reception, buffer management, and SD card write operations.

The double-buffering scheme uses two alternating 1000-sample buffers. While the ISR fills the active buffer with incoming IMU data, the main loop monitors whether the active buffer has reached capacity. Upon detecting a full buffer, the main loop atomically switches the active buffer pointer, allowing the ISR to continue writing to the alternate buffer without interruption, and then flushes the completed buffer to the SD card. This architecture ensures zero-loss data capture at the 500 Hz sampling rate, even during the relatively slow SD card write operations.

The GPS module operates asynchronously via UART serial communication. In each iteration of the main loop, the firmware checks for available serial data using the TinyGPS library and parses any incoming NMEA sentences to extract real-time position, velocity, and time data, which are synchronized with the vibration measurements for georeferencing.

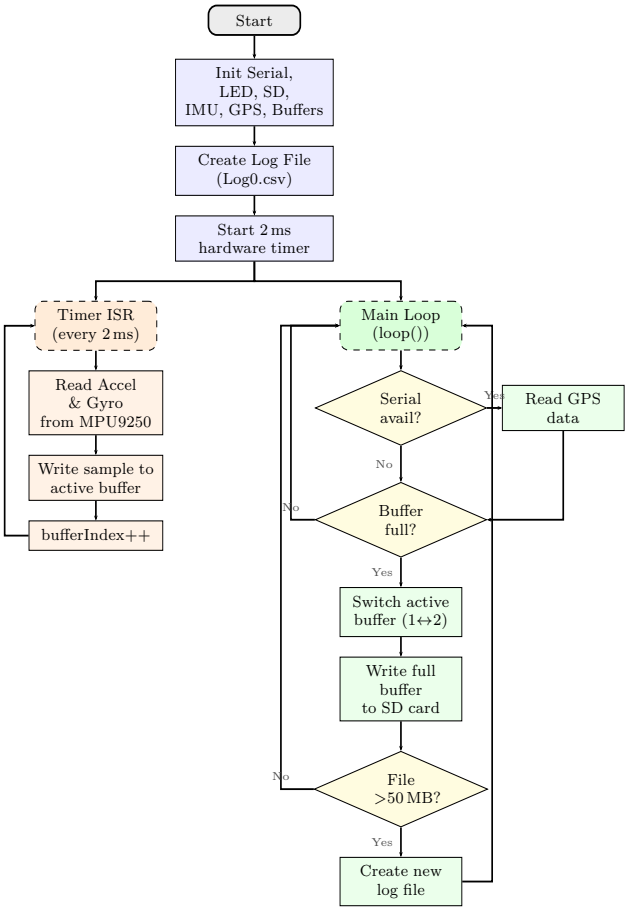


Fig. 2. Software architecture of the data acquisition system. The firmware uses a concurrent dual-context design: a hardware timer ISR (left) samples the IMU at 500 Hz, while the main loop (right) handles GPS reception, buffer management, and SD card storage.

C. Calibration and Data Integrity

Calibration: Static calibration was performed during system initialization to remove the initial DC bias using the sensor's built-in auto-offset function. To address temperature drift and constant alignment errors during the 5-hour continuous field test, a high-pass filter with a 1 Hz cutoff frequency was applied during post-processing, effectively removing low-frequency drift components without affecting the target vibration frequencies (120–220 Hz).

Saturation and Data Integrity: While the ±16 g range accommodated the majority of vibrations, minor signal clipping (saturation) was observed during high-impact events over rail joints. To ensure accuracy, average power calculations were derived exclusively from non-saturated data segments. However, spectral analysis was performed on both saturated and non-saturated segments, confirming that clipping did not shift the dominant resonant frequencies (∼138 Hz and ∼198 Hz).

D. Field Deployment

The DAQ system was powered by a rechargeable battery and mounted on the bogie bearing housing (axle-box region) of a freight wagon, identified through preliminary assessments as the location exhibiting the highest vibration amplitudes. This mounting position captures realistic axle-box vibration profiles representative of in-service conditions.

Field measurements were conducted along a 400 km segment of the Mongolian railway over a continuous 5-hour duration. Data were recorded under two representative operating conditions:

- Loaded condition: Fully loaded freight train carrying 80 tons of coal.
- Unloaded condition: Empty (unloaded) return wagon.

As shown in Fig. 3, the data logger was mounted on the bogie to capture realistic axle-box vibration profiles under in-service operating conditions. These empirical measurements provided essential input for assessing the vibration energy harvesting potential.

The deployment produced a comprehensive time-synchronized dataset of acceleration, gyroscope, GPS, and speed data for analyzing vibration behavior under different operating conditions.



Fig. 3. Installation of the data logger on a freight train bogie for vibration measurement. Left: the freight wagon (MGL 88002076). Right: the data logger mounted on the bogie bearing housing (axle-box region).

III. Data Processing and Spectral Analysis

A. Preprocessing

The acquired triaxial acceleration data were preprocessed to remove DC offsets by subtracting their mean values and converting from gravitational units (g) to SI acceleration units (m/s^2) using a scale factor of 9.81 m/s^2. Multiple CSV files containing triaxial acceleration, gyroscope, GPS, and speed information were merged into a unified dataset.

The dataset was segmented based on operational state. RMS velocity, derived from the integration of acceleration data and cross-referenced with GPS speed, was used to classify segments as 'loaded' or 'unloaded'. This speed-based segmentation confirmed that the dominant vibration frequencies shifted based on train mass rather than speed. The combined data were then segmented into speed intervals of 30, 50, and 70 km/h to enable condition-specific spectral analysis.

B. Power Spectral Density Estimation

The power spectral density (PSD) was computed using Welch's method, implemented through the scipy.signal.welch function in Python. Welch's method was selected for its reduced variance compared to single-segment FFT estimates, which is advantageous for long-duration, non-stationary railway measurements. The following parameters were applied uniformly across all analyses:

- Hann window for spectral smoothing
- FFT length $N_{\mathrm{FFT}} = 4096$
- 50% window overlap
- Mean detrending applied to each Welch segment to suppress residual DC components
- Sampling rate: 500 Hz, yielding a frequency resolution of approximately 0.12 Hz (500/4096)
- Analysis frequency range: 0–250 Hz to capture the primary wagon–track vibration modes

For each axis and speed interval, the PSD curves were generated and used to identify key spectral features. The dominant frequency was defined as the frequency corresponding to the maximum PSD amplitude within the analysis range. Additionally, the full width at half maximum (FWHM) of the dominant peak was computed to quantify the effective resonance bandwidth for subsequent harvester design. All PSDs were visualized using a consistent color scheme, and axes were consistently formatted using scientific notation for clarity.

This PSD-based methodology provides a consistent, reproducible framework for identifying the principal vibration characteristics required for subsequent vibration energy harvesting analysis.

IV. Results

A. Time-Domain Vibration Response

The raw acceleration signals were preprocessed by removing the DC offset to eliminate baseline drift and converted from gravitational units (g) to SI acceleration units (m/s^2). Representative steady-state intervals corresponding to three operating speeds (30 km/h, 50 km/h, and 70 km/h) were extracted for analysis.

Initial analysis of the triaxial acceleration data revealed that the vertical axis (Z-axis) exhibited significantly higher vibration magnitudes compared to the longitudinal (X-axis) and lateral (Y-axis) directions under both loading conditions. This dominance of vertical dynamics is consistent with the primary wheel–rail interaction mechanism at the axle-box location and confirms the Z-axis as the most relevant direction for PVEH design.

Fig. 4 and Fig. 5 show the time-domain acceleration of unloaded and loaded freight wagons in the X, Y, and Z directions together with GPS-derived speed. The vertical axis shows the largest vibration amplitude compared with the horizontal axes. The time-domain response provides direct insight into vibration amplitude variation with train speed, enabling identification of transient events and establishing the foundation for frequency-domain analysis.

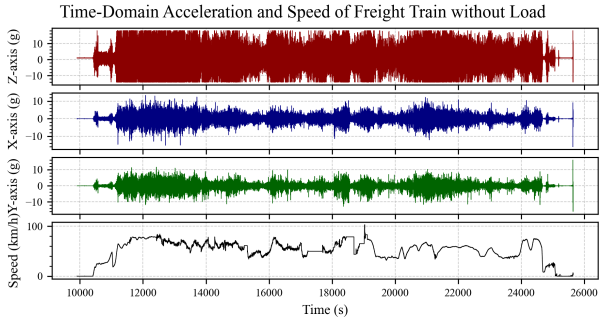


Fig. 4. Time-domain acceleration of an unloaded freight wagon in the X, Y, and Z directions and speed.

B. Spectral Characteristics: Unloaded Condition

PSD analysis of the unloaded freight wagon revealed a well-defined dominant resonance peak in the Z-axis within the 136–140 Hz range, as shown in Fig. 6. The X and Y axes do not exhibit comparable resonance features. Table I summarizes the identified peak frequencies and corresponding FWHM values at three operating speeds.

The predominant frequency remained relatively stable across all three speeds, with variations of less than 4 Hz, indicating that the resonance mode is governed by the structural dynamics of the wheel–rail–bogie system rather

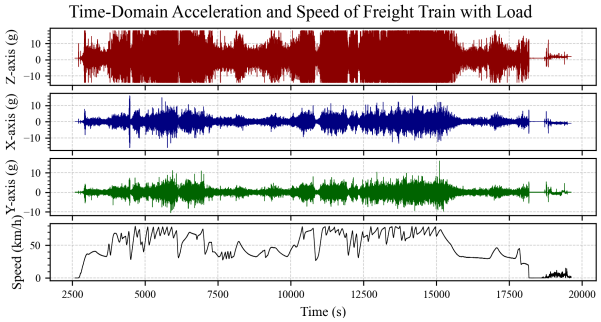


Fig. 5. Time-domain acceleration of a loaded (80 t) freight wagon in the X, Y, and Z directions and speed.

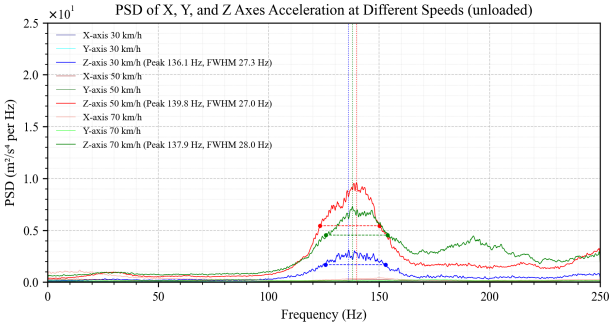


Fig. 6. Power spectral density of X-, Y-, and Z-axis acceleration at three train speeds (30, 50, 70 km/h) under unloaded conditions. The Z-axis shows a clear resonance peak; the dominant frequency and FWHM are indicated.

than train velocity. The spectral magnitude, however, increased substantially with speed, reflecting stronger dynamic excitation at higher velocities. While the spectral energy was relatively low at 30 km/h, the spectral peaks around 139 Hz were substantially amplified at 50 km/h and 70 km/h. These results collectively demonstrate that train speed primarily influences the intensity of vibrations, while the dominant frequency remains largely invariant.

The measured FWHM values of approximately 27–28 Hz indicate a moderately broad but consistent resonance band across speeds, which is favorable for PVEH as it relaxes the frequency-matching requirement. This finding suggests that the fundamental vibration modes are primarily governed by the wheel–rail interaction and the structural dynamics of the track, rather than being dependent on train velocity.

TABLE I
Z-axis dominant frequency: unloaded wagon.

Speed	Peak Freq. (Hz)	FWHM (Hz)
30 km/h	136.1	27.3
50 km/h	139.8	27.0
70 km/h	137.9	28.0

C. Spectral Characteristics: Loaded Condition

Under loaded conditions (80 t coal), the Z-axis PSD revealed a distinctly different dominant resonance centered in the 197–198 Hz range, as shown in Fig. 7, representing a shift of approximately 60 Hz relative to the unloaded case. The Z-axis shows a strong and well-defined dominant resonance peak within the 180–210 Hz range, whereas the X and Y axes do not exhibit comparable resonance features. Table II presents the corresponding results.

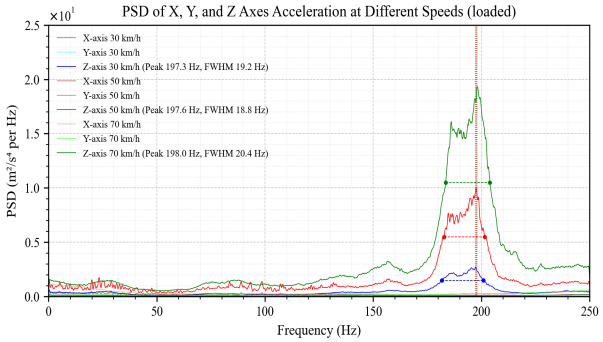


Fig. 7. Power spectral density of X-, Y-, and Z-axis acceleration for the loaded (80 t) train at three operating speeds (30, 50, 70 km/h). A dominant Z-axis resonance peak is observed near 198 Hz.

TABLE II
Z-axis dominant frequency: loaded wagon (80 t).

Speed	Peak Freq. (Hz)	FWHM (Hz)
30 km/h	197.3	19.2
50 km/h	197.6	18.8
70 km/h	198.0	20.4

The loaded condition peak frequency is similarly speed-invariant (within 0.7 Hz across speeds). The dominant Z-axis peak remains nearly unchanged across the three speeds, while the spectral magnitude increases substantially with velocity. The FWHM values of 19–20 Hz are narrower than the unloaded case, suggesting that the additional mass produces a more sharply defined resonance. This combination of a speed-invariant dominant frequency and only small variations in peak bandwidth is advantageous for vibration-based energy harvesting, as it enables the design of resonant harvesters that maintain performance over a range of train speeds without requiring adaptive tuning mechanisms.

The frequency shift between loaded and unloaded conditions can be explained by a simple mass–spring model. The natural frequency of the bogie–axle system is approximated as $f_n = \frac{1}{2\pi}\sqrt{k_{eq}/m_{eff}}$, where k_{eq} is the equivalent suspension stiffness and m_{eff} is the effective vibrating mass. Under loaded conditions, the 80 t coal payload increases the contact-patch normal force, which substantially stiffens the wheel–rail Hertzian contact (the dominant high-frequency spring element) and excites a higher-order vertical mode of the bogie substructure rather than the rigid-body sprung-mass mode observed in the unloaded case. This load-dependent modal shift is well-documented in railway vehicle dynamics [16], [17].

D. Comparison of Loading Conditions

The comparative spectral analysis reveals two key findings with direct implications for PVEH design:

1) Load-dependent frequency shift: The dominant Z-axis vibration frequency shifts from ∼138 Hz (unloaded) to ∼198 Hz (loaded), a difference of approximately 60 Hz. This shift is consistent across all measured speeds and is attributed to changes in the effective mass and boundary conditions of the bogie–axle system.
2) Speed-invariant peak frequency: Within each loading condition, the dominant frequency varies by less than 4 Hz across the 30–70 km/h operating range, confirming that train mass—not velocity—governs the resonance mode.

The X- and Y-axis responses do not exhibit comparable consistent peaks, confirming that vertical dynamics dominate the vibration environment in both loading scenarios. A comparison across all three axes confirmed that vertical vibrations consistently exhibited the largest amplitudes, emphasizing the dominant contribution of vertical dynamics to the overall railway vibration environment. These insights are particularly relevant for PVEH applications, suggesting that energy harvesters can be tuned to a stable resonance frequency around ∼138 Hz (unloaded) or ∼198 Hz (loaded), thereby ensuring consistent performance across various train speeds.

This frequency shift can be explained by a simple mass–spring model. The natural frequency of the wagon body suspension is approximated as $f_n = \frac{1}{2\pi}\sqrt{k/m_{\text{eff}}}$, where k is the secondary suspension stiffness and m_{eff} is the effective sprung mass. When the wagon is loaded, m_{eff} increases substantially while k remains essentially constant, resulting in a downward shift of the dominant natural frequency. This behavior is well-documented in railway vehicle dynamics [16], [17].

These characteristics establish the two target frequencies ($f_{\text{target1}} \approx 138$ Hz, $f_{\text{target2}} \approx 198$ Hz) for a dual-resonance PVEH system. In a realistic deployment scenario on Mongolian coal haulage routes, the freight wagon operates loaded during the outbound trip to the mine and returns unloaded. A dual-frequency harvester ensures continuous power availability throughout the entire logistics cycle.

E. Estimated Power Output

To quantitatively estimate the harvestable power, a single-degree-of-freedom (SDOF) lumped-parameter model was applied. At resonance, the maximum extractable electrical power is given by [14], [15]:

$$P_{max} = \frac{m \cdot A^2}{8 \cdot \zeta_T \cdot \omega_n} \tag{1}$$

where m is the proof mass, A is the RMS base acceleration, ζ_T is the total damping ratio, and $\omega_n = 2\pi f_n$. Substituting the measured $A \approx 1\,\mathrm{g}$ ($9.81\,\mathrm{m/s^2}$), $m = 10\,\mathrm{g}$, and $\zeta_T = 0.03$ yields $P_{max} \approx 4.6\,\mathrm{mW}$ at $f_n = 138\,\mathrm{Hz}$ (unloaded) and $P_{max} \approx 3.2\,\mathrm{mW}$ at $f_n = 198\,\mathrm{Hz}$ (loaded). These estimates align with the literature values cited above and confirm milliwatt-scale harvesting capability under both operating conditions.

V. Discussion

A. Implications for Energy Harvester Design

The identified vibration signatures have direct implications for the design of piezoelectric energy harvesters for Mongolian freight train applications. The presence of two distinct, stable, and well-separated resonance modes motivates a dual-frequency PVEH architecture—for example, a cantilever with two proof masses tuned to ∼138 Hz and ∼198 Hz—to ensure energy capture throughout the entire logistics cycle comprising loaded outbound and empty return trips.

A single-frequency harvester tuned to only one mode would lose efficiency under the opposite operating scenario. The measured FWHM values (19–28 Hz) indicate that a harvester with reasonably broad bandwidth can operate effectively even if the resonant frequency is not perfectly matched, though optimal performance requires tuning within approximately ±10 Hz of each target frequency.

The increasing spectral magnitude with speed suggests that the harvested power will scale with train velocity. At typical operating speeds of 50–70 km/h, the acceleration amplitudes are sufficient to drive milliwatt-level power generation from compact piezoelectric devices, as demonstrated by prior studies on similar railway PVEH systems achieving 0.2–1.1 mW under comparable excitation levels [5], [7]. Specifically, researchers have achieved output powers of 4.9 mW at 0.2 mm/7 Hz/5 g wheel-rail excitation, indicating significant potential for enhanced energy generation from railway vibrations [7].

B. Feasibility for IoT Power Supply

Typical axle-box IoT monitoring nodes—comprising a triaxial accelerometer, temperature sensor, low-power microcontroller, and LoRa/FSK transmitter—require approximately 0.4–1.0 mW average power with short bursts of 8–25 mW during data transmission. The stability and energy content of the measured vibration environment indicate that duty-cycled, partially self-powered operation of such nodes is feasible using appropriately designed PVEH systems.

It is important to note that in a practical implementation, additional power-management circuitry such as rectifiers, switches, and energy-storage elements would be required to condition and store the harvested energy. These components introduce conversion losses, so that only a fraction of the raw harvested power would be effectively available to power a wireless sensor node. The power levels assessed here should therefore be interpreted as an upper-bound estimate of the electrical energy that can be extracted from the mechanical vibrations. Optimization of the interface circuit and duty-cycled operation of the sensor node are necessary to enable practical self-powered deployment.

Even under conservative interface efficiency of 50 %, the harvested power exceeds the average node consumption by more than an order of magnitude, confirming energy-positive operation with sufficient margin for storage buffering during low-vibration intervals such as wagon stops and slow-speed maneuvers.

TABLE III
Energy Budget for IoT Node Operation

Parameter	Value
Active power (LoRa Tx burst)	∼15 mW for 200 ms
Sleep mode current	∼5 μA @ 3.3 V (16 μW)
Duty cycle (1 Tx / 5 min)	∼80 μW average
PVEH output, unloaded (Eq. 1)	∼4.6 mW
PVEH output, loaded (Eq. 1)	∼3.2 mW
Storage buffer (0.1 F supercap @ 3.3 V)	545 mJ (∼50 Tx)

TABLE IV
10-Year Total Cost of Ownership Comparison (50 IoT Nodes)

Item	Battery	PVEH
Hardware unit cost (USD)	8	60
Replacements per node (10 yr)	4	0
Labor per replacement (USD)	30	0
Total per node (USD)	160	60
Fleet total (50 nodes)	8,000	3,000

C. Comparison with Existing Technologies

Compared to conventional battery-powered sensors that require frequent replacements in remote railway sections, PVEH offers continuous energy scavenging from the operational environment [1]. Piezoelectric mechanisms provide advantages in simplicity, light weight, and adaptability to low-amplitude vibrations over electromagnetic or triboelectric alternatives [4]. The two distinct dominant vibration signatures identified in this study (∼138 Hz unloaded, ∼198 Hz loaded) motivate a dual-resonant PVEH structure to ensure continuous power availability throughout the entire freight logistics cycle.

D. Economic Comparison: Battery vs. PVEH

A 10-year total cost of ownership (TCO) analysis for a fleet of 50 sensor nodes (Table IV) further supports PVEH adoption. Although PVEH hardware is more expensive upfront, the elimination of recurring battery replacements and associated labor yields approximately a 62 % lifetime cost reduction. This advantage is particularly significant on the long-distance, remote Mongolian freight routes considered in this study, where maintenance access costs dominate the lifecycle expense.

VI. Conclusion

This paper presented the design, deployment, and analysis of a field vibration measurement campaign on the Mongolian freight railway to assess the feasibility of piezoelectric vibration energy harvesting for IoT applications. A custom ESP32-S3/MPU9250-based data acquisition system with a concurrent dual-context firmware architecture successfully captured triaxial vibration data at 500 Hz along a 400 km route under both loaded and unloaded conditions.

Spectral analysis using Welch's PSD method revealed two distinct, speed-invariant dominant vibration frequencies: ∼138 Hz for the unloaded wagon and ∼198 Hz for the loaded (80 t) wagon. Key findings include:

- The dominant Z-axis vibration frequency is load-dependent, shifting approximately 60 Hz between unloaded and loaded conditions.
- Within each loading condition, the peak frequency is speed-invariant (variation <4 Hz across 30–70 km/h), confirming that train mass governs the resonance mode.
- The FWHM of the resonance peaks ranges from 19–28 Hz, providing a moderately broad bandwidth favorable for resonant energy harvesting.
- Vertical (Z-axis) vibrations consistently dominate over longitudinal and lateral components.

These load-dependent but speed-invariant resonance characteristics provide the critical design parameters for a dual-frequency piezoelectric energy harvester capable of operating across the full freight logistics cycle. The measured vibration amplitudes and spectral stability confirm the viability of PVEH-based self-powering for duty-cycled, low-energy IoT monitoring nodes in this remote railway environment.

Future work will focus on the finite element design and experimental validation of a dual-resonance PVEH device tailored to these identified target frequencies, including integration with power management circuits and energy storage for practical deployment.

Acknowledgment

This research was supported by the "Capacity Building Project for School of Information and Communication Technology at Mongolian University of Science and Technology in Mongolia" (Contract No. P2019-00124), funded by the Korea International Cooperation Agency (KOICA). The authors acknowledge the Railway Authority of Mongolia for supporting the field measurements and providing access to freight trains used for data collection.

References

[1] G. Jing, R. Zhang, R. Connolly, D. P. Hou, W. and Aela, P., "Piezoelectric energy harvesters for railways: Recent trends and future opportunities," Multiscale and Multidisciplinary Modeling Experiments and Design, vol. 8, 2024.

[2] M. Gao, J. Cong, J. Xiao, J. He, Q. Li, S. Wang, Y. Yao, Y. Chen, R. and Wang, P., "Dynamic modeling and experimental investigation of self-powered sensor nodes for freight rail transport," Applied Energy, vol. 257, p. 113969, 2020.

[3] F. Rigo, M. Migliorini, and A. Pozzebon, "Piezoelectric sensors as energy harvesters for ultra low-power IoT applications," Sensors, vol. 24, pp. 2587–2587, 2024.

[4] L. Qi, H. Pan, Y. Pan, D. Luo, J. Yan, and Z. Zhang, "A review of vibration energy harvesting in rail transportation field," iScience, vol. 25, no. 3, p. 103849, 2022.

[5] B. Dziadak, M. Kucharek, and J. Starzyński, "Powering the WSN node for monitoring rail car parameters, using a piezoelectric energy harvester," Energies, vol. 15, pp. 1641–1641, 2022.

[6] C. A. Nelson, S. R. Platt, D. Albrecht, V. Kamarajugadda, and M. Fateh, "Power harvesting for railroad track health monitoring using piezoelectric and inductive devices," Proc. SPIE, vol. 6928, 2008.

[7] M. Gao, P. Wang, Y. Cao, R. Chen, and C. Liu, "A railborne piezoelectric transducer for energy harvesting of railway vibration," J. Vibroengineering, vol. 18, pp. 4647–4663, 2016.

[8] X. Gao, C. Qiu, G. Li, M. Ma, S. Yang, Z. Xu, and F. Li, "High output power density of a shear-mode piezoelectric energy harvester based on single crystals," Applied Energy, vol. 271, p. 115193, 2020.

[9] C. Liu, W. Zhang, Z. Junsheng, L. Ren, and X. Hong-quan, "Design and evaluation of a remote measurement system for the online monitoring of rail vibration signals," Proc. Inst. Mech. Eng. F, J. Rail Rapid Transit, vol. 230, pp. 724–733, 2014.

[10] R. Goodall, R. Dixon, R. Hamadache, M. and Ward, C., "Railways discovering mechatronics and monitoring – an overview," IFAC-PapersOnLine, vol. 53, pp. 8488–8493, 2020.

[11] M. Urbaniak, M. Myrcik, J. Juda, M. and Mandrysz, J., "The application of mobile devices for measuring accelerations in rail vehicles: Methodology and field research outcomes in tramway transport," Sensors, vol. 25, pp. 4635–4635, 2025.

[12] A. P. Gonzalo, M. Entezami, M. Weston, P. Roberts, and F. P. G. Márquez, "Railway track and vehicle onboard monitoring: A review," e3s conf., 2023.

[13] M. Sreenivasulu, V. U. Shree, and P. C. Reddy, "Acoustic energy harvesting through multilayer piezoelectric harvester model," Int. J. Innov. Technol. Explor. Eng., vol. 9, pp. 1818–1856, 2020.

[14] C. B. Williams and R. B. Yates, "Analysis of a micro-electric generator for microsystems," Sensors and Actuators A: Physical, vol. 52, no. 1–3, pp. 8–11, 1996.

[15] S. Roundy, "Energy scavenging for wireless sensor nodes with a focus on vibration to electricity conversion," Ph.D. dissertation, Dept. Mech. Eng., Univ. California, Berkeley, CA, USA, 2003.

[16] S. Iwnicki, Ed., Handbook of Railway Vehicle Dynamics. Boca Raton, FL, USA: CRC Press, 2006.

[17] V. K. Garg and R. V. Dukkipati, Dynamics of Railway Vehicle Systems. Toronto, ON, Canada: Academic Press, 1984.

[18] A. Erturk and D. J. Inman, Piezoelectric Energy Harvesting. Chichester, UK: John Wiley & Sons, 2011.

[19] S. P. Beeby, M. J. Tudor, and N. M. White, "Energy harvesting vibration sources for microsystems applications," Meas. Sci. Technol., vol. 17, no. 12, pp. R175–R195, Dec. 2006.

[20] S. Priya et al., "A survey on piezoelectric energy harvesting: Materials, methods, and circuits," Energy Harvesting and Systems, vol. 4, no. 1, pp. 3–39, 2017.

Дээд мөчний нөхөн сэргээх эмчилгээг дэмжих VR тоглоомын хөгжүүлэлт

О.Оюунбат
МУИС, МТЭС
Мэдээлэл, компьютерын ухааны тэнхим
(Мультимедиа технологи IV)
Улаанбаатар, Монгол
22B1NUM0312@stud.num.edu.mn

Х.Оюундолгор
МУИС, МТЭС
Мэдээлэл, компьютерын ухааны тэнхим
(профессор, PhD)
Улаанбаатар, Монгол
okhorloo@num.edu.mn

Хураангуй— Дээд мөчний үйл ажиллагаа алдагдах нь харвалт болон бусад гэмтлийн дараах түгээмэл хүндрэл бөгөөд нөхөн сэргээх эмчилгээнд өвчтөний идэвхтэй, тогтмол оролцоо нэн чухал байдаг. Энэхүү судалгааны ажлаар дээд мөчний нөхөн сэргээх эмчилгээг дэмжих зорилготой виртуал бодит байдал (VR) технологид суурилсан тоглоомын техникийн шийдэл, архитектур болон тоглоомчилсон (gamification) элементүүдийг бололвсруулав. Meta Quest VR төхөөрөмж болон Unreal Engine 5 платформ дээр хөгжүүлсэн тус систем нь хэрэглэгчийн гарын хөдөлгөөний кинематикийг бодит цагийн горимд тооцоолж, эмчилгээний дасгалуудыг сонирхолтой тоглоомын хэлбэрт хөрвүүлсэн. Системийн үндсэн давуу тал нь хөдөлгөөнийг илүү нарийвчлалтай тооцоолох, өвчтөн үргэлжлүүлэн тоглох сэдлийг нэмэгдүүлэх, мөн VR орчин дахь хөдөлгөөний дотор муухайрах (VR sickness) эрсдэлийг бууруулах техникийн оновчтой шийдлүүдэд оршино. Мөн VR орчинд гүйцэтгэлийг оновчлохын тулд текстурын хэмжээ, гэрэлтүүлэг, моделийн полигоны тоог бууруулах зэрэг оптимизацийн аргуудыг хэрэгжүүлсэн.

Түлхүүр үгс: виртуал бодит байдал, нөхөн сэргээх эмчилгээ, дээд мөчний хөдөлгөөн, цохилтын дасгал, Unreal Engine

I. Удиртгал

Дэлхийн Эрүүл Мэндийн Байгууллагын (ДЭМБ) мэдээлснээр нөхөн сэргээх эмчилгээ нь хүний үйл ажиллагааны чадавхыг сайжруулах, өдөр тутмын амьдралын чанарыг дээшлүүлэхэд зайлшгүй шаардлагатай эрүүл мэндийн тусламж үйлчилгээний салшгүй хэсэг юм [1]. Ялангуяа тархины цус харвалтын дараах дээд мөчний саажилтын үед уламжлалт нөхөн сэргээх эмчилгээ нь удаан хугацаа шаарддаг бөгөөд өвчтөнөөс маш их тэвчээр, идэвх зүтгэл шаарддаг. Нөхөн сэргээх эмчилгээ нь аливаа гэмтэл, өвчин, хөдөлгөөний чадвар алдагдсаны дараах биеийн үйл ажиллагааг сэргээхэд чиглэсэн давтамжтай болон зорилтот хөдөлгөөний дасгалууд дээр тулгуурладаг тул хүмүүст нэг хэвийн дасгал хөдөлгөөнийг удаан хугацаанд давтан хийхээс уйдах, сонирхол, идэвх буурах, дасгал хөдөлгөөнөө таслах зэрэг асуудлууд түгээмэл тохиолддог.

Сүүлийн жилүүдэд виртуал бодит байдал (VR) технологи нь эрүүл мэндийн салбарт өргөнөөр ашиглагдаж эхэлсэн бөгөөд хэрэглэгчийг бүрэн виртуал орчинд оруулж, биеийн хөдөлгөөнийг бодит цагт илрүүлэн, харааны болон дууны хариу үйлдэл үзүүлэх боломжийг бүрдүүлдэг. VR-д суурилсан нөхөн сэргээх эмчилгээ нь уламжлалт эмчилгээтэй харьцуулахад дээд мөчний хөдөлгөөний үйл ажиллагаа, өдөр тутмын чадварыг сайжруулахад эерэг нөлөө үзүүлж байгааг олон судалгаа харуулсан [2], [3], [4].

Түүнчлэн сүүлийн үеийн судалгаагаар боксын дасгал хөдөлгөөн нь дээд мөчний үйл ажиллагааг сэргээхэд чухал нөлөөтэй болохыг тогтоосон. Тухайлбал, тархины цус харвалтын дараах өвчтөнүүдэд хийгдсэн боксын дасгал нь уламжлалт дасгалтай харьцуулахад гарын хөдөлгөөний үйл ажиллагаа, булчингийн хүч, мөрний хөдөлгөөний далайцыг илүү сайжруулсан байна [5]. Мөн Beat Saber зэрэг VR тоглоомуудыг нөхөн сэргээх эмчилгээний судалгаанд ашиглаж, 900 минутаас дээш хугацаанд тоглосон хоёр оролцогчид хөдөлгөөний чанар нь сайжирсан болохыг тогтоосон [6].

Энэхүү судалгааны гол зорилго нь дээд мөчний үйл ажиллагааг нөхөн сэргээхэд чиглэсэн цохилтын хөдөлгөөнд суурилсан VR тоглоомыг хөгжүүлэх, түүний техникийн шийдэл болон тоглоомын элементүүдийн үр нөлөөг судлан туршихад оршино.

II. Судалгааны арга зүй ба тоглоомын хөгжүүлэлт

A. Технологийн сонголт ба хөгжүүлэлтийн орчин

Энэхүү системийг Meta Quest 2 болон түүнээс хойшхи үеийн бие даасан (standalone) VR төхөөрөмжүүдэд зориулан хөгжүүлсэн. Виртуал орчинд өвчтөнүүдэд түгээмэл тохиолддог дотор муухайрах (VR sickness) шинж тэмдгээс урьдчилан сэргийлэхийн тулд дүрсний солигдох хурдыг (frame rate) дор хаяж 72 FPS (frames per second) байлгах шаардлага тавигддаг. Системийн гүйцэтгэлийн бүх оптимизаци, график рендеринг нь энэхүү давтамжийг тогтвортой барих зорилтод чиглэсэн.

Тоглоомын логик болон орчны хөгжүүлэлтийг Unreal Engine 5 (UE5) платформ дээр гүйцэтгэсэн бөгөөд программчлалын хувьд Blueprint визуал скриптийн системийг ашиглан системийн уян хатан байдлыг хангасан [7],[8].

B. Тоглоомын архитектур ба орчин

Тоглоом нь хэрэглэгчийн сонирхлыг татах, эмчилгээний уйтгартай байдлыг эвдэх зорилгоор олон төрлийн орчин болон хүндрэлийн түвшинтэй. Уг тоглоом нь Circus, Alien, Zombie, Vampire, Human гэсэн 5 өөр сэдэвчилсэн виртуал орчинтой байх ба орчин бүр Easy, Medium, Hard гэсэн гурван хүндрэлийн түвшний сонголттой байх тул өвчтөний нөхөн сэргээх үе шаттай уялдан тохируулга хийгдэнэ.

Тоглоомын үндсэн дүрэм нь виртуал орчинд гарч ирэх хөдөлгөөнт объектуудтай харилцан үйлчлэл хийхэд суурилна.



Зураг 1: Эхлэлийн хэсэг.

C. Хөдөлгөөн илрүүлэх алгоритм ба кинематик тооцоолол

Дээд мөчний нөхөн сэргээх эмчилгээнд хөдөлгөөний нарийвчлал нь хамгийн чухал хэмжигдэхүүн юм. Хэрэглэгчийн гарын хөдөлгөөнийг Unreal Engine дэх Motion Controller компонентоор тасралтгүй хянаж, гарын байрлал болон хурдыг бодит цагт тооцоолсон.

Хөдөлгөөний хурдыг систем 1/72 секундийн давтамжтайгаар дараах байдлаар тооцоолно.

$$v = \frac{\Delta x}{\Delta t}$$

Үүнд, Δx нь гарын өмнөх болон дараах байрлалын зөрүү (орон зайн шилжилт), Δt = 1/72 секунд буюу нэг фреймийн хугацаа юм. Хөдөлгөөний хурдыг тооцолохдоо виртуал орон зайд хэрэглэгчийн гар нэг байрлалаас өөр нөгөө байрлал руу шилжихэд зарцуулсан хугацааг туулсан зайд нь харьцуулан авч үзсэн.

Тархины саажилттай эсвэл гэмтэлтэй өвчтөнүүдийн санамсаргүй, сул хөдөлгөөнөөс үүдэлтэй буруу илрүүлэлтийг (false positive) шүүхийн тулд хурдны босго утгыг $v > 20$ байхаар тохируулж, зөвхөн зорилтот цохилтын хөдөлгөөнийг ялгасан. Үүнд 20 гэх утга нь UE5 дахь дотоод нэгж болох 20 см/с -ийг илэрхийлнэ. Уг утгаар гарын чичиргээ, санамсаргүй хөдөлгөөн зэргийг шүүж цохилт мөн эсэхийг илрүүлнэ.

Цохилт амжилттай илэрсэн тохиолдолд виртуал объектод үйлчлэх хүчийг хэрэглэгчийн гарын чиглэл болон хурдны хэмжээнээс хамааруулан дараах байдлаар тодорхойлсон.

$$F = d \times v$$

Энэхүү тэгшитгэлд d нь цохилтын чиглэлийн вектор, v нь хурдны скаляр хэмжээ юм. Уг илэрхийлэл нь Ньютоны хоёрдугаар хууль дахь физикийн хүчний тэгшитгэл биш бөгөөд энэ нь Unreal Engine-ийн физик хөдөлгүүрт (Chaos Physics) объектын хариу үйлдлийг бүрдүүлэхэд хэрэглэсэн хоёр функцийн параметрийг товчилсон тэмдэглэгээ юм. Энэхүү тэмдэглэгээний бодит хэрэгжүүлэлтийг дараах байдлаар тодорхойлно.

$$F = SetPhysicsLinearVelocty(d) + AddForce(v)$$

Виртуал объектод цохилтын хариу үйлдлийг үзүүлэхийн тулд Unreal Engine-ийн хоёр API функцийг хослуулан ашигласан. Эхний функц нь *SetPhysicsLinearVelocity*(d) бөгөөд объектын одоогийн шугаман хурдны векторыг шидэлтийн чиглэлийн вектор d-ээр шууд солино. Энэ нь мөргөлдөөний агшинд объектод импульс дамжуулсантай адил үр дүнг өгч, гарын хөдөлгөөний чиглэлд тохирсон шилжилтийг өдөөдөг. Хоёр дахь функц нь *AddForce*(v) бөгөөд объектод нэмэлт хүчийг чиглэлийн дагуу v скаляр хэмжээтэйгээр үйлчилнэ.

Энэхүү физик харилцан үйлчлэлийн аргачлал нь хэрэглэгчийн бодит хөдөлгөөнийг виртуал орчинд динамик байдлаар тусган, харилцан үйлчлэлийг илүү бодит мэт мэдрүүлэх боломжийг бүрдүүлж өгсөн.

III. Тоглоомын элемент ба хэрэглэгчийн идэвх оролцоо

Нөхөн сэргээх эмчилгээний амжилтын гол үндэс нь өвчтөнийг сэдэлжүүлэх явдал юм. Судалгаанаас үзэхэд видео тоглоомын дизайны зарчмууд болон тоглоомын элементүүд нь өвчтөний эмчилгээнд оролцох идэвхийг эрс нэмэгдүүлдэг [9], [10], [11]. Тус системд дараах тоглоомын механизмуудыг нэвтрүүлсэн. Үүнд:

Онооны систем: Хэрэглэгч виртуал объектуудыг устгах бүрд үйлдэл гүйцэтгэсэн хугацаа болон хөдөлгөөний чанараас хамаарч оноо авна.

Streak механик: Дараалсан амжилттай цохилтын тоо тасралтгүй нэмэгдэх тусам "streak" коэффициент өсөж, үржигдэх замаар нэмэлт оноо бодогдохоор тоглоомын логикийг хэрэгжүүлсэн. Энэ нь анхаарал төвлөрөлтийг сайжруулахад чухал үүрэгтэй.

Ultimate чадвар: Тодорхой тооны объектыг амжилттай устгасны дараа "ultimate bar" дүүрч, хэрэглэгчдэд богино хугацаанд илүү хүчтэй, том талбайг хамарсан үйлдэл хийх боломжийг олгоно. Энэхүү шагналын систем нь хэрэглэгчийн идэвхийг хадгалах, ядаргааг мартуулах, идэвхийг сэргээх зорилготой [11].



Зураг 2: Тоглоомын оноо тооцох дүрэм.

IV. Үр дүн ба хэрэгжүүлэлт

A. Виртуал орчны хөгжүүлэлт ба оптимзаци

Системийн виртуал орчныг хэрэглэгчийн сонирхлыг татахуйц Circus, Alien, Zombie, Vampire, Human гэсэн таван өөр орчноос бүрдэх, тус тусын түвшинтэй байхаар UE5 платформ дээр зохион байгуулж хөгжүүлсэн. Энэхүү архитектур нь санах ойн ачаалыг үлэмж хэмжээгээр бууруулж, зөвхөн тухайн үед тоглох шаардлагатай орчныг санах ойд ачаалах боломжийг бүрдүүлсэн.

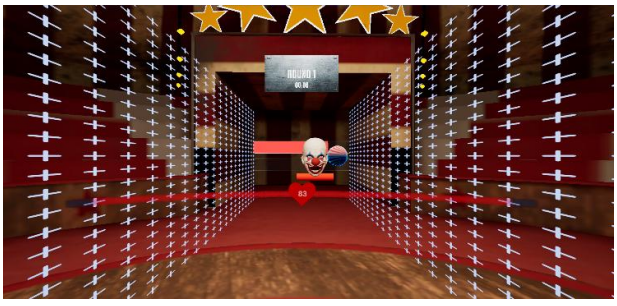
Бие даасан VR төхөөрөмж (Meta Quest 2) дээр дүрс солигдох хурдыг (хамгийн багадаа 72 FPS) тасралтгүй, тогтвортой хадгалах нь өвчтөнд дотор муухайрах (VR sickness) шинж тэмдэг илрэхээс сэргийлэх гол нөхцөл юм. Тиймээс системийн хөгжүүлэлтийн шатанд график рендерингийн дараах оптимизацийн аргуудыг амжилттай хэрэгжүүлэв. Үүнд:

- Виртуал орчны объектууд болон дүрүүдэд цөөн өнцөгт бүхий (low-poly) 3D загваруудыг сонгон ашиглаж, график процессорын (GPU) ачааллыг бууруулсан.
- Системд ашиглагдсан нийт текстурын хэмжээг 1024×1024 пикселээс 256×256 болон 512×512 пиксел хүртэл оновчтойгоор бууруулж, видео санах ойн (VRAM) зарцуулалтыг багасгасан.
- Бодит хугацааны гэрэлтүүлгээс үүсэх тооцооллын ачааллыг арилгахын тулд орчин бүрд хөдөлгөөнгүй (static) гэрэлтүүлгийн аргыг ашиглан гэрлийн мэдээллийг текстур руу буулгасан. Түүнчлэн, тухайн орчны сэдэвчилсэн онцлог, уур амьсгалаас хамааруулан цэгэн болон чиглэлтэй гэрлүүдийг хослуулан байрлуулж, харагдах байдлын чанарыг алдагдуулахгүйгээр гүйцэтгэлийг боломжит түвшинд хүргэсэн.

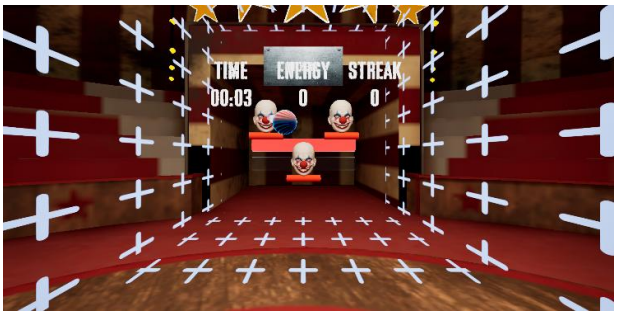
B. Тоглоомын горимууд ба эмчилгээний зорилт

Хөдөлгөөний эмчилгээний ялгаатай зорилтуудад нийцүүлэн системийг гурван үндсэн горимоор ажиллуулахаар хөгжүүлсэн.

1. Classic горим: Нийт 6 үе шаттай бөгөөд зорилтот объектуудыг устган оноо цуглуулна. Хүндрэлийн түвшин аажмаар нэмэгдэх алгоритм нь хэрэглэгчийн ур чадвар болон сорилт хоорондын тэнцвэрийг хангана [9].
2. Soccer горим: Хөл бөмбөгийн зарчимд суурилсан 4 үе шаттай горим. Энэ нь зөвхөн цохих бус, бөмбөгийг тодорхой чиглэлд оновчтой залах буюу орон зайн баримжаа, нарийвчлалыг шаарддаг.
3. Fitness горим: Дээд мөчний тэсвэр тэвчээрийг сайжруулахад чиглэсэн. Нэгэн зэрэг 3-6 виртуал объектыг виртуал орчинд байрлуулж, өвчтөнөөс илүү хурдтай, өргөн далайцтай хөдөлгөөн хийхийг шаарддаг эрчимт дасгалын хэлбэр юм.







Зураг 3: Тоглоомын горимууд (a) Classic горим; (b) Soccer горим; (c) Fitness горим.

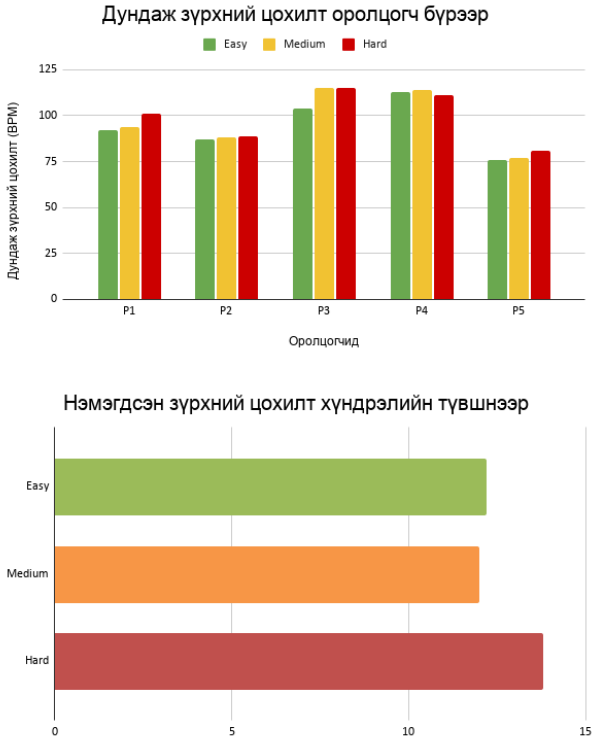
V. Туршилтын үр дүн ба өгөгдлийн шинжилгээ

A. Туршилтын тохиргоо

Системийн анхдагч үр нөлөөг үнэлэх зорилгоор эрүүл, насанд хүрсэн 5 оролцогч Classic горимын Easy, Medium, Hard гурван хүндрэлийн түвшинд тус бүр нэг удаа тоглож, нийт 15 тоглолтын үр дүнг бүртгэн авч шинжилгээ хийсэн. Тоглолт тус бүрд зүрхний цохилт (HR), идэвхтэй илчлэг (kcal), эхлэх ба дуусах хугацаа, цуглуулсан оноо, мөн нийт цохилтын тоо, дундаж болон дээд цохилтын хурд (UE5-ийн дотоод нэгжээр) гэсэн хэмжилтийн утгуудыг бүртгэн авсан. Хурдны босго утгыг $v > 20$ байхаар авч санамсаргүй чичиргээг шүүсэн бөгөөд энэ түвшнээс дээш бүртгэгдсэн цохилтын утгуудыг цаашдын шинжилгээнд ашигласан.

B. Зүрхний цохилтын хариу үйлдэл

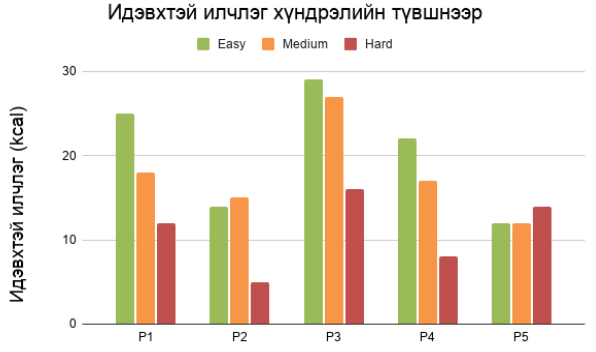
Easy түвшинд дундаж зүрхний цохилт 94.4 BPM, Medium түвшинд 97.6 BPM, Hard түвшинд 99.4 BPM байв. Хүндрэлийн түвшин нэмэгдэх тусам зүрхний цохилт жигд өсөж, энгийн үеийн цохилттой харьцуулсан зөрүү Hard түвшинд хамгийн өндөр буюу 13.8 BPM-д хүрсэн нь манай систем сонгосон хүндрэлийн түвшинд тохирсон физиологийн ачааллыг өвчтөнд өгч чадаж буйг харуулж байна. Зураг 4-т оролцогч тус бүрийн дундаж зүрхний цохилт болон тоглолтын өмнөх үеэс зүрхний цохилт нэмэгдсэн үзүүлэлтийг харуулав.

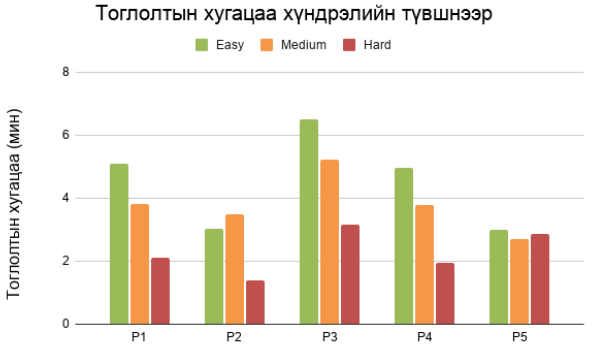


Зураг 4: Зүрхний цохилтын дүн шинжилгээ. (a) Оролцогч тус бүрийн дундаж зүрхний цохилт хүндрэлийн түвшнээр; (b) Тоглолтын өмнөх үеэс нэмэгдсэн зүрхний цохилтын график.

C. Илчлэгийн зарцуулалт ба тоглоомын хугацаа

Тоглолт тутамд зарцуулсан идэвхтэй илчлэг Easy түвшинд дунджаар 20.4 ккал, Medium түвшинд 17.8 ккал, Hard түвшинд 11.0 ккал болж буурсан байна. Энэ нь Easy түвшний тоглолтын дундаж үргэлжлэх хугацаа 4.51 минут, Medium 3.80 минут, Hard 2.30 минут байсантай шууд холбоотой буюу хүндрэл нэмэгдэх тутам тоглоом богино хугацаанд эрчимтэй явагдаж байгааг илтгэнэ. Энэ шинж нь нөхөн сэргээх дасгалын хувьд эерэг шинж юм. Easy түвшин нь өвчтөнд урт хугацааны дасгалын тэсвэрийг сургахад тохиромжтой бол Hard түвшин нь хурдан, эрчимтэй цохилтын чадварыг сориход зориулагдсан. Зураг 5-т оролцогч тус бүрийн идэвхтэй илчлэг болон тоглолтын үргэлжлэх хугацааг хүндрэлийн түвшнээр харьцуулан үзүүлэв.

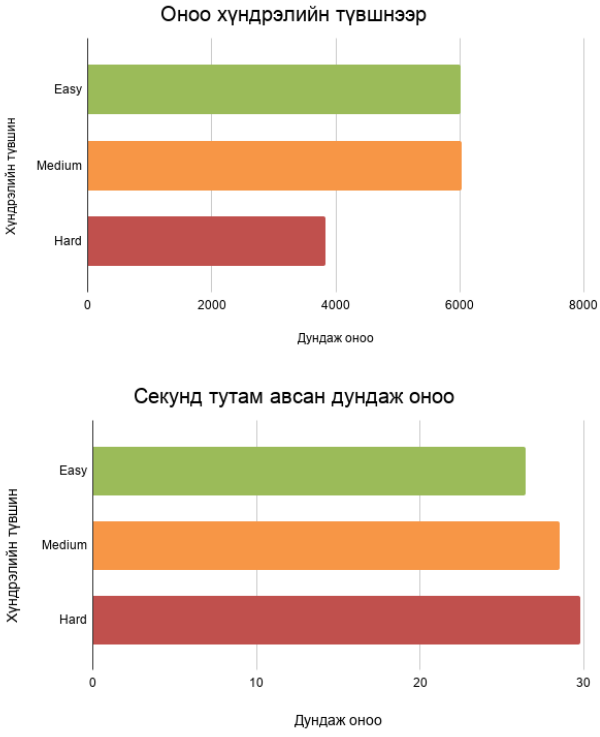




Зураг 5: (a) Оролцогч тус бүрийн идэвхтэй илчлэгийн зарцуулалт; (b) Тоглолтын үргэлжлэх хугацааны график хүндрэлийн түвшнээр.

D. Тоглоомын онооны шинжилгээ

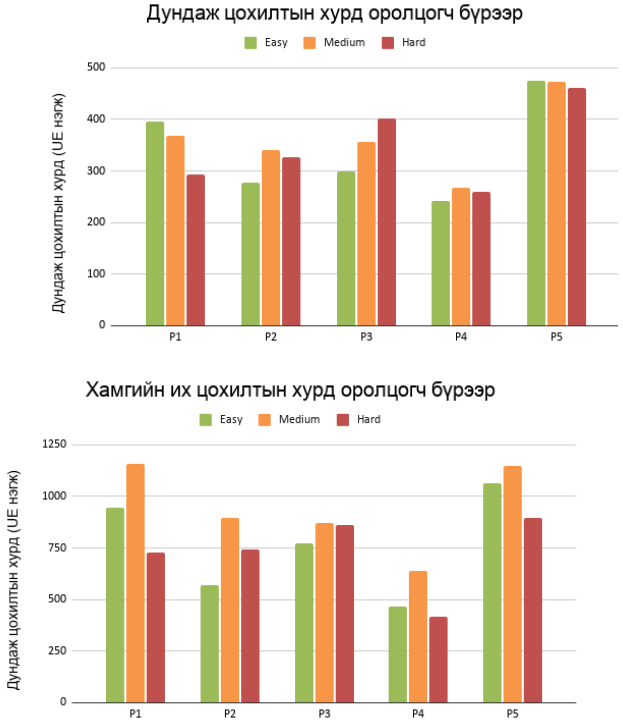
Classic горимд авсан дундаж оноо Easy 6020, Medium 6037, Hard 3828 байсан бөгөөд Hard түвшний оноо хамгийн их хэлбэлзэлтэй гарсан нь оролцогчдын хооронд ур чадварын ялгаа их байсныг илэрхийлж байна. Харин эсрэгээс нь секунд тутамд авах оноогоор тооцвол Easy 26.4, Medium 28.5, Hard 29.8 буюу хүндрэлийн түвшин нэмэгдэх тутам секундэд авах оноо ихсэж байв. Энэ нь Hard түвшинд оноог илүү өндөр коэффициенттойгоор өсгөх дизайн нь зорилтот үр дүнгээ үзүүлж байгааг харуулж байна. Зураг 6-т онооны болон секунд тутамд авсан онооны үзүүлэлтийг хүндрэлийн түвшнээр үзүүлэв.



Зураг 6: (a) Тоглоомын онооны үзүүлэлт хүндрэлийн түвшнээр; (b) Секунд тутамд авсан онооны график

E. Цохилтын хурдны шинжилгээ

Таван оролцогчоос авсан нийт 1301 цохилтын өгөгдөлд хийсэн шинжилгээгээр дундаж цохилтын хурд хүндрэлийн түвшнээр Easy 338, Medium 361, Hard 349 байв. Харин цохилтын дээд хурд нь системийн гүйцэтгэлийн боломжит утгаар тодорхойлогдоно. Бүх оролцогчдын хувьд Medium түвшинд бүртгэгдсэн дээд утга хамгийн өндөр буюу 1157 нэгжид хүрсэн нь систем хэрэглэгчдээс эрчимтэй цохилт өдөөж чадаж байгааг илтгэнэ. Оролцогчдын хувьд P5 (474 нэгж дундажтай) болон P1 (396 нэгж дундажтай) хамгийн өндөр хурдтай, P4 (242–267 нэгж) хамгийн зөөлөн цохилт хийсэн нь хувь хүний биеийн хүчний ялгаатай байдлыг харуулж байна. Зураг 7-т оролцогч тус бүрийн дундаж болон дээд цохилтын хурдыг хүндрэлийн түвшнээр харьцуулсныг үзүүлэв.



Зураг 7: Цохилтын хурдны харьцуулсан шинжилгээ. (a) Оролцогч тус бүрийн дундаж цохилтын хурд хүндрэлийн түвшнээр; (b) Дээд цохилтын хурд хүндрэлийн түвшнээр.

VI. ДҮГНЭЛТ

Энэхүү судалгааны ажлаар VR технологид суурилсан, дээд мөчний хөдөлгөөнийг нөхөн сэргээх эмчилгээг дэмжих цохилтын хөдөлгөөнд суурилсан тоглоомыг Unreal Engine 5 платформ дээр амжилттай хөгжүүлсэн. Хөдөлгөөн таних математик загварчлал болон тоглоомын элементүүдийг оновчтой хослуулсан нь өвчтөнд өгөгдөх физик ачааллыг сонирхолтой хэлбэрт оруулах боломжийг бүрдүүлсэн.

Эрүүл, насанд хүрсэн 5 оролцогчийн дунд явуулсан анхдагч туршилтын үр дүнгээс үзэхэд манай систем нь хүндрэлийн түвшинд тохирсон физиологийн ачааллыг өвчтөнд өгч чадаж байгаа нь батлагдсан. Тухайлбал, дундаж зүрхний цохилт Easy түвшний 94.4 BPM-ээс Hard түвшинд 99.4 BPM хүртэл өсөж, тоглолтын өмнөх үетэй харьцуулсан зөрүү Hard түвшинд хамгийн өндөр буюу 13.8 BPM-д хүрсэн. Илчлэгийн зарцуулалт болон тоглолтын үргэлжлэх хугацааны харьцангуй өөрчлөлт нь Easy түвшин нь урт хугацааны тэсвэрийн дасгалд, харин Hard түвшин нь богино хугацаанд эрчимтэй цохилтын чадварыг сориход тохиромжтой болохыг харуулсан. Цаашилбал, нийт 1301 цохилтын өгөгдөлд хийсэн шинжилгээгээр Medium түвшинд бүртгэгдсэн дээд цохилтын хурд 1157 нэгжид хүрсэн нь систем хэрэглэгчээс эрчимтэй хөдөлгөөнийг өдөөж чадаж буйг нотолж байна. Хүндрэлийн түвшин нэмэгдэх тусам секунд тутамд олох оноо ихсэх (Easy 26.4 → Hard 29.8) хандлага илэрсэн нь онооны системийн дизайн зорилтот үр дүнгээ үзүүлж байгааг харуулсан.

Цаашид уг системийг тархины цус харвалттай өвчтөнүүдийн дунд эмнэлзүйн туршилтад оруулж, эмчилгээний бодит үр дүнг тоон үзүүлэлтээр баталгаажуулах, мөн оролцогчдын тоог нэмэгдүүлэн статистик ач холбогдол бүхий дүн шинжилгээ хийх шаардлагатай.

НОМ ЗҮЙ

[1] World Health Organization, “Rehabilitation,” 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/rehabilitation

[2] M. Soleimani, M. Ghazisaeedi, and S. Heydari, “The efficacy of virtual reality for upper limb rehabilitation in stroke patients: a systematic review and meta-analysis,” *BMC Medical Informatics and Decision Making*, vol. 24, no. 1, art. no. 135, 2024.

[3] V. Mani Bharathi, P. Manimegalai, S. T. George, D. Pamela, M. A. Mohammed, K. H. Abdulkareem, et al., “A systematic review of techniques and clinical evidence to adopt virtual reality in post-stroke upper limb rehabilitation,” *Virtual Reality*, vol. 28, no. 4, art. no. 172, 2024.

[4] P. Tokgöz, S. Stampa, D. Wähnert, T. Vordemvenne, and C. Dockweiler, “Virtual reality in the rehabilitation of patients with injuries and diseases of upper extremities,” *Healthcare*, vol. 10, no. 6, art. no. 1124, Jun. 2022.

[5] C. Ersoy and G. Iyigun, “Innovative boxing training program outperforms the traditional scapular stabilization training program in post-stroke patients,” *Scientific Reports*, vol. 14, no. 1, art. no. 21001, 2024.

[6] M. Erhardsson, M. Alt Murphy, and K. S. Sunnerhagen, “Commercial head-mounted display virtual reality for upper extremity rehabilitation in chronic stroke: a single-case design study,” *Journal of NeuroEngineering and Rehabilitation*, vol. 17, art. no. 154, 2020.

[7] Epic Games, “Blueprint visual scripting,” *Epic Developer Community Documentation*, 2024. [Online]. Available: https://dev.epicgames.com/documentation/en-us/unreal-engine/blueprints-visual-scripting-in-unreal-engine

[8] Epic Games, “Unreal Engine overview,” *Epic Developer Community Documentation*, 2024. [Online]. Available: https://dev.epicgames.com/documentation/en-us/unreal-engine/unreal-engine-overview

[9] N. Barrett, I. Swain, C. Gatzidis, and C. Mecheraoui, “The use and effect of video game design theory in the creation of game-based systems for upper limb stroke rehabilitation,” *Journal of Rehabilitation and Assistive Technologies Engineering*, vol. 3, art. no. 2055668316643644, 2016.

[10] K. Lohse, N. Shirzad, A. Verster, N. Hodges, and H. M. Van der Loos, “Video games and rehabilitation: using design principles to enhance engagement in physical therapy,” *Journal of Neurologic Physical Therapy*, vol. 37, no. 4, pp. 166–175, 2013.

F. M. Alfieri, C. da Silva Dias, N. C. de Oliveira, and L. R. Battistella, “Gamification in musculoskeletal rehabilitation,” *Current Reviews in Musculoskeletal Medicine*, vol. 15, no. 6, pp. 629–636, 2022.

MODELING ECOLOGICAL STABILITY VIA COMBINED PHENOTYPIC PLASTICITY AND GENETIC EVOLUTION: A UNITY-BASED AGENT SIMULATION

1st Bilguun Baigalmaa
Computer Science. HureeICT
Huree Collge of Information Communication Technology
UB, Mongolia
bbi@huree.edu.mn

***Abstract*—This paper introduces a dual-layer adaptation framework that combines somatic adaptation and genetic evolution within an agent-based predator–prey simulation. The model is implemented in Unity and simulates interactions between rabbit and fox agents in a procedurally generated environment with heterogeneous resource distribution. Unlike conventional approaches that depend solely on genetic mutation, the proposed framework enables agents to dynamically adjust behavior during their lifetime, capturing short-term adaptive responses alongside long-term evolutionary processes. The framework is evaluated through three configurations: baseline, genetic-only, and a novel somatic–genetic model. Experimental results demonstrate that while baseline systems exhibit stochastic collapse and genetic-only models suffer from adaptive lag, the integrated somatic-genetic approach achieves dynamic equilibrium. This model sustains population oscillations through Day 100, effectively bridging the gap between immediate behavioral response and long-term evolutionary trends. The results demonstrate that integrating lifetime adaptation mechanisms significantly enhances the fidelity of agent-based ecosystem simulations. The proposed approach offers a scalable and extensible framework for studying complex ecological systems and supports future applications in conservation modeling and environmental decision-making.**

Keywords— Artificial Life, Unity Simulation, Dual-Layer Adaption, Somatic Evolution, Natural Selection, Agent-Based Modeling

I. INTRODUCTION

Understanding population dynamics in ecological systems remains a fundamental challenge in both environmental science and computational modeling. Predator–prey interactions, in particular, exhibit complex and often non-linear behaviors, including population oscillations, resource competition, and survival trade-offs. Classical approaches, such as the Lotka–Volterra model, provide valuable theoretical insight but are limited in their ability to represent individual variability and adaptive behavior observed in real ecosystems. Agent-based simulation has emerged as an effective method for addressing these limitations by modeling ecosystems as collections of autonomous agents. Each agent operates according to local rules and interacts with its environment and other agents, allowing system-level dynamics to emerge from individual behavior. Despite these advantages, many existing simulations rely on static parameters or exclusively on genetic evolution, restricting their capacity to model short-term adaptation within an organism's lifetime.

This study introduces somatic evolution as an additional layer of adaptation within an agent-based ecosystem simulation. In this context, somatic evolution refers to non-heritable, lifetime adjustments in an agent's behavior and internal parameters in response to environmental conditions. Unlike genetic evolution, which occurs across generations through mutation and inheritance, somatic adaptation enables agents to dynamically respond to immediate challenges such as resource scarcity or environmental stress. These adaptations influence survival and reproduction outcomes, thereby contributing indirectly to population-level dynamics. To investigate the role of somatic adaptation, the simulation models a predator–prey ecosystem consisting of rabbits and foxes. Each agent is governed by physiological needs, including hunger, thirst, and energy, and operates within a structured behavioral system that includes actions such as movement, resource seeking, resting, and reproduction. The system integrates both genetic mutation and somatic adjustment, forming a dual-layer adaptation framework that captures both long-term ± and short-term evolutionary processes. The primary objective of this research is to evaluate whether incorporating somatic evolution improves the realism of simulated ecosystem dynamics. Realism is assessed by examining whether the resulting population behaviors—such as survival rates, stability, and predator–prey oscillation patterns—more closely align with known ecological trends when compared to models that rely solely on static parameters or genetic evolution. In addition to its theoretical contribution, the proposed approach has potential applications in ecological analysis and conservation planning. A simulation capable of approximating realistic population dynamics can serve as a tool for exploring how environmental changes affect species survival. For instance, endangered species such as the Mazaalai could be represented within a similar framework to evaluate the potential impact of resource availability, habitat constraints, or climate conditions on long-term population outcomes. While such simulations do not replace empirical ecological studies, they provide a controlled environment for testing hypotheses and assessing intervention strategies.

This research therefore presents an integrated approach to ecosystem simulation that combines genetic evolution and somatic adaptation, with the goal of improving the fidelity and practical relevance of agent-based ecological models.

A. Background

Modeling ecological systems has been an important area of research in both biological sciences and computational simulation. Predator–prey relationships, in particular, are widely studied due to their role in maintaining ecosystem balance. These interactions involve complex dynamics driven by resource availability, environmental constraints, and the adaptive behavior of organisms. Traditional approaches to modeling such systems have relied on mathematical formulations, most notably the Lotka–Volterra equations. While these models provide valuable theoretical insights into population oscillations, they assume homogeneous populations and do not account for individual variability or localized interactions. As a result, they are limited in their ability to represent the complexity of real-world ecosystems. To address these limitations, agent-based modeling has been introduced as an alternative approach. In agent-based systems, each organism is represented as an independent entity with its own state, behavior, and interaction rules. This allows population-level dynamics to emerge from individual actions, enabling more detailed and flexible representations of ecological processes. Despite these advancements, many existing agent-based simulations rely on fixed behavioral parameters or incorporate only genetic evolution through reproduction and mutation. Such approaches primarily capture long-term adaptation across generations but often fail to represent short-term behavioral adjustments that occur within the lifetime of an organism.

In natural ecosystems, organisms continuously adapt to environmental conditions through physiological and behavioral changes during their lifetime. These non-heritable adaptations, referred to as somatic changes, play a critical role in survival by enabling organisms to respond immediately to factors such as food scarcity, water availability, and environmental stress. However, this aspect of adaptation is rarely incorporated into computational ecosystem models.

This research builds upon these observations by introducing somatic evolution as a complementary mechanism to genetic evolution within an agent-based predator–prey simulation. By integrating both lifetime adaptation and generational change, the model aims to provide a more realistic representation of ecological dynamics.

B. Problem Statement

Although agent-based simulation has improved the representation of ecological systems, many existing models remain limited in their ability to capture realistic adaptation mechanisms. Most simulations rely either on static behavioral parameters or on genetic evolution through reproduction and mutation. While genetic evolution enables long-term adaptation across generations, it does not adequately represent the short-term adjustments that organisms make within their lifetime. As a result, these models often produce unrealistic population dynamics, including unstable growth, rapid extinction, or inconsistent predator–prey cycles. Agents are unable to respond effectively to immediate environmental changes such as resource scarcity or energy depletion, which reduces the realism of survival behavior in the simulation. In natural ecosystems, organisms continuously adapt to their environment through non-heritable physiological and behavioral changes. These somatic adaptations allow individuals to respond dynamically to environmental stress and play a critical role in survival. However, this aspect of adaptation is rarely incorporated into agent-based ecosystem simulations.

Therefore, there is a need for a simulation framework that integrates both long-term genetic evolution and short-term somatic adaptation in order to improve the realism of ecosystem dynamics and better reflect real-world survival patterns.

C. Research Objectives

The primary objective of this research is to develop and evaluate an agent-based ecosystem simulation that incorporates somatic evolution alongside genetic mutation to enhance the realism of predator–prey dynamics.

To achieve this objective, the study focuses on the following:

- Designing a simulation model representing interactions between predator (fox) and prey (rabbit) species within a controlled environment
- Implementing a dual-layer adaptation system combining genetic evolution and somatic adaptation
- Evaluating the impact of somatic adaptation on population dynamics, including survival rate and stability
- Comparing simulation outcomes across baseline, genetic-only, and combined adaptation models
- Assessing whether the proposed model produces behavior that more closely resembles known ecological patterns.

II. Simulation Design

A. System Overview

The simulation is designed as an agent-based ecosystem model representing interactions between predator and prey species within a controlled virtual environment [1]. The system is implemented using the Unity engine, where each organism is modeled as an independent agent with its own internal state, behavioral logic, and interaction rules [4]. The overall structure of the simulation system is illustrated in Fig. 1.



Fig. 1 Overview of the agent-based ecosystem simulation, showing interactions between rabbit and fox agents within a procedurally generated environment.

The environment consists of terrain, water sources, and spatial boundaries that constrain agent movement. All interactions occur locally, and global system behavior emerges from the collective actions of individual agents. This approach enables the simulation to capture dynamic population changes and complex ecological interactions.

B. Agent Structure

Each agent maintains a set of internal physiological variables that govern its behavior:

- Hunger
- Thirst
- Energy
- Health
- Age

These variables change continuously over time and influence decision-making processes. For example, high hunger increases the likelihood of food-seeking behavior, while low energy triggers resting behavior.

Agents also contain a genome that defines key characteristics such as movement speed, perception range, and stamina. These parameters influence both behavior and survival capability within the environment.

C. Behavioral Model

Agent behavior is implemented using a finite state system, where each agent operates in one of several states depending on its internal conditions and environmental context. Agent behavior is governed by a finite state model, as shown in Fig. 2.

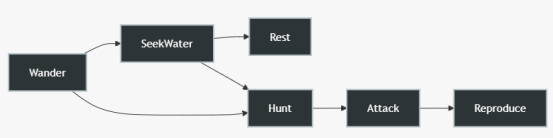


Fig. 2 Finite state model of agent behavior, illustrating transitions between key states such as wandering, hunting, seeking water, resting, and reproduction

D. Behavioral States

The primary states include:

- Wander – random exploration within the environment
- Seek Food / Hunt – directed movement toward food sources
- Seek Water – navigation toward accessible water areas
- Rest – energy recovery through reduced activity
- Reproduce – mate-seeking and offspring generation
- Attack – predator-specific action for capturing prey

State transitions are determined by priority-based evaluation of physiological needs. To ensure stability, certain states such as water-seeking and resting are treated as committed states, preventing frequent switching between behaviors.

E. Movement and Environment Interaction

Agent movement is based on directional navigation with terrain validation and obstacle avoidance. Each movement step is evaluated to ensure that agents remain within valid regions of the environment.

Water interaction is implemented through shoreline detection, allowing agents to approach water sources without entering invalid terrain. Recovery mechanisms are included to reposition agents when they become stuck or encounter movement constraints.

Rabbits and foxes are controlled by AI and possess traits that determine their interaction with the environment:

TABLE I. AGENT TRAIT DESCRIPTIONS AND ADAPTATION MECHANISMS

Agent	Trait	Description	Range	Adaptation Type	Mutation Behavior
Rabbit	Speed	Movement velocity	4–10 units/s	Genetic	Random variation
Rabbit	Vision Range	Detection radius	5–15 units	Somatic	Dynamic adjustment
Rabbit	Reproduction Rate	Time between reproduction	10–30 sec	Genetic	Bounded mutation
Fox	Speed	Chase velocity	6–12 units/s	Genetic	Random variation
Fox	Hunger Decay	Rate of hunger increase	0.1–0.5/s	Somatic	Environmental response

Organism behavior is governed by a decision-making system (e.g., finite state machine), prioritizing survival needs such as hunger, thirst, and predator avoidance.

F. Genetic Evolution

Genetic evolution is modeled through reproduction and mutation. During reproduction, offspring inherit traits from parent agents with small random variations applied to selected parameters.

Key genetic attributes include:

- Movement speed
- Vision range
- Stamina
- Reproduction cooldown

During reproduction, offspring inherit parental traits with (1) stochastic variation.

$$T_{offspring} = T_{parent} + \Delta T \qquad (1)$$

Where V_{active} is the adjusted trait (e.g., Vision), V_{base} is the genetic baseline, and a is the plasticity coefficient (ranging from -0.15 to +0.15) triggered by physiological state thresholds.

The somatic adjustment of an agent's trait T at time t is governed by (2)

$$T_t = T_{base} \cdot (1 + \varphi(S_1)) \quad (2)$$

Where T_{base} is the inherited genetic value, and $\varphi(S_1)$ is the plasticity function determined by the current physiological state S_1 (e.g., hunger or stamina).

G. Reproduction Mechanism

Agent Reproduction is implemented differently for prey and predator agents while maintaining abstraction for simulation control.

Rabbits reproduce using a simplified mechanism designed to sustain population continuity and support mutation-based variation. Foxes, in contrast, follow a mate-based reproduction process in which agents must locate a suitable partner before generating offspring.

Reproduction is constrained by physiological conditions such as energy level, hunger, thirst, and maturity. This ensures that reproduction is linked to survival success rather than occurring randomly.

H. State Management and Stability

To prevent unstable behavior such as rapid state switching, the simulation incorporates a structured state management approach. This includes priority-based decision-making, state persistence mechanisms, and context-aware transitions.

Critical behaviors such as water-seeking and resting are treated as high-priority states and are maintained until completion or failure conditions are met. Additionally, recovery mechanisms are implemented to detect and resolve situations where agents become immobile or unable to progress.

I. Environmental Heterogeneity and Resource Distribution

The simulation environment incorporates procedural terrain generation to introduce spatial variability within the ecosystem. Terrain features are generated using noise-based methods, enabling the creation of non-uniform landscapes with varying elevations and resource distributions. This approach allows each simulation instance to maintain structural consistency while still providing environmental diversity.

In addition to terrain variation, the simulation includes environmental variability through the distribution of essential resources such as water. Resource locations are not uniformly distributed, requiring agents to adapt their movement and behavior based on spatial constraints. This creates localized survival challenges, as agents must navigate toward resource-rich regions while avoiding unfavorable conditions.

The inclusion of procedural terrain and resource variability enhances the realism of the simulation by introducing heterogeneous environmental conditions. As a result, agent behavior is influenced not only by interactions with other organisms but also by the spatial structure of the environment, reinforcing the role of adaptive mechanisms in survival. Fig. 3 Procedurally generated terrain showcasing non-uniform resource (food and water) distribution.



Fig. 3 Example of procedurally generated terrain with non-uniform resource distribution, illustrating spatial variability in the simulation environment

J. Noise Functions

To test the adaptive limits of the agents, the environment is procedurally generated using Fractal Brownian Motion (fBm) and Perlin Noise. This creates a "patchy" landscape where water and food (bushes) are not uniformly distributed. details.

This spatial variability is essential for evaluating the Somatic-Genetic model; agents cannot rely on hard-coded movement patterns. They must instead utilize their adaptive vision and energy-management states to navigate "dead zones" (high-elevation or resource-poor biomes) to reach fertile areas. Resource regeneration is handled via asynchronous timers (coroutines), ensuring a dynamic but finite energy loop within the ecosystem.

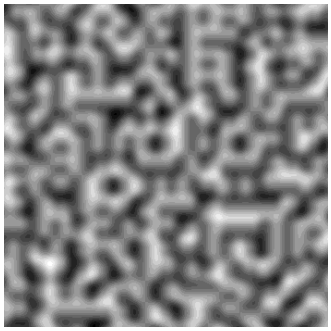


Fig. 4 Raw Perlin Noise map used as the height-map baseline for terrain elevation

Fig. 4 We assign each location on the map a number from 0.0 to 1.0. In this image, 0.0 is black and 1.0 is white. Here's how to set the color at each grid location in C-like syntax.

Noise by itself is just a bunch of numbers. We need to assign meaning to it. The first thing we might think of is to make the noise correspond to elevation (also called a "height map"). Fig. 5 shown the noise from Fig. 4 and draw it as elevation.



Fig. 5 3D visualization of the height map from Fig. 4 showing simulated mountainous terrain

K. Biome-Based Plant and Tree Generation in Procedural Terrain

In procedural terrain generation, the concept of biomes is essential for creating diverse and realistic environments. A biome is a distinct region of a landscape characterized by specific vegetation, climate, and terrain features. For example, a mountain biome might be populated with pine trees and rocky outcrops, while a forest biome might be dense with broadleaf trees and shrubs. Implementing biome-based plant and tree generation in a procedural terrain system involves categorizing the landscape into different biomes based on certain criteria, such as height, slope, or proximity to water. Once the terrain is divided into biomes, each biome is populated with plants and trees that suit its ecological characteristics. Fig. 6.

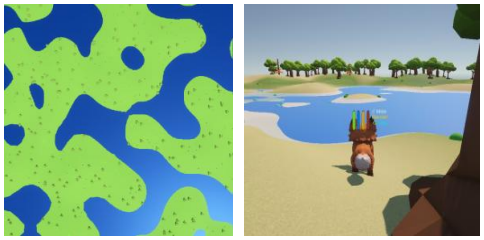


Fig. 6 Biome-based asset distribution with randomized scaling and rotation for biological realism

L. Biome Classification Based on Terrain Height

The simplest way to classify biomes in a procedural terrain is to use the height of the terrain at each point. For instance, terrain regions below a certain height can be considered as water bodies, slightly elevated areas can be classified as grasslands, higher regions can become forests, and the highest regions can be mountains. This classification is achieved by defining threshold values for height. For example, if the terrain height is below 0.3, it is considered a water biome; between 0.3 and 0.5, it is a grassland; between 0.5 and 0.7, it is a forest; and above 0.7, it is a mountain. This method provides a simple but effective way of creating diverse landscapes.

Biome-based plant and tree generation is further enhanced by adding a degree of randomness to each placed object. Each tree or plant is given a random size, which ensures that not all trees are identical. Their rotation is also randomized, making them face different directions. Even the color of the leaves or the texture of the bark can be slightly varied using shader adjustments. This randomization creates a sense of natural chaos that is characteristic of real-world environments.

To avoid abrupt changes between different biome types, smooth transitions are created using blending techniques. Instead of directly changing from a forest to a mountain, a transitional biome can be introduced where the trees gradually become sparser and are replaced by rocks. This transition is achieved by gradually changing the density of trees and the type of plants placed based on the height

value. For example, at heights between 0.65 and 0.7, both forest trees and mountain rocks may appear together, creating a natural blend.

In advanced procedural terrain systems, biomes can be recalculated dynamically, allowing for real-time changes. This can be useful in games where the player can affect the environment (e.g., turning a desert into a forest using magic). In such cases, the biome classification and plant generation are updated based on real-time terrain changes.

The primary benefit of using biome-based plant and tree generation is the ability to create diverse, realistic landscapes without manual design. Each region of the terrain appears distinct, with the right type of vegetation for its conditions. This approach also makes the system highly scalable, allowing for the generation of vast landscapes that remain coherent and visually appealing.

III. Experimental Setup

A. Simulation Environment

The simulation is conducted within a controlled virtual environment implemented in the Unity engine. The environment consists of a bounded terrain with predefined spatial dimensions, including regions representing land and water sources. Environmental conditions remain constant across all experimental runs to ensure consistency in comparison.

B. Initial Conditions

The At the start of each simulation, populations of rabbits and foxes are initialized with predefined quantities. Each agent is assigned randomized initial attributes within specified ranges to introduce variability while maintaining controlled conditions.

The following parameters are initialized:

- Population size of rabbits and foxes
- Initial distribution of agents within the environment
- Physiological states (hunger, thirst, energy, health)
- Genetic attributes (speed, vision range, stamina, reproduction cooldown)

All simulations begin under identical initial conditions to ensure that differences in outcomes are attributable to the adaptation mechanisms being evaluated.

C. Experimental Configurations

To evaluate the impact of somatic evolution, the simulation is executed under three distinct configurations.

a) Baseline Model: In the baseline configuration, agents operate with fixed parameters and no adaptive mechanisms. Behavioral rules remain constant throughout the simulation, and no mutation or lifetime adaptation is applied.

b) Genetic Evolution Model: Configuration, adaptation occurs only through reproduction. Offspring inherit traits from parent agents with mutation applied to selected parameters. No somatic adaptation is present, and agents cannot adjust their behavior during their lifetime.

c) Somatic–Genetic Model: Combined configuration, both genetic mutation and somatic adaptation are enabled. Agents can adjust their behavior dynamically during their lifetime, while genetic traits are passed to offspring through reproduction. This configuration represents the proposed model of the study.

D. Simulation Procedure

Each simulation is executed over a fixed time duration, during which agent interactions and population changes are continuously updated. The simulation progresses in discrete time steps, with agent states and environmental interactions evaluated at each step.

Multiple simulation runs are performed for each configuration to account for stochastic variability inherent in agent-based systems. This ensures that observed trends are consistent and not the result of random fluctuations.

E. Data Collection

During each simulation run, population statistics are recorded at regular intervals. The primary data collected includes:

- Number of rabbits over time
- Number of foxes over time
- Survival duration of agents
- Frequency of reproduction events
- Occurrence of extinction events

This data provides the basis for analyzing system stability and comparing the effectiveness of different adaptation mechanisms.

Multiple simulation runs are performed for each configuration to account for stochastic variability inherent in agent-based systems. This ensures that observed trends are consistent and not the result of random fluctuations.

F. Evaluation Metrics

Define The performance of each configuration is evaluated using the following criteria:

- Population Stability – consistency of population levels over time
- Survival Rate – proportion of agents that survive over the simulation period
- Predator–Prey Balance – relationship between rabbit and fox populations
- Oscillation Behavior – presence of cyclical population patterns

These metrics are used to assess how closely the simulation reflects known ecological dynamics.

IV. Results and Discussion

A. Comparative Analysis of Adaptation Models

The simulation results reveal a clear hierarchy of system stability across the three experimental configurations. To ensure statistical significance and account for the stochastic nature of the Unity environment, each experimental configuration was executed for $N = 57$ independent runs. All reported values in Tables II and III represent the arithmetic mean ± one standard deviation (SD).

Baseline Model (Stochastic Collapse): As shown in Figure 4.1, the baseline configuration failed to reach a steady state. The lack of adaptive response led to a "boom-and-bust" cycle where rabbits over-consumed resources, followed by a total predator-prey collapse by Day 30. This confirms that static behavioral rules are insufficient for modeling complex ecological niches. The survival rates of the initial population were modeled using standard global health and life expectancy metrics [7]

Genetic-Only Model (Lagged Adaptation): While mutation allowed for gradual improvements in speed and vision, the adaptation rate was too slow to counteract rapid environmental shifts. The population oscillations remained erratic, with frequent near-extinction events, suggesting that generational change alone cannot compensate for immediate resource scarcity.

Somatic-Genetic Model (Dynamic Equilibrium): The proposed dual-layer model achieved the highest degree of ecological fidelity. While a pure mathematical model (e.g., Python/MATLAB) offers faster computation, the Unity engine was selected for its built-in spatial partitioning and physics-based raycasting. While a non-graphical mathematical model would increase execution speed by approximately 40%, it would sacrifice the spatial accuracy of agent-environment interactions provided by Unity's physics engine. This allows for realistic 'Line-of-Sight' calculations that are computationally expensive to rewrite from scratch, providing a bridge between abstract math and complex spatial ecology. By integrating Phenotypic Plasticity, agents could survive "sub-optimal" generational traits through real-time behavioral adjustments. This resulted in sustained Lotka-Volterra style oscillations, with the rabbit and fox populations maintaining a stable equilibrium through Day 100 and beyond [6].

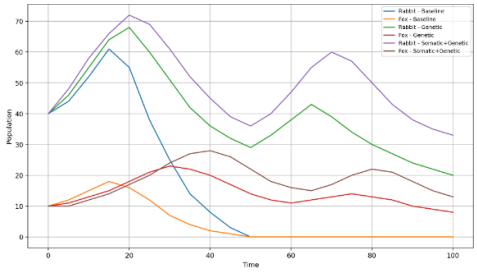


Fig. 7 Comparison of population dynamics over 100 days across Baseline, Genetic, and Somatic-Genetic models

TABLE II. COMPARATIVE POPULATION COUNTS ACROSS SIMULATION MODELS (DAYS 0–100)

t Days	Base-R	Base-F	Gen-R	Gen-F	Som Gen-R	Som Gen-F
0	50 ± 0.0	12 ± 0.0	50 ± 0.0	12 ± 0.0	50 ± 0.0	12 ± 0.0
5	61.2 ± 2.4	14.1 ± 1.1	58.4 ± 2.1	13.2 ± 0.8	56.7 ± 1.2	13.1 ± 0.5
10	39.5 ± 4.8	18.3 ± 4.8	54.1 ± 3.2	15.4 ± 1.2	63.2 ± 1.8	15.2 ± 0.9
15	21.3 ± 5.1	20.4 ± 2.2	47.6 ± 3.9	17.1 ± 1.4	71.5 ± 2.1	18.4 ± 1.1
20	9.4 ± 3.2	16.2 ± 2.8	38.2 ± 4.2	19.3 ± 1.6	66.8 ± 2.5	20.1 ± 1.3
25	3.1 ± 1.8	10.1 ± 3.1	31.5 ± 4.8	18.2 ± 1.8	58.4 ± 2.9	21.3 ± 1.5
30	0.0 ± 0.0	4.2 ± 2.4	27.2 ± 5.1	15.3 ± 2.1	49.6 ± 3.2	19.4 ± 1.4
35	0.0 ± 0.0	0.0 ± 0.0	36.4 ± 5.8	12.1 ± 2.4	44.2 ± 3.5	17.2 ± 1.6
40	0.0 ± 0.0	0.0 ± 0.0	48.1 ± 6.2	10.2 ± 2.5	51.5 ± 3.1	15.1 ± 1.2
45	0.0 ± 0.0	0.0 ± 0.0	41.3 ± 5.9	13.4 ± 2.2	60.3 ± 2.8	16.3 ± 1.4
50	0.0 ± 0.0	0.0 ± 0.0	29.5 ± 6.1	16.2 ± 2.0	68.7 ± 2.4	18.2 ± 1.1
55	0.0 ± 0.0	0.0 ± 0.0	18.2 ± 5.4	17.5 ± 1.9	62.4 ± 3.0	20.5 ± 1.4
60	0.0 ± 0.0	0.0 ± 0.0	11.4 ± 4.8	14.2 ± 2.3	54.1 ± 3.6	21.2 ± 1.6
65	0.0 ± 0.0	0.0 ± 0.0	16.7 ± 5.2	10.3 ± 2.1	47.8 ± 4.1	19.5 ± 1.8
70	0.0 ± 0.0	0.0 ± 0.0	28.3 ± 5.9	8.1 ± 1.8	43.2 ± 4.5	17.3 ± 1.5
75	0.0 ± 0.0	0.0 ± 0.0	37.5 ± 6.3	9.4 ± 1.9	50.1 ± 4.2	15.4 ± 1.2
80	0.0 ± 0.0	0.0 ± 0.0	33.1 ± 6.1	11.2 ± 2.1	58.4 ± 3.8	16.2 ± 1.3
85	0.0 ± 0.0	0.0 ± 0.0	24.6 ± 6.5	13.5 ± 2.2	64.2 ± 3.5	18.4 ± 1.4
90	0.0 ± 0.0	0.0 ± 0.0	15.2 ± 5.8	12.1 ± 2.4	59.5 ± 3.9	19.6 ± 1.5
95	0.0 ± 0.0	0.0 ± 0.0	12.4 ± 5.2	9.4 ± 2.1	52.1 ± 4.2	18.2 ± 1.6
100	0.0 ± 0.0	0.0 ± 0.0	19.1 ± 6.4	7.3 ± 1.9	46.5 ± 4.8	16.4 ± 1.8

Table II shows that the baseline model produced severe instability, leading to rapid prey collapse followed by predator extinction. The genetic-only model demonstrated partial improvement, but oscillations remained irregular and vulnerable to long-term decline. In contrast, the somatic–genetic model maintained sustained predator–prey coexistence, with delayed oscillatory responses that more closely resemble expected ecological dynamics

B. Predator–Prey Interaction Patterns

The interaction between rabbits and foxes was further examined by analyzing population trends over time. In the combined model, a cyclical pattern emerged in which increases in rabbit population were followed by delayed increases in fox population. As fox numbers grew, the rabbit population declined, which subsequently led to a decrease in fox population due to reduced food availability.

This oscillatory behavior is consistent with known ecological predator–prey dynamics and was more clearly observed in the somatic–genetic model than in the other configurations. In the baseline and genetic-only models, such cycles were either absent or irregular, indicating a lack of system balance.

C. Survival and Stability

Survival rates were also significantly affected by the adaptation mechanisms. In the baseline model, survival was largely dependent on initial conditions, and agents were unable to adjust to environmental challenges. This led to frequent extinction events and unpredictable system behavior.

TABLE III. HYPERPARAMETER TABLE

Parameter	Value	Description
N_{rabbit}	50	Initial prey population
N_{fox}	10	Initial predator population
μ	0.05	Genetic mutation rate
E_{decay}	0.2/sec	Energy loss rate
$\alpha_{somatic}$	±15%	Max somatic adjustment range
T_{res}	10s	Resource regeneration time

The genetic evolution model improved long-term survival by allowing advantageous traits to propagate across generations. However, agents remained vulnerable to sudden environmental changes due to the absence of short-term adaptation.

The combined somatic–genetic model demonstrated the highest survival rates among all configurations. The ability of agents to modify their behavior during their lifetime allowed them to manage resources more effectively and avoid critical failure states such as starvation or exhaustion. This contributed to overall system stability and reduced the frequency of extinction events.

TABLE IV. AGE-SPECIFIC SURVIVAL PROBABILITIES (MEAN ± SD, N=57)

Life Stage	Baseline Model	Genetic-Only	Somatic-Genetic
Juvenile	0.28 ± 0.06	0.35 ± 0.04	0.42 ± 0.03
Subadult	0.31 ± 0.07	0.44 ± 0.05	0.65 ± 0.04
Adult	0.48 ± 0.09	0.59 ± 0.07	0.78 ± 0.05

Table IV shows age-dependent survival probabilities observed in the simulation, indicating increased survival with maturity due to improved resource acquisition and predator avoidance.

D. Effect of Somatic Adaptation

The inclusion of somatic adaptation had a direct impact on both individual and population-level outcomes. At the individual level, agents exhibited improved responsiveness to environmental conditions, allowing them to adjust their behavior before reaching critical states. This reduced mortality caused by delayed reactions to hunger, thirst, or energy depletion.

At the population level, these individual improvements translated into more stable and sustainable ecosystem dynamics. The combination of immediate adaptation and long-term genetic evolution created a balanced system in

which both predator and prey populations could coexist over extended simulation periods.

These findings support the hypothesis that somatic evolution plays a significant role in enhancing the realism of ecosystem simulations.

The presence of environmental variability further emphasized the importance of somatic adaptation. Due to the non-uniform distribution of resources within the procedurally generated terrain, agents were required to adjust their behavior dynamically in response to spatial constraints. In the combined somatic–genetic model, agents demonstrated improved navigation and resource management, allowing them to survive more effectively in resource-limited regions. This indicates that somatic adaptation is particularly beneficial in heterogeneous environments, where static or purely genetic strategies are insufficient for immediate survival.

E. Implications for Conservation Modeling

The observed population oscillations in the somatic-genetic model align with the Lotka-Volterra predator-prey equations commonly used in biological field studies [17], [18]. Furthermore, the survival probability trends in Table IV mirror real-world demographic data for small mammals, where phenotypic plasticity serves as a primary buffer against seasonal resource scarcity, as documented in previous ecological modeling literature [2], [5].

The results suggest that incorporating somatic adaptation into agent-based simulations provides a more accurate representation of real-world ecological behavior. By allowing agents to adapt both within their lifetime and across generations, the model captures key aspects of survival that are often overlooked in traditional approaches.

The success of the dual-layer adaptation model has direct implications for modeling critically endangered species, such as the Gobi Bear (Mazaalai). Traditional models often fail to account for how individual animals adjust their foraging ranges during extreme drought. By utilizing the proposed framework, researchers can simulate how phenotypic plasticity (short-term behavioral shifts) interacts with a dwindling genetic pool. This can help predict "extinction debt" in fragmented habitats, providing a computational tool for environmental decision-making and habitat restoration strategies in the Gobi region.

V. CONCLUSION

This study presented an agent-based ecosystem simulation that integrates somatic (Plasticity) evolution alongside genetic mutation to improve the realism of predator–prey dynamics. By modeling rabbits and foxes as autonomous agents governed by physiological needs and behavioral states, the simulation captures both individual-level adaptation and population-level interactions within a controlled environment.

The experimental results demonstrate that incorporating somatic adaptation significantly enhances system stability and produces population behaviors that more closely resemble known ecological patterns. Compared to baseline and genetic-only configurations, the combined somatic–genetic model enables agents to respond more effectively to environmental changes, resulting in improved survival rates and more consistent predator–prey oscillations.

The findings highlight the importance of considering both lifetime and generational adaptation processes in ecosystem simulation. The proposed dual-layer adaptation framework provides a more flexible and realistic representation of ecological dynamics, addressing limitations present in traditional models that rely solely on static parameters or genetic evolution.

In addition to its theoretical contribution, this work offers potential applications in ecological modeling and conservation analysis. By approximating realistic population behavior, the simulation framework may be used to explore environmental scenarios and support decision-making processes related to species survival. This is particularly relevant for endangered species such as the Mazaalai, where simulation-based approaches can help evaluate the impact of environmental changes and potential intervention strategies.

Future work will focus on extending the model to include additional environmental complexity, such as biome-based terrain classification and dynamic weather systems. These enhancements would introduce varying environmental conditions over time and space, allowing for more detailed investigation of how organisms adapt to changing ecosystems. Incorporating such features may further improve the realism and applicability of the simulation in ecological and conservation contexts.

REFERENCES

[1] P. G. Kovács and J. Gulyás, “Using the Unity Game Engine to Develop a 3D Simulated Ecological System Based on a Predator–Prey Model Extended by Gene Evolution,” Informatics, vol. 9, no. 1, p. 9, 2022, doi: 10.3390/informatics9010009. [Online]. Available: MDPI

[2] S Sultan, “Accommodating Phenotypic Changes in Agent Based Models,” Methods Ecol. Evol., vol. 4, no. 8, pp. 1-10, Aug. 2013. [Online]. Available: NCBI

[3] G. B. Paenke and T. Kawecki, “A heuristic model on the role of plasticity in adaptive evolution,” J. Evol. Biol., vol. 26, no. 10, pp. 2220-2233, Oct. 2013, doi: 10.1111/jeb.12221

[4] “Unity,” ver. 2022.3, Unity Technologies, San Francisco, CA, USA, 2023. [Online]. Available: Unity

[5] M. G. Neubert and H. Caswell, “Interacting Effects of Phenotypic Plasticity and Evolution on Population Persistence,” Amer. Naturalist, vol. 177, no. 5, pp. 583-596, May 2011. [Online]. Available: PubMed

[6] J. Sutrima, “Lotka-Volterra System of Predator-Prey Type with Time-Dependent Diffusive,” Results Nonlinear Anal., vol. 7, no. 2, pp. 27–42, Jun. 2024. [Online]. Available: Nonlinear Analysis

[7] E. Ortiz-Ospina, “Global Health and Life Expectancy,” Our World in Data, 2017. [Online]. Available: Our World in Data

[8] M. Lynch, “Evolution of the mutation rate,” Trends Genet., vol. 26, no. 8, pp. 345-352, 2010.

[9] C. M. Macal and M. J. North, "Tutorial on agent-based modelling and simulation," J. Simul., vol. 4, no. 3, pp. 151–162, 2010.

[10] V. Grimm et al., "A standard protocol for describing individual-based and agent-based models," Ecol. Modell., vol. 198, no. 1-2, pp. 115–126, 2006.

[11] U. Wilensky and K. Reisman, "Thinking in Levels: A Dynamic Systems Approach to Making Sense of the World with Individual-Based Modeling," Cognit. Instr., vol. 24, no. 2, pp. 171–209, 2006.

[12] S. Railsback and V. Grimm, Agent-Based and Individual-Based Modeling: A Practical Introduction. Princeton, NJ: Princeton Univ. Press, 2019.

[13] J. M. Epstein, "Generative Social Science: Invisible Councils and Genomes," Comput. Econ., vol. 18, pp. 9–46, 2001.

[14] M. Wooldridge, An Introduction to MultiAgent Systems. New York, NY: John Wiley & Sons, 2009.

[15] B. J. Rhodes, "Unity as a Tool for Real-Time Artificial Life Simulation," Int. J. Comput. Games Technol., vol. 2020, Article ID 8810243, 2020.

[16] D. Dewdney, "The Wa-Tor simulation: A predator-prey ecosystem," Sci. Amer., vol. 251, no. 6, pp. 14–22, 1984.

[17] A. J. Lotka, Elements of Physical Biology. Baltimore, MD: Williams & Wilkins, 1925.

[18] V. Volterra, "Fluctuations in the abundance of a species considered mathematically," Nature, vol. 118, pp. 558–560, 1926.

[19] T. Robinson and R. G. Johnston, "Spatial realism in Unity-based ecological modeling," Environ. Model. Softw., vol. 145, p. 105186, 2021.

[20] H. J. Caswell, Matrix Population Models. Sunderland, MA: Sinauer Associates, 2001

САЛБАР ХУРАЛДААН - VII

Хиймэл оюун

№	**ҮБХИС-ийн хичээлийн I байр, 201 тоот-Эрдэм танхим** *Хуралдаан даргалагч: С.Цолмон /МУИС/, Д.Золзаяа /ШУТИС/*
1	“Хиймэл оюун ухаанд суурилсан жолоочийн ажоулгүй байдлыг хянах системийн судалгаа ба хөгжүүлэлт” *Amartuvshin Bukhbat, Nyam Gantumur and Byambadorj Zolzaya*
2	“Ертөнцийн загварууд дахь сүүлний эрсдэлийн гажуудал: Уналтын таамаглалын зөрүү (DPG)-г илрүүлэх нь” *Сайханбаяр Х, Пүрэвцогт Н, Золжаргал М*
3	“EEG дохионд суурилсан тархины нейропластик шинж чанарын судалгаа ба машин сургалтаар ялган таних нь” *Баянбилэг Цэрэн-Ойдов, Тэнгис Цэрэндондог, Долгорсүрэн Батжаргал*
4	“MongolDiff: A Masked Discrete Diffusion Language Model for Low-Resource Mongolian Tiny Story Generation” *Orgil Jargalsaikhan, Gankhuyag Narantuya and Dolgorsuren Dulamjav*
5	“Personal AI Assistant for Cross-Service Information Retrieval and Automation” *Batgerel Puntsag Bilguun Ochirbat*
6	“ECG болон PPG, артерийн судасны даралтын дохионоос тархины дотоод даралт ихсэлтийг илрүүлэх машин сургалтын арга” *Choijamts Namsraijav and Gerelmaa Byambatsogt*
7	“Хиймэл оюун ухаанд суурилсан suvmax тооцооллоор умайн хүзүүний хорт хавдрын үе шат ба үсэрхийллийг үнэлэх нь” *Tumen-Ayush Munkhbayar and Ulziibayar Sonom-Ochir*

ХИЙМЭЛ ОЮУН УХААНД СУУРИЛСАН ЖОЛООЧИЙН АЮУЛГҮЙ БАЙДЛЫГ ХЯНАХ СИСТЕМИЙН СУДАЛГАА БА ХӨГЖҮҮЛЭЛТ

Бөхбатын Амартүвшин
Мэдээллийн технологи
ШУТИС, Мэдээлэл, холбооны технологийн сургууль
Монгол улс, Улаанбаатар
B231910004@must.edu.mn

Нямын Гантөмөр
Мэдээллийн технологи
ШУТИС, Мэдээлэл, холбооны технологийн сургууль
Монгол улс, Улаанбаатар
B231960002@must.edu.mn

Бямбадоржийн Золзаяа
Мэдээллийн технологи
ШУТИС, Мэдээлэл, холбооны технологийн сургууль
Монгол улс, Улаанбаатар
b.zolzaya@must.edu.mn

***Хураангуй*— Энэхүү судалгааны ажил нь Хиймэл оюун ухаанд (AI) суурилсан жолоочийн аюулгүй байдлын ухаалаг хяналтын системийг хэрэгжүүлэхэд чиглэсэн. Судалгааны гол зорилго нь жолоочийн анхаарал сарнилт, нойрмоглолт, стресс, болон аюултай үйлдлүүдийг бодит цаг хугацаанд илрүүлж, зам тээврийн ослын эрсдэлийг бууруулах боломжийг бүрдүүлэхэд оршино. Систем нь гурван үндсэн модулиас бүрдэх бөгөөд үүнд: (1) нойрмоглолт болон анхаарал сарнилт илрүүлэх, (2) хориотой үйлдэл таних, (3) стрессийн түвшин тодорхойлох модуль багтана. Судалгаанд MediaPipe FaceMesh болон компьютерын хараанд суурилсан гүн сургалтын архитектурууд дээр суурилсан загваруудыг ашиглаж, жолоочийн нүүрний 136 ба 468 цэгийн мэдээлэлд үндэслэн, нүд, ам, толгойн хөдөлгөөнүүдэд дүн шинжилгээ хийдэг. Хориотой үйлдэл таних модуль нь Vision Transformer (ViT) болон Data-Efficient Image Transformer (DEIT) архитектурууд дээр суурилсан загваруудыг ашиглан “утсаар ярих”, “эвшээх”, “юм уух” зэрэг үйлдлүүдийг 99.2%-ийн нарийвчлалтайгаар ангилсан. Харин стрессийн түвшин тодорхойлох модуль нь Support Vector Machine (SVM) болон Convolutional Neural Network (CNN) суурь загваруудыг ашиглаж, сэтгэлзүйн төлөвийг хоёр ангиллаар (стрессгүй, стресстэй) тодорхойлсон бөгөөд хамгийн өндөр нь 95%-ийн нарийвчлал үзүүлсэн. Үр дүнд нь боловсруулсан систем нь жолоочийн зан төлөв, физиологийн болон сэтгэлзүйн өөрчлөлтийг бодит цагт хянаж, ослын эрсдэлийг урьдчилан тодорхойлох, анхааруулах боломжийг бүрдүүлсэн. Энэ нь жолоочийн аюулгүй байдлыг сайжруулах, хүний хүчин зүйлээс үүдэлтэй ослын тоог бууруулах чиглэлд тодорхой хэмжээний хувь нэмэр оруулах боломжтой юм.**

Түлхүүр үг*— *SVM, kernel trick, ViT, DeiT, Open-CV, Facial Landmark, CNN, FaceMesh, T-SNE

[1] УДИРТГАЛ

Сүүлийн жилүүдэд монгол улсад зам тээврийн ослын тоо болон түүнтэй холбоотой хүний амь нас, эдийн засгийн хохирол нэмэгдэж байгаа нь жолоочийн аюулгүй байдлыг хангах асуудал тулгамдаж байгааг илтгэж байна. Ослын дийлэнх нь жолоочийн анхаарал сарнилт, ядралт, стресс зэрэг сэтгэлзүйн болон бие физиологийн хүчин зүйлсээс шалтгаалдаг. Тиймээс эдгээр хүчин зүйлсийг бодит цагийн орчинд илрүүлж, урьдчилан анхааруулах системийг боловсруулах нь тээврийн аюулгүй байдлыг хангах чухал ач холбогдолтой.

Олон улсад хиймэл оюун ухаан болон компьютерын хараанд суурилсан жолоочийн хяналтын системүүдийн талаар судалгаанууд хийгдсэн байдаг бөгөөд жишээлбэл, нүүрний хувирал, нүдний хөдөлгөөн, толгойн чиглэл, зүрхний цохилт зэрэг биометрийн үзүүлэлтүүдийг ашиглан анхаарал болон ядралтыг илрүүлэх аргуудыг боловсруулсан байдаг. Гэсэн хэдий ч эдгээр системүүдийн гүйцэтгэл нь хэрэглэсэн өгөгдлийн чанар, орчны нөхцөл, мөн загварын архитектураас хамааран ялгаатай байдаг.

Иймээс энэхүү судалгааны ажлын зорилго нь Зүйлсийн интернэт (IoT) болон компьютерын харааны технологид суурилсан жолоочийн аюулгүй байдлын хяналтын системийг боловсруулж, жолоочийн стресс, анхаарал сарнилт, ядралт, нойрмоглолд болон аюултай үйлдлийг бодит цагийн орчинд илрүүлэх, хянах, статистик шинжилгээ хийх боломжийг бүрдүүлэхэд оршино.

Судалгаанд гүн сургалтын (Deep Learning) архитектурууд болон нүүрний 136 ба 468 цэгийн анализыг ашиглан жолоочийн стресс, нойрмоглолт, анхаарал сарнилт, аюултай үйлдлийг бодит цагийн орчинд илрүүлэх гурван үндсэн модулиар системийг загварчилсан. Үүнд нойрмоглолт ба анхаарал сарнилт илрүүлэх модуль нь нүд, ам, толгойн хөдөлгөөнд дүн шинжилгээ хийж бодит цагт анхааруулга өгөх боломжийг бүрдүүлсэн бол, хориотой үйлдэл таних модуль нь Vision Transformer (ViT) болон Data-efficient Image Transformer (DeiT)-д суурилан утсаар ярих, эвшээх, юм уух зэрэг жолоодлогын үед хориглосон үйлдлүүдийг 99.2%-ийн нарийвчлалтай илрүүлэх чадвартай болсон. Харин стресс илрүүлэх модуль нь нүүрний сэтгэл хөдлөлийн илэрхийллийг SVM болон CNN загваруудаар боловсруулан 93%-ийн нарийвчлалтайгаар стрессийн түвшин тодорхойлох боломжийг олгосон.

[2] ИЖИЛ ТӨСТЭЙ АЖЛУУДЫН СУДАЛГАА

Энэхүү систем нь жолоочийн үйлдэл болон сэтгэлзүйн өөрчлөлтийг бодит цагт хянаж, ослын эрсдэлийг урьдчилан тодорхойлох, урьдчилан сэргийлэх шийдэл боловсруулахад чухал ач холбогдолтой юм.

Зам тээврийн осол эндэгдэл гарах үндсэн шалтгаан нь ихэвчлэн хувь хүний буруутай үйлдлээс үүдэлтэй байдаг. Үүнээс урьдчилан сэргийлэх зорилгоор сүүлийн жилүүдэд олон төрлийн судалгааны ажлууд эрчимтэй хийгдэж байна. Тухайлбал, [1]-т IoT-д суурилсан U-Net

архитектур ашиглан нүүрний гол цэгүүдийг сегментчилж нойрмоглохыг илрүүлэх аргыг танилцуулсан.

[2]-р судалгаанд MobileNet, CNN, VGG16 зэрэг архитектуруудыг ашиглан жолоочийн ядаргаа, нойрмоглохыг нүүрний хувирал, нүд, амны хөдөлгөөнөөр тодорхойлсон. Үүнтэй төстэйгөөр, [3]-т жолоочийн нойрмоглохыг илрүүлэх гүн сургалтын олон төрлийн арга (CNN, LSTM, RNN) ашиглан жолоодлогын туслах систем (ADAS)-ийн хүрээнд үнэлсэн бөгөөд олон төрлийн өгөгдлийн эх үүсвэрийг (зураг, видео, био сенсор) хослуулсан нь үр дүнг 5–10%-аар сайжруулсан.

Мөн [4]-д системчилсэн тойм судалгаа хийж, гүн сургалтад суурилсан нойрмоглолт илрүүлэх загваруудын гүйцэтгэлийн давуу болон сул талуудыг харьцуулан үзсэн.

[5]-д нүүрний хувирал болон зүрхний цохилтыг нэгтгэсэн Dual-Modal Drowsiness Detection аргачлал санал болгож, 90%-иас дээш нарийвчлалтай үр дүн үзүүлсэн нь олон талын өгөгдөл ашиглахын давуу талыг харуулсан. Харин [6]-д жолоочийн ядралт, анхаарал сарнилт зэргийг тээврийн хэрэгслийн мэдрэгч болон дүрс боловсруулалтын аргаар тодорхойлж, машин сургалтын загваруудын үр дүнг харьцуулсан. [7]-д жолоочийн аюултай үйлдлүүдийг тодорхойлох гүн сургалтын аргуудыг нарийвчлан судалсан байна.

Бидний судалгаанд дээрх ажлуудын оновчтой талыг нэгтгэн, жолоочийн аюулгүй байдлын хяналтын системийг боловсруулахаар зорьсон. Энэхүү систем нь стрессийн түвшин, нойрмоглолт, аюултай үйлдлийг бодит цагийн орчинд илрүүлж статистик үзүүлэлт болгон гаргадаг. Ингэснээр систем нь зөвхөн нойрмоглолт бус стресс ба анхаарал сарнилт зэрэг олон хэмжигдэхүүнийг хянаж, зам тээврийн ослоос урьдчилан сэргийлэх боломжтойг харуулж байна.

[3] АШИГЛАСАН АРГА

Энэхүү жолоочийн аюулгүй байдлыг хянах систем нь дараах 3 модулиас бүрдэнэ.

1. Нойрмоглолт болон анхаарал сарнилт илрүүлэх модуль
2. Хориотой үйлдэл илрүүлэх модуль
3. Стрессийн түвшин таних модуль

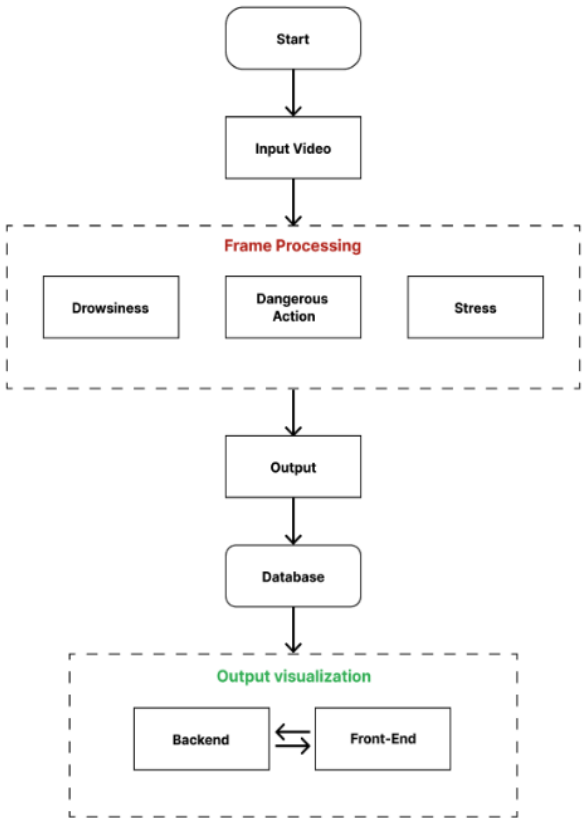


Fig. 1. Системийн үйл ажиллагааны диаграмм

1 - р зураг системийн үйл ажиллагааны диаграммыг харуулж байгаа бөгөөд бидний тодорхойлсон 3 модуль нь оролтоос эхлэн зэрэгцээ байдлаар ажиллах юм.

Монгол улсын замын хөдөлгөөний дүрэм [8] - д зааснаар, тодруулбал 3.7 - ийн Жолоочийн үүрэгт тухайн жолоочийн зүгээс тээврийн хэрэгсэл жолоодож байх үедээ гаргах дараах үйлдлүүдийг хориглосон байдаг. Үүнд:

- а/ тухайн ангиллын тээврийн хэрэгслийг жолоодох эрхгүй (жолоодлогын дадлага хийхээс бусад тохиолдолд) , эсхүл согтууруулах ундаа, мансууруулах эм, сэтгэцэд нөлөөт бодис хэрэглэсэн, мөн хөдөлгөөний аюулгүй байдлыг хангаж чадахааргүй өвчтэй буюу ядарсан үедээ тээврийн хэрэгсэл жолоодох;
- г/ эрх бүхий байгууллагаас тогтоосон жолоочийн хөдөлмөр, амралтын горимыг зөрчиж тээврийн хэрэгсэл жолоодох;
- д/ тээврийн хэрэгслийн хөдөлгөөний үед гар утас болон түүнтэй ижил төстэй төхөөрөмж, хэрэгслийг гарын оролцоотойгоор ашиглах;

3.1 Нойрмоглолт болон анхаарал сарнилт илрүүлэх модуль

Ихэнх зам тээврийн осол, эндэгдлийн гол шалтгаан нь жолоочийн анхаарал сарнилт, удаан хугацаанд амралтгүй жолоо барьснаас үүдэлтэй ядаргаа, нойрмоглолттой шууд холбоотой байдаг. Иймээс энэхүү судалгаанд боловсруулсан модуль нь жолоочийн нүүрний хувирал,

нүд, амны хөдөлгөөн, мөн толгойн чиглэлд тулгуурлан нойрмоглолт болон анхаарал сарнилтыг бодит цагийн орчинд илрүүлэх зорилготой юм.

Жолоочийн нүүрний хөдөлгөөнийг тодорхойлохын тулд **MediaPipe Face Mesh** санг ашиглаж, хүний нүүрний **468** цэгийн координатыг илрүүлэн тухайн цэгүүдийн байрлалын өөрчлөлтөд дүн шинжилгээ хийсэн.



Fig. 2. *Толгойн чиглэлийн дүрслэл*

2 - зурагт толгойн чиглэл буюу Head Pose Estimation-ийг тодорхойлохдоо хамар, эрүү, нүд болон амны ирмэгийн цэгүүдийг ашиглан 2D зураглалаас 3D нүүрний загварын вектор дүрслэл хэрэглэсэн аргыг харуулж байна. Үүнд Perspective-n-Point (PnP) алгоритм ашиглагдсан бөгөөд энэ нь камерын координатын системд үндэслэн дэлхийн координатын системд орших нүүрний байрлал, өнцөг, чиглэлийг тодорхойлох боломжийг олгодог.

1. Perspective projection model:

$$sp_c = K\,[R \mid T\,]p_w$$

Энд $p_w = [x\ y\ z\ 1]^T$ нь homogeneous world point, $p_w = [u\ v\ 1]^T$ нь homogeneous Image point, K нь камерын дотоод параметрийн матриц, R болон T нь эргэлтийн болон шилжилтийн матриц юм.

Энэхүү аргачлалаар толгойн хөдөлгөөн, чиглэл (жишээлбэл, замаас баруун, зүүн тийш удаан харсан эсэх) тодорхойлогдож, жолоочийн анхаарал сарнисан нөхцөлд систем анхааруулга өгөх боломж бүрдэнэ.



Fig. 3. Нүд, амны цэгэн дүрслэл

3 - р зурагт нойрмог болон ядралттай байдлыг илрүүлэхийн тулд нүд болон амны хүрээний цэгүүдийг дүрслэн харуулсан байна. Дээрх цэгүүд дээр тооцоолол хийхийн тулд дараах үндсэн параметрүүдийг ашиглав:

- **EAR (Eye Aspect Ratio):** Нүдний өндрийн болон өргөний харьцааг тооцоолон нүдний анилт, нээлттэй байдлыг илрүүлдэг. EAR-ийн утга тодорхой босгоос доош буурахад нүд аньсан гэж үзнэ.
- **PERCLOS** (Percentage of Eye Closure): Тодорхой хугацаанд нүд хаалттай байсан хувь хэмжээг тооцох үзүүлэлт бөгөөд жолоочийн нойрмоглолтын түвшнийг үнэлэхэд өргөн хэрэглэгддэг.

  PERCLOS нь дараах байдлаар тооцогдоно:

$$PERCLOS = \frac{N_{closed}}{N_{total}} \circ 100$$

- **Yawning Detection:** Амны хүрээний цэгүүдийн хоорондын зай, харьцаанд тулгуурлан эвшээх үйлдлийг илрүүлнэ.

Эдгээр үзүүлэлтүүдийг ашиглан систем нь нойрмоглох, анхаарал сарних зэрэг үйлдлүүдийг илрүүлэх боломжтой.

3.2 Хориотой үйлдэл таних модуль

Хориотой үйлдэл таних модулийг хэрэгжүүлэхдээ бид 2 төрлийн суурь загварыг ашигласан.

3.2.1ViT based google/vit-tiny-patch16-224

ViT (Vision transformer) нь Transformer encoder архитектурт (BERT - тэй төстэй) суурилсан загвар бөгөөд ImageNet-21k 1000 ангиллын 1 сая жишээтэй өгөгдлийн

багцаар урьдчилан сургагдсан ба дараа нь ImageNet-1k өгөгдлийн багцаар дахин сургагдсан.

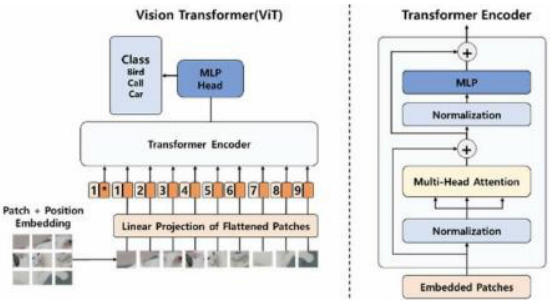


Fig. 4. Vision transformer архитектур

Зураг нь тухайн загварт оролт болж ороход 4 - р зурагт харуулсантай ижил тогтмол хэмжээтэй хэсгүүдэд (patches) хуваагдаж ордог бөгөөд тус бүрийн хэмжээ нь 16x16 харьцаатай пиксель байна. Эдгээр хэсгүүдийг шугаман байдлаар хувиргадаг. Мөн дарааллын эхэнд [CLS] token нэмэгддэг ба энэ нь ангиллын үүрэг гүйцэтгэхэд ашиглагддаг. Үүн дээр нэмээд absolute position embedding хийгддэг. Ингэснээр загвар нь хэсгүүдийн дарааллын байрлалыг ойлгож чаддаг.

Хэрвээ бидэнд шошготой зурагийн өгөгдлийн багц байгаа бол урьдчилсан сургасан энкодерийн дээр нэг шугамтай давхарга нэмж энгийн ангилагчийг сургах боломжтой юм.

***3.2.2** DeiT facebook/deit-tiny-patch16-224*

DeiT архитектуртай загвар нь илүү үр ашигтайгаар сургагдсан ViT юм. Ерөнхий суурь бүтэц нь ижил боловч загварын сургалтын арга техник нь харьцангуй өөр юм. Учир нь энэхүү загвар шууд ImageNet-1k өгөгдлийн багц дээр урьдчилан сургагдсан юм. Мөн илүү бага хэмжээтэй (patches) хувилбар, encoder давхаргын тоог бууруулсан байдал нь тухайн архитектурт суурилсан загвар цөөн тооны зураг дээр суралцах боломжийг олгодог байна.

Манай судалгааны хүрээнд тус тус загварын хамгийн жижиг хувилбаруудыг суурь загвар болгон авч ашигласан.

- DeiT facebook/deit-tiny-patch16-224 5сая параметр
- ViT based google/vit-tiny-patch16-224 5.8 сая параметр

Жолоочийн зүгээс тээврийн хэрэгсэл жолоодож байх үедээ гаргаж болох хориотой үйлдэл таних модуль нь дээрх заалтуудаас 3.7.а ядралтай байгааг илэрхийлэх, 3.7.д гар утас ашиглах утсаар ярихыг илрүүлэхийг зорьсон.

Дараах хоёр загвар нь тус бур 1000 ангилалтай 1 сая жишээтэй Image-Net21k өгөгдлийн багц дээр сурсан бөгөөд объект таних чадвартай.

Манай судалгааны хүрээнд дараах загваруудыг объект таних байдлаас нь үйлдэл таних боломжтой болгон сайжруулан сургасан бөгөөд сургасны дараагаар загвар нь дараах үйлдлүүдийг ангилах чадамжтай болох юм:

- Утсаар ярих үйлдэл
- Эвшээх үйлдэл
- Хуванцар савнаас юм уух үйлдэл
- Бусад тохиолдолд хэвийн байх үйлдэл

***3.2.3** Сургалтын өгөгдлийн багц*

Суурь загваруудын гол онцлог нь тухайн зураг болон фрэймээс зөвхөн объектыг таньдаг байсан бөгөөд энэ нь судалгааны явцад тодорхой хэмжээний сорилт болж байсан бөгөөд бидний зорилтот нөхцөл байдалд буюу жолооны ард байгаа жолооч нь тухайн хориотой үйлдлүүдийг хийж байгаа зурган өгөгдөл байсан юм.

Тухайн нөхцөлд болон зорилтот ангилалд багтсан өгөгдлийн багцыг тодорхой хэмжээгээр цахим эх сурвалжуудаас цуглуулсан боловч дараачийн асуудал болох өгөгдлийн багцын ангиллын тэнцвэртэй бус байдал тулгарсан.

Учир нь ViT болон DeiT архитектуртай загварууд нь өгөгдлөөр өлсгөлөн (data hungry) байдаг бөгөөд, дан ганц цахим эх сурвалжаас цуглуулсан өгөгдлийн багцаар загварыг сайжруулахад хангалтгүй гэж үзсэн бөгөөд өгөгдлийн багцын ангиллын тоо хэмжээ болон тэнцвэргүй байдлыг шийдэхийн тулд, бид өөрсдөө жолоо барьж буй нийт 51 мин 48 секунд хугацааны бичлэг хийсэн. Нийт 3108 секунд урттай бөгөөд, 186480 фрэйм бөгөөд, нэг ижил фрэймийг олон давхцуулж ашиглахаас сэргийлэн 1 секунд бүрээс 6 фрэймийг авсан ашигласан.

TABLE I. ТЭНЦВЭРГҮЙ ӨГӨГДЛИЙН БАГЦЫН ТОО ХЭМЖЭЭ

Тэнцвэргүй өгөгдлийн багцын тоо хэмжээ			
Ангилал	**Гаргаж авсан**	**Цуглуулсан**	**Нийт**
Drinking	1272	363	1635
Other_activities	378	2826	3204
Talking_phone	667	3264	3931
Yawning	172	1543	1715

Нийт 18648 фрэйм гаргаж авснаас нийт 2489 зураг зорилтод 4 ангилалд 1 - р хүснэгтэд харуулсан байдлаар хувиарлагдсан.

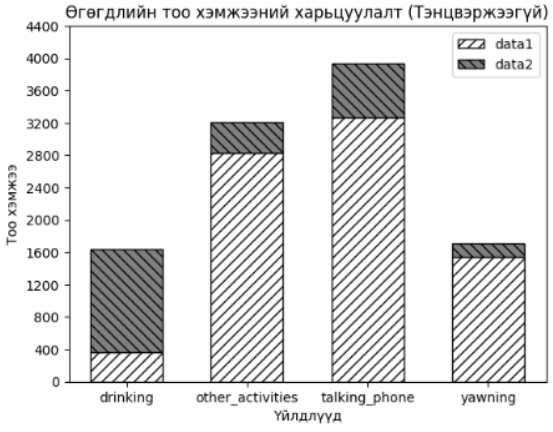


Fig. 5. Тэнцвэргүй өгөгдлийн багцын харьцуулалт

5 - р зурагт data1 нь цахим эх сурвалжаас цуглуулсан ба data2 гаргаж авсан зургийн тоо хэмжээг ангилал бүр дээр тэнцвэржээгүй байх үед харуулж байна.

TABLE II. ТЭНЦВЭРТЭЙ ӨГӨГДЛИЙН БАГЦЫН ТОО ХЭМЖЭЭ

Тэнцвэртэй өгөгдлийн багцын тоо хэмжээ			
Ангилал	**Гаргаж авсан**	**Цуглуулсан**	**Нийт**
Drinking	1272	363	1635
Other_activities	378	1257	1635
Talking_phone	667	968	1635
Yawning	172	1463	1635

2 - р хүснэгтэд 4 ангилал дахь тэнцвэржсэн өгөгдлийн багцын тоо хэмжээг харуулж байгаа бөгөөд

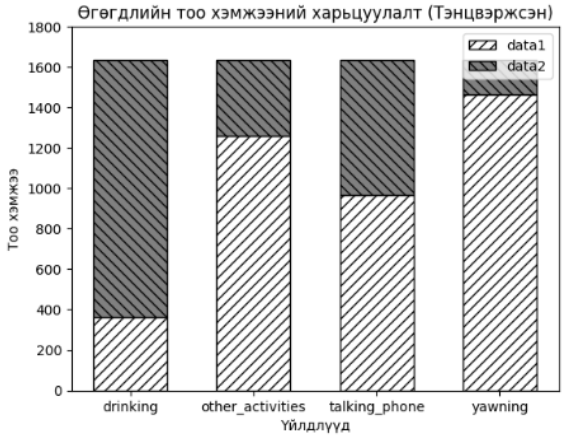


Fig. 6. Тэнцвэржсэн өгөгдлийн харьцуулалт

6 - р зурагт тэнцвэржсэн өгөгдлийн багцын ангилал тус бүр дээрх цуглуулсан болон гаргаж авсан зургийн эзлэх тоо хэмжээ болон харьцааг харуулж байна.



Fig. 7. ”drinking” ангиллын жишээ зураг

7 – р зурагт “drinking” ангилал дахь сургалтын жишээ зургийг харуулж байна.

3 - р хүснэгтэд бидний цахим эх сурвалжаас цуглуулсан өгөгдлийн багцын ангилал болон харгалзах тоо хэмжээг харуулж байна. 8-р зурагт ангилал бүрийн жишээ зургийг харуулав.

Эндээс ангилал тус бүрээс нийт 8004 зураг авч тэнцваржүүлсэн ба 8:1:1 харьцаагаар хуваан ашигласан.



Fig. 8. Сургалтын өгөгдлийн багцын ангилал тус бүрийн жишээ

3.3.2 Жижиг загвар - SVM загвар

Хүний нүүрний нийт 136 цэгийг урьдчилан сургагдсан гүн сургалтын загвар ашиглан илрүүлж, тухайн цэгүүдийн координатуудыг тоон хэлбэртэй онцлог шинжүүд болгон хувиргасан. Гарсан онцлогуудыг ашиглан SVM (Support Vector Machine) ангилагчийг сургасан.

Энэ аргын давуу тал нь шууд том нейрон сүлжээ сургалгүйгээр урьдчилан сургагдсан загварын гаргасан онцлог шинжүүдийг ашиглан илүү хөнгөн, тооцооллын хувьд үр ашигтай загвар байгуулах боломжийг олгож буй явдал юм.

3.3.3 Дунд хэмжээний загвар

Бид мөн CNN (Convolutional neural network) загварыг авч үзсэн. Ингэхдээ тухайн зураг ангиллын даалгаварт тохирсон загварын давхаргын архитектурыг авч ашигласан. Тухайн загвар нь нийт 38 давхаргатай бөгөөд ойролцоогоор 2.5 сая параметртэй юм.

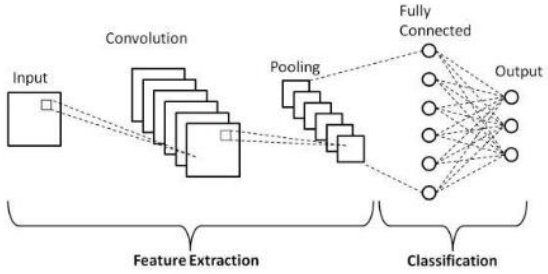


Fig. 9. CNN загварын архитектур

Энэхүү загвар нь Keras санг ашиглан хэрэгжүүлсэн гүн CNN архитектур юм. Загварын Feature Extraction хэсэг нь 5 шатлалтай Convolutional Block-уудаас бүрдэх ба эдгээр нь дүрсний онцлог шинжүүдийг шат дараалан ялгаж авдаг. Classification хэсэг нь нэг Dense давхарга бүхий бүтэцтэй бөгөөд эцэст нь softmax идэвхжүүлэлтийн

функцийг ашиглан өгөгдлийг n ангиллын аль нэгэнд хамааруулдаг.

[4] Үр дүн

4.1 Хориотой үйлдэл таних модуль

Сургалтын өгөгдлийн багц нь нийт 4 ангилал тус бурд 1635 жишээтэй бөгөөд, энэхүү өгөгдлийн багцыг сургалтын, баталгаажуулалтын, тестийн өгөгдөлд зориулж 8:1:1 харьцаатай хуваасан.

TABLE III. Загварын гиперпараметрууд

Гиперпараметрууд	
Epochs	10
Learning_Rate	0.00005
Batch_size	64
Steps	100
Warmup_ratio	0.05

4.1.1 ViT архитектуртай загвар

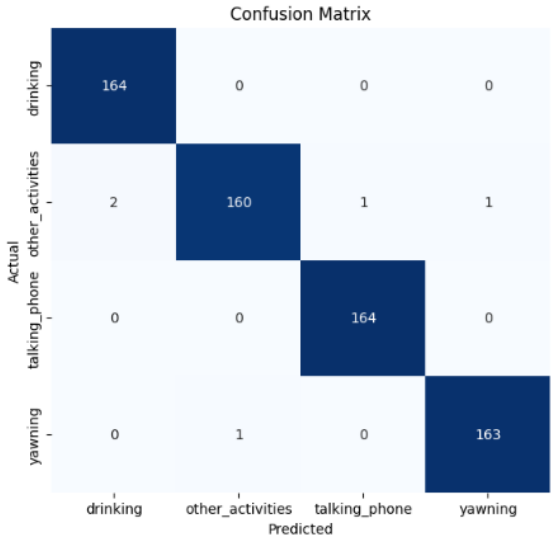


Fig. 10. ViT архитектуртай загварын тестийн өгөгдөл дээрх гүйцэтгэл

10 - р зургаас харахад загвар тестийн багц дээр хангалттай сайн үзүүлэлт үзүүлсэн байна.

TABLE IV. ViT загварын үнэлгээний хүснэгт

Accuracy	0.9924
Precision	0.9924
Recall	0.9924
F1 score	0.9924

4.1.2 DeiT архитектуртай загвар

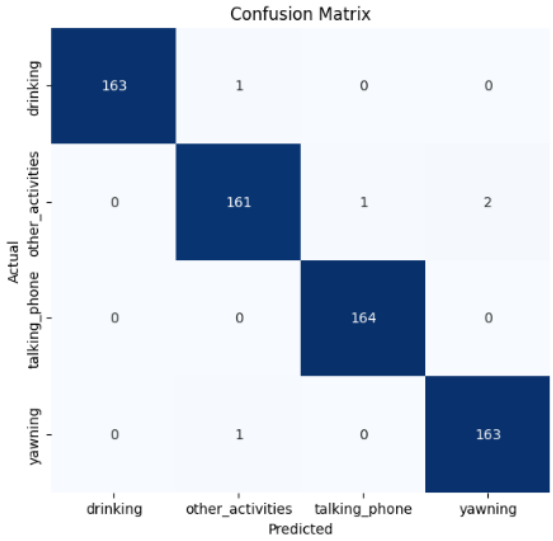


Fig. 11. DeiT архитектуртай загварын тестийн өгөгдөл дээрх гүйцэтгэл

TABLE V. DeiT загварын үнэлгээний хүснэгт

Accuracy	0.9924
Precision	0.9924
Recall	0.9924
F1 score	0.9924

5, 6 - р хүснэгтүүд болон 10, 11 – р зургуудад харуулснаар ViT болон DeiT архитектуртай загвар тус бүр ижил үр дүнг үзүүлсэн байна.

Энэ нь дараах загварууд нь нэгэн ижил transformers архитектур дээр суурилсан боловч сургалтын техник, стратегиараа ялгардаг болохыг баталж байна. Манай тохиолдолд ижил өгөгдлийн багц болон 4 – р хүснэгтэд харуулсантай ижил гиперпараметрүүдээр загваруудыг тус тус сургасан болно.

Мөн бид баталгаажуулах үүднээс бодит байдал дээр давхар камераар загваруудыг тус тус туршихад мөн ижил хэмжээний үр дүнг үзүүлсэн ба зорилтод ангиллуудыг ангилж чадаж байсан.

4.2 Стрессийн түвшин тогтоох модуль

4.2.1 SVM RBF kernel

TABLE VI. SVM RBF kernel

kernel	*C*	*gamma*
RBF	*0.01, 0.1, 1, 10, 100*	*0.001, 0.01, 0.1, 1, auto, scale*

Бид SVM загварын RBF кернелийн хамгийн оновчтой параметрийг олохын тулд 7 - р хүснэгтэд харуулснаар параметрүүдийг авч үзэн grid search байдлаар хамгийн сайн үзүүлэлт үзүүлсэн загварыг тодорхойлсон. Эндээс хамгийн сайн үр дүн үзүүлсэн парамтетр нь {C: 100, gamm: 0.01} байсан.

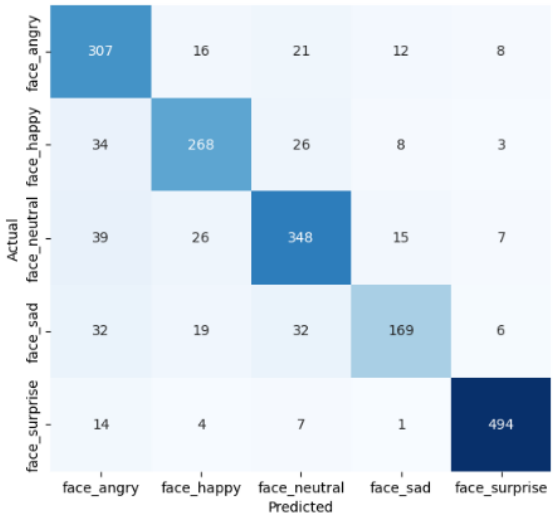


Fig. 12. SVM загварын хувилбаруудын 5 ангилал дээрх төөрөгдлийн матриц

12 – р зураг дээрх тухайн загварын төөрөгдлийн матрицаас дараах зүйл ажиглагдаж байна. Тестийн багц дээр загвар нь Angry байх сэтгэл хөдлөлийг sad, neutral болон sad ангиллууд дээр тодорхой хэмжээний төөрөгдлүүд үүссэн байна.

	precision	recall	f1-score	support
angry	0.72	0.84	0.78	364
happy	0.80	0.79	0.80	339
neutral	0.80	0.80	0.80	435
sad	0.82	0.66	0.73	258
surprise	0.95	0.95	0.95	520
accuracy			0.83	1916
macro avg	0.82	0.81	0.81	1916
weighted avg	0.83	0.83	0.83	1916

Fig. 13. SVM загварын үнэлгээний хэмжүүрийн хүснэгт

13-р зураг тестийн өгөгдлийн багц дээрх загварын гүйцэтгэлийг харуулсан ба 83% нарийвчлалыг үзүүлсэн байна.

4.2.2 Дунд хэмжээний загвар

Бидний хувьд энэхүү сэтгэл хөдлөлөөр ангилах даалгаварт SVM нь хараахан жижиг байгаа гэж үзэн CNN загвар руу шилжсэн. 3.3.3 – д дурдсан 38 давхаргатай загварыг нийт 100 epoch сургахад дараах үр дүнгүүд гарсан байна.

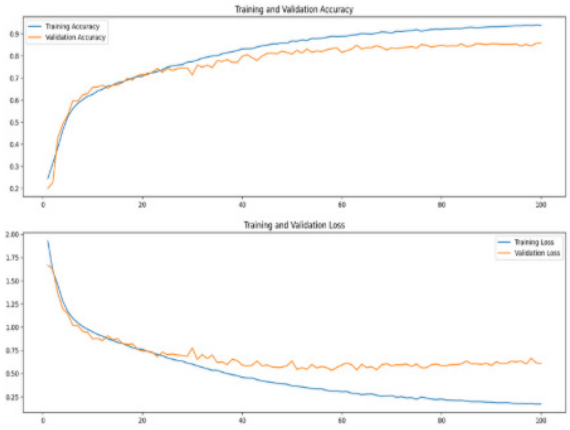


Fig. 14. CNN загварын 5 ангилал дээрх алдагдал болон нариивчлалын график

14 - р зургаас харахад загвар сургалтын болон баталгаажуулалтын багц дээр үр дүн хоёулаа тогтвортой, ойролцоо түвшинд байна.

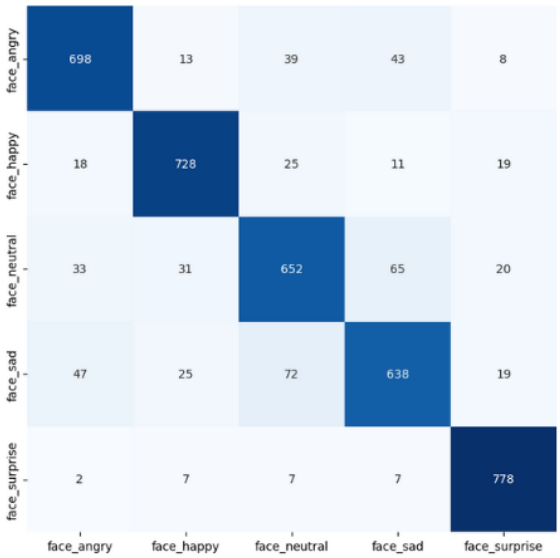


Fig. 15. CNN загварын төөрөгдлийн матриц

15 – р зурагт харуулснаар CNN загвар дээрх төөрөгдлийн матрицыг харуулж байгаа бөгөөд өмнөх SVM загварын үр дүнгээс харьцангуй илүү үр дүн үзүүлсэн байгаа боловч neutral болон sad ангилал дээр тодорхой хэмжээний төөрөгдөл байгаа ажиглагдаж байна.

	precision	recall	f1-score	support
Angry	0.87	0.87	0.87	801
Happy	0.91	0.91	0.91	801
Neutral	0.82	0.81	0.82	801
Sad	0.84	0.80	0.82	801
Surprise	0.92	0.97	0.95	801
accuracy			0.87	4005
macro avg	0.87	0.87	0.87	4005
weighted avg	0.87	0.87	0.87	4005

Fig. 16. CNN загварын тестийн багц дээрх нарийвчлал

Тухайн загварыг нийт 100 epoch сургасны дараа тестийн багц дээрх үнэлгээг 16 – р зурагт харуулж байгаа бөгөөд 87% нарийвчлалтай байгаа SVM загвараас харьцангуй илүү үр дүнг үзүүлсэн байна.

4.2.3 Том загварыг сайжруулах арга

Бид тухайн ангиллууд дээрх тодорхой хэмжээний төөрөгдлийг арилгахын тулд хориотой үйлдэл таних модуль дээрх ашигласан DeiT суурь загварыг ижил гипер параметртэй тохируулан таван сэтгэл хөдлөлийг илэрхийлж буй өгөгдлийн багц дээр нийт 10 epoch сургасан.

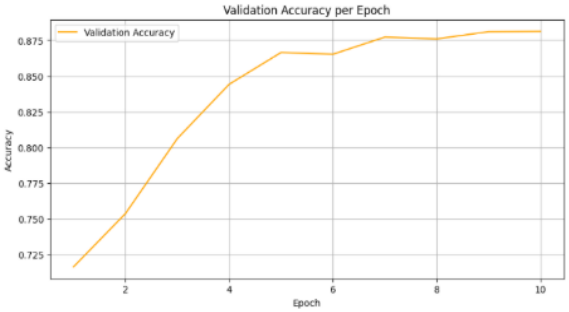


Fig. 17. DeiT загварын баталгаажуулалтын нарийвчлал

17 – р зураг сургалтын явцын баталгаажуулалтын нарийвчлалыг харуулж байна. Эндээс баталгаажуулалтын багцын нарийвчлал тогтворгүй байгаа нь тухайн өгөгдлийн зургийн хэмжээтэй хамааралтай байж болох бөгөөд сургалтын зураг нь 48x48 пиксел хэмжээтэй юм.

Сургалтын нарийвчлалыг epoch бүрд тооцоход, сургалтын орчин, GPU – ийн санах ой дүүрч сургалт зогсож байсан тул оруулаагүй болно.

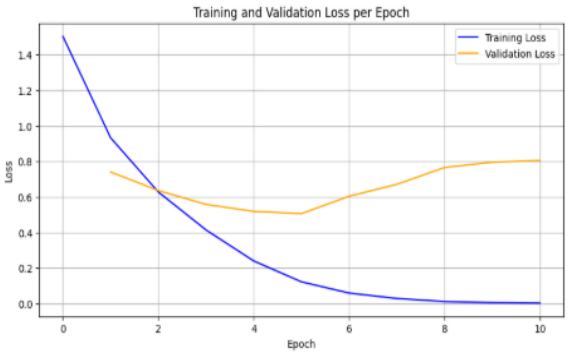


Fig. 18. DeiT загвар дээрх сургалтын болон баталгаажуулалтын алдагдал

18 – р зураг дээр сургалтын болон баталгаажуулалтын алдагдлыг харуулж байгаа бөгөөд дээр дурдсантай ижил зүй тогтол илэрч байна.

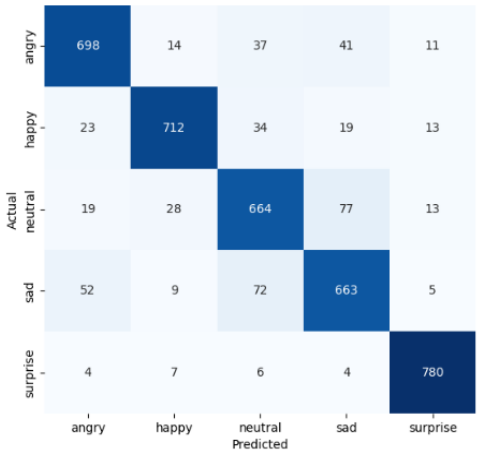


Fig. 19. DeiT загварын тестийн багц дээрх төөрөгдлийн матриц

19 – р зураг DeiT загварын төөрөгдлийн матрицыг харуулж байгаа бөгөөд энд тухайн өмнөх загваруудаас илүү ангилал хоорондын төөрөгдлийг харуулсан байна. Учир нь DeiT архитектур нь 16x16 patch тай бөгөөд энэ нь 48x48 зургаас 9 токен үүсгэх ба, онцлог шинж чанаруудыг олж чадахгүй байгааг илэрхийлж байна.

	precision	recall	f1-score	support
angry	0.88	0.87	0.87	801
happy	0.92	0.89	0.91	801
neutral	0.82	0.83	0.82	801
sad	0.82	0.83	0.83	801
surprise	0.95	0.97	0.96	801
accuracy			0.88	4005
macro avg	0.88	0.88	0.88	4005
weighted avg	0.88	0.88	0.88	4005

Fig. 20. DeiT загварын үнэлгээний хүснэгт

20 – р зураг харуулснаар загвар нь тестийн багц дээр 88% нарийвчлал үзүүлсэн байна.

4.2.4 Ангилал хоорондын төөрөгдлийн шинжилгээ

Өмнөх үр дүнгээс харахад зарим ангиллуудад төөрөгдөл үүсэж байв. Үүнийг тодорхой болгохын тулд зурган өгөгдлийн багцаас нүүрний 136 онцлог шинж чанарыг (facial landmarks) гарган авч, тухайн шинжүүд болон холбогдох ангиллын мэдээллийг агуулсан онцлог матриц үүсгэв.

Энэхүү матриц дээр хэмжээсийг багасгах зорилгоор Stochastic Neighbor Embedding (t-SNE) аргыг ашиглан онцлог шинжүүдийн уялдаа хамаарлыг шинжилсэн. t-SNE нь өндөр хэмжээсийн өгөгдлийн цэгүүдийн хоорондын ижил төстэй байдлыг магадлалын хэлбэрээр тооцож, үүнийг бага хэмжээст орон зайд хадгалж дүрслэдэг.

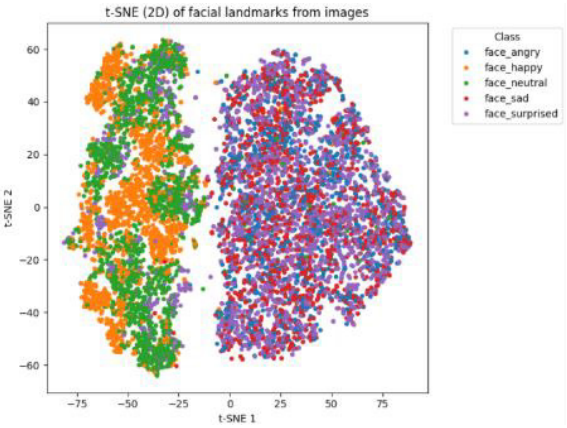


Fig. 21. Стрессийн өгөгдлийн багц дээрх T-SNE хувиргалт

21-р зураг стрессийн өгөгдлийн багц дээр t-SNE хувиргалт хийсний дараах өгөгдлийн хоёр хэмжээст дүрслэлийг үзүүлэв. Зургийн үр дүнгээс харахад face_angry, face_sad, face_surprised ангиллууд хоорондоо ойролцоо байрлаж, face_happy болон face_neutral ангиллууд мөн ижил төстэй бүлэглэл үүсгэж байна.

Үүний үндсэн дээр ижил төстэй ангиллуудыг нэгтгэн, зорилтот ангиллын тоог хоёроор хязгаарлав. Энэ нь тухайн хүний нүүрний зургаас стресстэй эсвэл стрессгүй эсэхийг тодорхойлох хоёр ангиллын асуудал болгон хувиргасан гэсэн үг юм.

Ингэснээр 21-р зураг үзүүлсэнчлэн, сөрөг сэтгэл хөдлөлүүдийг (angry, sad, surprised) "стресстэй", харин эерэг сэтгэл хөдлөлүүдийг (happy, neutral) "стрессгүй" ангилалд хамааруулж, загваруудыг энэ хоёр ангилал дээр тулгуурлан дахин сургах арга зүйг бид авч үзсэн.

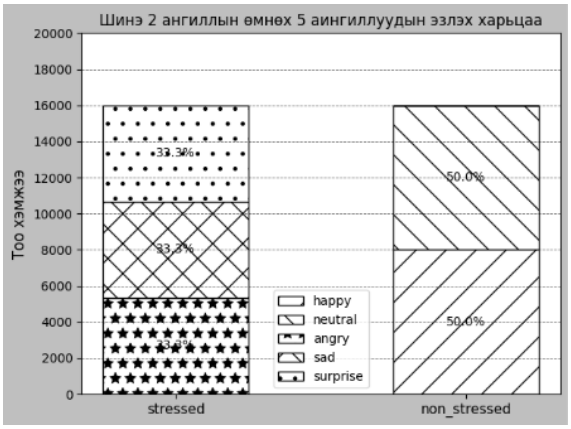


Fig. 22. SVM загварын 2 ангилал дээрх төөрөгдлийн матриц

22 – р зураг шинэ 2 ангилалд хамаарах өмнөх 5 ангиллууд болон тэдгээрийн харгалзах эзлэх хувийг харуулж байгаа бөгөөд мөн адил 8:1:1 харьцаатай сургалт, баталгаажуулалт, тестийн багцад хуваасан.

4.2.5 SVM загварын үр дүн

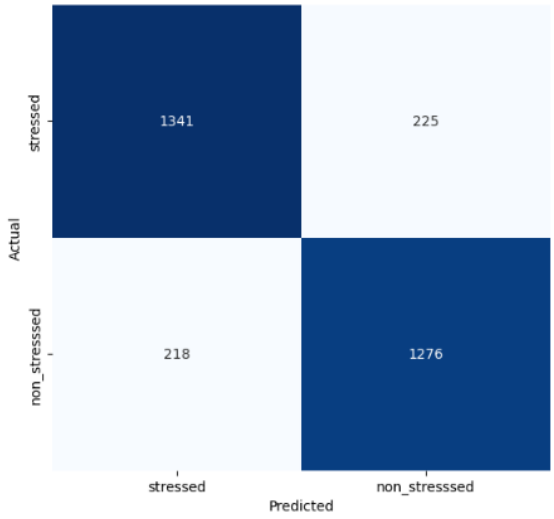


Fig. 23. SVM загварын 2 ангилал дээрх төөрөгдлийн матриц

	precision	recall	f1-score	support
stressed	0.86	0.86	0.86	1566
non_stresssed	0.85	0.85	0.85	1494
accuracy			0.86	3060
macro avg	0.86	0.86	0.86	3060
weighted avg	0.86	0.86	0.86	3060

Fig. 24. SVM RBF кернел дээрх тестийн багц дээрх үнэлгээний матриц

4.2.6 CNN загварын үр дүн

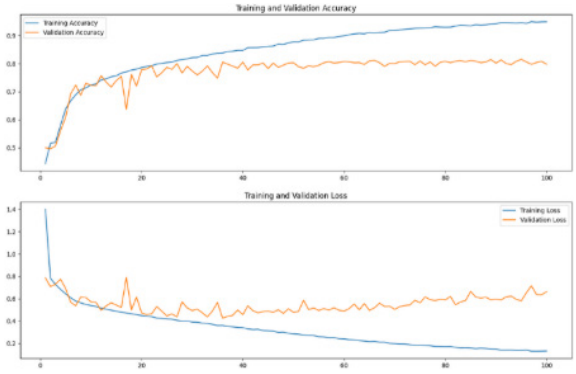


Fig. 25. Дунд хэмжээний загварын 2 ангилал дээрх алдагдал болон нарийвчлалны график

25 - р зураг харуулснаар баталгаажуулалт болон сургалтын алдагдлын зөрүү бага зэрэг нэмэгдсэн байна.

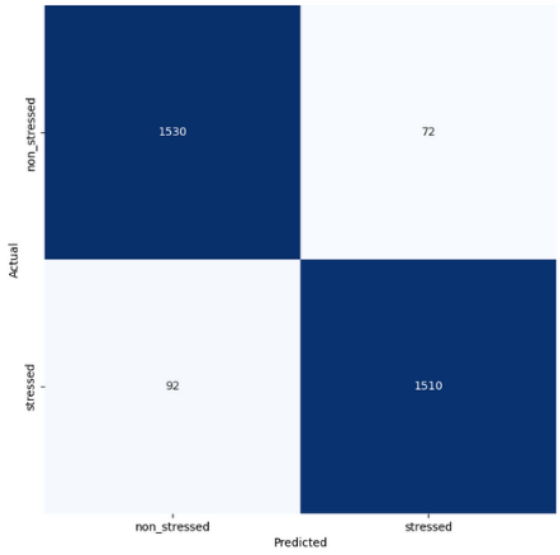


Fig. 26. CNN загварын 2 ангилал дээрх төөрөгдлийн матриц

26 - р зураг харуулснаар тухайн загварын төөрөгдлийн матриц нь бусад загвартай харьцангуй илүү үр дүнг үзүүлж байна.

	precision	recall	f1-score	support
non_stressed	0.94	0.96	0.95	1602
stressed	0.95	0.94	0.95	1602
accuracy			0.95	3204
macro avg	0.95	0.95	0.95	3204
weighted avg	0.95	0.95	0.95	3204

Fig. 27. CNN загварын үнэлгээний хүснэгт

27 – р зураг харуулснаар CNN загвар нь тухайн бууруулсан хоёр ангилал дээр хамгийн өндөр нарийвчлалыг үзүүлсэн байна.

4.2.7 DeiT загварын үр дүн

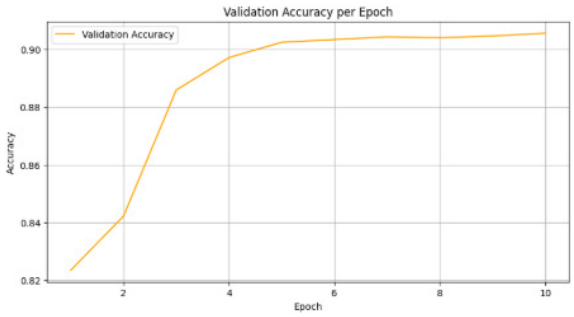


Fig. 28. DeiT загвар дээрх баталгаажуулалтын багцын нарийвчлал



Fig. 29. DeiT загвар дээрх баталгаажуулалт болон сургалтын алдагдал

28 болон 29 - р зургаас харахад таван ангилал дээрх сургалтын үетэй харьцангуй баталгаажуулалтын нарийвчлал өссөний хажуугаар алдагдал бурсан харагдаж байна.

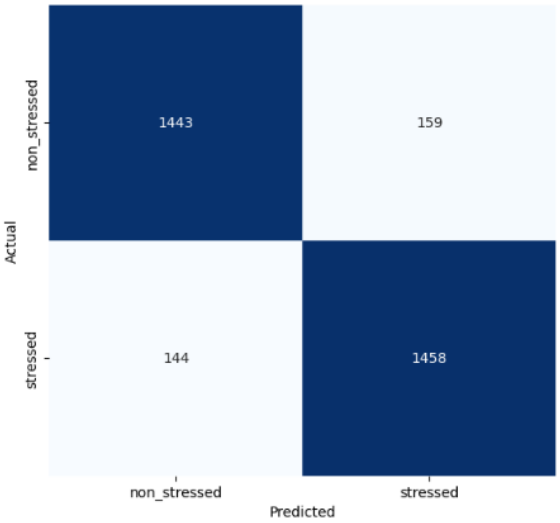


Fig. 30. DeiT загварын 2 ангилал дээрх төөрөгдлийн матриц

30 – р зураг харуулснаар загварын үр дүн сайжирсан боловч CNN загвараас харьцангуй муу үр дүнг үзүүлсэн байна.

	precision	recall	f1-score	support
non_stressed	0.91	0.90	0.90	1602
stressed	0.90	0.91	0.91	1602
accuracy			0.91	3204
macro avg	0.91	0.91	0.91	3204
weighted avg	0.91	0.91	0.91	3204

Fig. 31. DeiT загвар дээрх үнэлгээний хүснэгт

31 – р зурагт харуулснаар загварын нарийвчлал өссөн хэдий ч төөрөгдөл хэвээр байгаа нь өгөгдлийн зургийн пикселийн хэмжээтэй холбоотой хэвээр байна.

4.3 Нийт загваруудын ерөнхий харьцуулалт

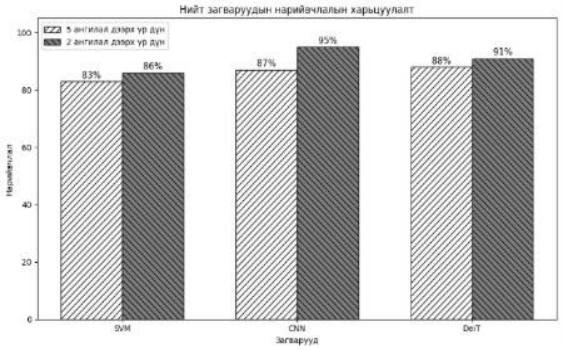


Fig. 32. Нийт загваруудын нарийвчлалын харьцуулалт

32 – р зурагт нийт загваруудын 5 болон 2 ангилал дээрх тестийн нарийвчлалыг харуулж байна. Бидний зорьсны дагуу тухайн жолоочийн стрессийг таних нь чухал бөгөөд дараах графикаас хамгийн өндөр нарийвчлал үзүүлсэн загвар CNN загварыг бид системд ашигласан.

4.4 Жолоочийн хяналтын системийн веб интерфейс

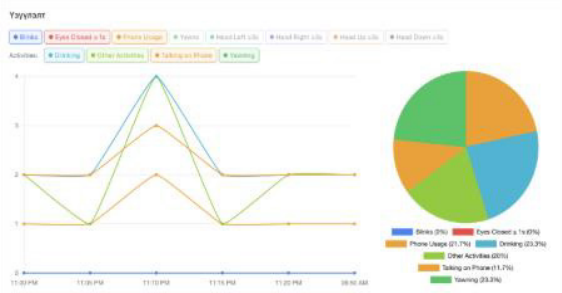


Fig. 33. Жолоочийн мэдээллийн график дүрслэл



Fig. 34. Стрессийн түвшний график дүрслэл

Зураг 33 - д жолоочийн хяналтын мэдээллийг цагийн интервалаар нарийвчлан харуулж байна. Үүнд нүд анилтын давтамж, толгойн хөдөлгөөн, эвшээх, утас оролдох, аюултай үйлдлүүдийн үзүүлэлтүүдийг цагийн интервалаар харуулж байгаа бол 34 - р зурагт жолооч тодорхой цагт стресстэй байсан үгүйл илэрхийлэх график харагдаж байна. Дээрх дэлгэрэнгүй мэдээллүүд дээр үндэслэн ослоос урьдчилан сэргийлэх шийдлүүдийг боловсруулах боломжтой.

[5] Дүгнэлт

Энэхүү судалгааны ажил нь Хиймэл оюун ухаанд суурилсан жолоочийн аюулгүй байдлын ухаалаг хяналтын системийг боловсруулж, жолоочийн сэтгэлзүйн болон зан үйлийн өөрчлөлтөөс үүдэлтэй ослын эрсдэлийг бууруулахад чиглэв. Судалгаанд компьютерын хараа, гүн сургалт, машин сургалтын алгоритмуудыг хослуулан ашигласнаар нойрмоглох, стрессдэх, анхаарал сарних болон хориотой үйлдлүүдийг бодит цагийн орчинд илрүүлэх боломжийг бүрдүүлсэн.

Нойрмоглолт ба анхаарал сарнилт илрүүлэх модулиар MediaPipe FaceMesh санг ашиглан жолоочийн нүүрний 468 цэгийн мэдээлэлд дүн шинжилгээ хийж, нүдний хувирал, амны хөдөлгөөн, толгойн чиглэл зэрэг үзүүлэлтүүдийн өөрчлөлтөд үндэслэн нойрмоглох болон анхаарал сарнисан төлөвийг тодорхойлсон. Энэ модуль нь жолооч анхаарал сарнисан үед анхааруулга өгөх боломжийг олгосон.

Хориотой үйлдэл таних модулийг Vision Transformer (ViT) болон Data-efficient Image Transformer (DeiT) архитектурт тулгуурлан хөгжүүлсэн бөгөөд энэхүү модуль нь “утсаар ярих”, “ус уух”, “эвшээх”, “бусад” гэсэн дөрвөн төрлийн үйлдлийг 99.2%-ийн нарийвчлалтайгаар ангилж чадсан.

Харин стрессийн түвшин тогтоох модулийн хувьд Support Vector Machine (SVM) болон Convolutional Neural Network (CNN) загваруудыг ашиглаж, нүүрний илэрхийлэлд суурилан эерэг болон сөрөг сэтгэл хөдлөлийг ялган тодорхойлох боломжийг бүрдүүлэв. Энэ модулийн хамгийн өндөр гүйцэтгэл 92.6%-ийн нарийвчлалтайгаар гарсан байна.

Эдгээр модулиудыг нэгтгэн боловсруулсан систем нь жолоочийн аюултай үйлдэл, анхаарал, ядралт, стрессийн түвшнийг бодит цагийн орчинд хянаж, аюул үүсэхээс өмнө анхааруулах боломжтой юм. Ийнхүү боловсруулсан хиймэл оюун ухаанд суурилсан жолоочийн хяналтын систем нь зам тээврийн ослоос урьдчилан сэргийлэх, аюулгүй байдлыг хангах, хүний хүчин зүйлтэй холбоотой эрсдэлийг бууруулахад үр ашигтай шийдэл болж байна.

Цаашид энэхүү системийн хөгжүүлэлтийг IoT-той хослуулан бодит байдал дээр, нэмэлт өгөгдлийн багцаар сайжруулан сургаж модулиудыг нэмэлтээр хөгжүүлснээр судалгааны үр дүнг практикт нэвтрүүлэх боломжтой гэж үзэж байна.

[6] Ном зүй

[1] Das, S., Pratihar, S., Pradhan, B., Jhaveri, R. H., & Benedetto, F. (2024). IoT-assisted automatic driver drowsiness detection through

facial movement analysis using deep learning and a U-Net-based architecture. Information, 15(1), 30. [CrossRef]

[2] Delwar, T. S., Singh, M., Mukhopadhyay, S., Kumar, A., Parashar, D., Lee, Y., Rahman, M. H., Sejan, M. A. S., & Ryu, J. Y. (2025). AI- and deep learning-powered driver drowsiness detection method using facial analysis. Applied Sciences, 15(3), 1102. [CrossRef]

[3] Yang, E.; Yi, O. Enhancing Road Safety: Deep Learning-Based Intelligent Driver Drowsiness Detection for Advanced Driver-Assistance Systems. Electronics 2024, 13, 708. [CrossRef]

[4] Fonseca, T., & Ferreira, S. (2025). Drowsiness detection in drivers: A systematic review of deep learning-based models. Applied Sciences, 15(16), 9018. [CrossRef]

[5] Chew, Y. X.; Abdul Razak, S. F.; Yogarayan, S.; Sayed Ismail, S. N. M. Dual-Modal Drowsiness Detection to Enhance Driver Safety. Computers, Materials & Continua 2024, 81(3), 4397–4417. [CrossRef]

[6] Liu, Y. (2024). The method of driver fatigue detection in intelligent driving. Theoretical and Natural Science, 51, 137–141. [CrossRef]

[7] Hou, J., Zhang, B., Zhong, Y., & He, W. (2025). Research progress of dangerous driving behavior recognition methods based on deep learning. World Electric Vehicle Journal, 16(2), 62. [CrossRef]

[8] Government of Mongolia. (2018). Монгол улсын замын хөдөлгөөний дүрэм [Road traffic rules of Mongolia] (Resolution No. 239). [CrossRef]

Ертөнцийн загварууд дахь сүүлний эрсдэлийн гажуудал: Уналтын таамаглалын зөрүү (DPG)-г илрүүлэх нь

Х.Сайханбаяр*, Н.Пүрэвцогт*, М.Золжаргал†
*Статистик, хэрэглээний математикийн тэнхим
†Мэдээлэл, компьютерийн ухааны тэнхим
МТЭС, МУИС, Улаанбаатар хот, Монгол улс
Email: {purevtsogt, zoljargal}@num.edu.mn

***Хураангуй*—TTD-MPC2 зэрэг бататгах сургалтын (Reinforcement Learning) агентууд нь детерминистик далд ертөнцийн загваруудыг (Deterministic Latent World Models) ашиглан төсөөлөлт замнал дээр төлөвлөлт хийх замаар өндөр тооцооллын үр ашигтайгаар оновчтой өгөөжид хүрдэг. Гэвч эдгээр загвар нь институцийн хөрөнгө оруулалтын эрсдэлийн гол хэмжүүр болох "Хамгийн их бууралт" (Maximum Drawdown - MaxDD) буюу сүүлний эрсдэлд (Tail Risk) хэрхэн нөлөөлдөг талаарх асуулт тодорхойгүй хэвээр байна. Бид бодит болон төсөөлөлт MaxDD-ийн зөрүүг хэмжих "Уналтын таамаглалын зөрүү" (Drawdown Planning Gap - DPG) хэмээх инвазив бус оношилгооны аргыг танилцуулж байна. Энэхүү арга нь шагналд төвлөрсөн уламжлалт аргуудаас ялгаатай нь сүүлний эрсдэлийг шууд оношлох боломжийг олгодог. Санхүүгийн урт хугацааны өгөгдөл дээр хийсэн бидний туршилтаар төлөвлөлтийн хугацаа уртсах тусам DPG монотон байдлаар өсөж, 5–10 алхамт төсөөллийн үед системтэйгээр эерэг утгатай болж байгааг тогтоолоо. Төлөвлөлтийн хугацааг нэмэгдүүлэх нь өгөөжид статистикийн хувьд ач холбогдолтой ахиц гаргалгүйгээр тооцооллын зардлыг огцом өсгөж, MaxDD болон калмарын харьцаа зэрэг эрсдэлийн үзүүлэлтүүдийг дордуулж байна. Мөн MPPI түүврийн тоог өөрчлөх үед калмарын харьцаа болон DPG үзүүлэлтүүд тогтвортой сайжрахгүй байгаа нь энэхүү эрсдэлийн "Сохор цэг" нь түүвэрлэлтийн санамсаргүй хэлбэлзлээс бус, харин загварын бүтцийн гажуудлаас (structural model bias) үүдэлтэйг нотолж буй явдал юм. Тэгхээр энэхүү DPG аргыг детерминистик далд ертөнцийн загварууд дахь сүүлний эрсдэлийн дутуу үнэлэлтийг илрүүлж, эрсдэлийг хянаж тохируулах практик ач холбогдол хэрэгсэл гэж үзэж байна.**

***Түлхүүр үгс*—Ертөнцийн Загвар (World Model), Загварт Суурилсан Бататгах Сургалт (Model Based Reinforcement Learning), Хамгийн их уналт (MaxDD), Санхүүгийн Эрсдэл (Financial Risk), Уналтын таамаглалын зөрүү (Drawdown Planning Gap), Сүүлний эрсдэл (Tail Risk)**

I. Удиртгал

"Хамгийн их бууралт" (Maximum Drawdown, MaxDD) [3] нь институцийн хөрөнгө оруулалтын менежментэд эрсдэлийг тодорхойлох үндсэн хэмжүүр юм. Энэ нь үнэлгээний хугацааны туршид оргилоос ёроол хүртэлх хамгийн их алдагдлыг илэрхийлж, эрсдэлийн хязгаар, хөрөнгийн хуваарилалт болон сангийн хөрөнгө татах нөхцөлийг шууд тодорхойлдог. Дисперсэд суурилсан эрсдэлийн үзүүлэлтүүдээс, тухайлбал стандарт хазайлтаас ялгаатай нь MaxDD нь замаас хамааралтай (path-dependent) захын статистик (Tail Statistic) бөгөөд ховор боловч эрс тэс бууралтын үйл явдлуудад онцгой мэдрэмтгий байдаг. Практикт нэг л хүнд бууралтын эпизод (тухайлбал, 2008 оны санхүүгийн хямрал эсвэл 2020 оны COVID-19-ийн шок) урт хугацааны стратегийн бүх цаг үеийн MaxDD-г тодорхойлж, эрсдэлийн удирдлагад ноцтой ноцтой нөлөө үзүүлэх боломжтой.

Сүүлийн жилүүдэд загварт суурилсан бататгах сургалт (Model-Based Reinforcement Learning, MBRL) [7], [8] нь санхүүгийн шийдвэр гаргалт, багцын удирдлага, арилжааны стратеги зэрэг тасралтгүй удирдлагын даалгаврууд ирээдүйтэй парадигм болон гарч ирж байна. TD-MPC2 [2] зэрэг агентууд нь детерминистик далд ертөнцийн загваруудыг (Deterministic Latent World Models) ашиглан далд орон зайд олон алхамт төсөөллийн замнал үүсгэж, MPPI [6] зэрэг төлөвлөгчөөр үйлдлийг сонгодог. Ийм хандлага нь орчинтой харилцах давтамжийг хэмнэж, урт хугацааны гүйцэтгэлийг өрсөлдөхүйц түвшинд хүргэдэг тул санхүүгийн хязгаарлагдмал, стационар бус өгөгдөл дээр бататгах сургалтыг хэрэглэхэд нэн тохиромжтой.

Гэвч ийм загваруудын сүүлний эрсдэлийн (Tail Risk) зөв калибрейшн хангалттай судлагдаагүй хэвээр байна. "Далд төлөвлөлтийн зөрүү" (Latent Planning Gap, LPG) [1] хэмээх сүүлийн үеийн ажлаар төлөвлөлтийн хугацаа уртсах тусам төсөөлсөн өгөөж нь бодит өгөөжөөс системтэйгээр их гарч, урамшууллын үүднээс хэт өөдрөг (optimistic) болж байгааг харуулсан. LPG нь ертөнцийн загвар дундаж өгөөжийг хэрхэн гажуудуулж байгааг хэмждэг ч замаас хамааралтай эрсдэлийн үзүүлэлтүүд, ялангуяа MaxDD болон Калмарын харьцаанд (Calmar ratio) [14] үзүүлэх нөлөөг эмпирик түвшинд тогтвортой шинжилж үзээгүй байна. Дундаж өгөөжийн хувьд сайн калибрейшнтэй ертөнцийн загвар хүртэл MaxDD-г үлэмж хэмжээгээр дутуу үнэлэх боломжтой бөгөөд энэ нь зөвхөн хэлбэлзэлд суурилсан үзүүлэлтүүдээр (жишээлбэл Шарпын харьцаа) хэмжихэд үл харагдана.

Манай эмпирик шинжилгээнээс харахад энэ зөрчил бодитоор илэрдэг. Төлөвлөлтийн хугацааг $H = 10$ болгосон үед French өгөгдлийн сан дээр Шарпын харьцаа +29.5%-иар өсөж (2.903 vs. 2.242) байхад Калмарын харьцаа хоёр

өгөгдлийн сангийн аль алинд нь буурч байна (ETF: 212.3 → 194.6, −8.3%; French: 163.6 → 150.6, −8.0%). Зөвхөн Шарпд тулгуурласан үнэлгээний дагуу $H = 10$-ыг илүүд үзэх нь өгөөжийг сайжруулж байгаа мэт харагдах боловч MaxDD болон Калмарыг бодитоор дордуулсан, эрсдэлийг үл тоомсорлосон сонголт болно. Зураг 2-ээс харахад H-ийн өсөлт нь Шарпын харьцааг бага зэрэг сайжруулж болох ч Калмарын харьцааг тогтвортой бууруулж байгааг илтгэнэ. Ертөнцийн загвар нэг хэмжүүрээр (Шарп) сайжирч байгаа мэт харагдах нь нөгөө хэмжүүрээр (Калмар) ноцтой муудсан байдлын шинж тэмдэг байж болох бөгөөд одоогийн оношилгооны практик ийм "Шарп–Калмарын салалт"-ыг системтэйгээр илрүүлэхгүй байна.

Энэхүү цоорхойг нөхөхийн тулд бид детерминистик ертөнцийн загварын сүүлний эрсдэлийн өөдрөг байдлыг оношлох, системд нөлөөлөхгүй (non-invasive) хэмжүүр болох "Уналтыг төлөвлөх зөрүү" (Drawdown Planning Gap, DPG)-г санал болгож байна. Концепцийн хувьд DPG(k) нь ижил төлөвлөгчийн (planner) гаргасан үйлдлийн дараалыг бодит орчин болон ертөнцийн загварт зэрэг хэрэгжүүлж, тухайн k хугацааны хүрээн дэх MaxDD-ийн зөрүүг хэмждэг. Бодлого (Policy) болон үйлдлийн дараалал ижил тул үүсэж буй аливаа ялгаа нь зөвхөн ертөнцийн загваруудын урт хугацааны таамаглалын алдаагаар тайлбарлагдана. Эерэг DPG(k) утга нь төсөөлсөн уналт бодит уналтаас бага байгааг, өөрөөр хэлбэл ертөнцийн загварууд уналтын хүндрэлийг дутуу үнэлж сүүлний эрсдэлийн (Tail Risk) хувьд хэт өөдрөг байгааг илтгэнэ.

Энэ ажилд бид DPG-г эмпирик оношилгооны хэрэгсэл байдлаар судалж, санхүүгийн хоёр урт хугацааны өгөгдлийн багц дээр TD-MPC2 төрлийн агентуудын зан төлөвийг шинжлэв. Бидний туршилтуудаас дараах гурван зүйлс тод ажиглагдав. Нэгдүгээрт, нэг алхамт эрсдэлийн хувьд DPG(1) нь үндсэндээ тэг орчим боловч төсөөллийн хугацааг $k = 10$ хүртэл сунгахад DPG(k) нь 8.8–31 дахин өсөж, бүх үр (seeds)-ийн хувьд хугацааны хүрээ уртсах тусам монотон өсөлттэй байна (Зураг 1). Хоёрдугаарт, төлөвлөлтийн хугацааг $H = 0$-ээс $H = 10$ болгох нь Шарпын харьцааг зарим тохиолдолд сайжруулж байгаа ч Калмарын харьцааг тогтвортой бууруулж, "Шарп–Калмарын салалт"-ыг бий болгож байна. Гуравдугаарт, төлөвлөлтийн хугацаа H, MPPI түүврийн тоо N-ийг ихэсгэх, эсвэл төсөөлсөн MaxDD-д энгийн λ-торгууль оноох зэрэг арга нь DPG(k)-ийн эерэг хазайлтыг найдвартай бууруулж чадахгүй байгаа нь энэхүү сүүлний эрсдэлийн "сохор цэг" нь түүвэрлэлтийн шуугиан бус, загварын бүтцийн гажуудал байж болзошгүйг харуулж байна.

Эцэст нь DPG-г детерминистик ертөнцийн загварт суурилсан санхүүгийн бататгах сургалтын агентуудын эрсдэлийн калибрейшныг оношлох практик хэрэгсэл гэж үзэж, системтэй эмпирик судалгаагаар сүүлний эрсдэлийг дутуу үнэлэх хандлагыг баримтаар харуулж байна. Цаашид DPG-г илүү өргөн орон зайд турших, мөн DPG-д мэдрэмтгий (DPG-aware) сургалтын зорилт болон төлөвлөлтийн схемийг зохион бүтээх нь чухал судалгааны чиглэл болно.

II. Холбогдох судалгаа

A. Санхүү дэх бататгах сургалт

Бататгах сургалтыг багцын удирдлага, арилжаа, эрсдэлийн хяналтад ашиглах чиглэлийн судалгаа эрчимтэй хөгжиж байна [11]. Олон арга барил нь DDPG эсвэл PPO зэрэг model-free аргуудыг ашиглан үнийн цуваа эсвэл дохионд үндэслэн арилжааны бодлогыг шууд сурч, хүлээгдэж буй өгөөж эсвэл Шарпын үзүүлэлтийг оновчлохыг зорьдог [11]. Мөн reward shaping, хэлбэлзлийн торгууль, эсвэл позицийн хязгаарлалт гэх мэт эрсдэлийн үндсэн хязгаарлалтуудыг нэвтрүүлдэг арга ч бий. Эдгээр арга нь backtest дээр өрсөлдөхүйц өгөөж үзүүлдэг ч сурсан бодлого нь MaxDD зэрэг замд суурилсан эрсдэлийн үзүүлэлтийн хүрээнд хэрхэн ажиллаж байгааг нарийвчлан хянадаггүй.

B. Загварт суурилсан бататгах сургалт ба ертөнцийн загварууд

Загварт суурилсан бататгах сургалт (MBRL) нь орчны урьдчилан таамаглах загварыг сурч, түүнийгээ төлөвлөлт эсвэл бодлогыг сайжруулахад ашиглахыг зорьдог [7], [8]. Ертөнцийн загварууд нь авсаархан latent динамикийг сурч, CEM, MPPI зэрэг түүвэрлэлтэд суурилсан төлөвлөгчдийг ашиглан үйлдлийг сонгодог бөгөөд ингэснээр тасралтгүй хяналтын даалгаварт орчинтой бага харилцан үйлчлэлээр өндөр гүйцэтгэл үзүүлдэг. Детерминистик далд ертөнцийн загварууд нь сургахад хялбар, тооцоололд хурдан байдаг ч ирээдүйн траекторийн тодорхойгүй байдлыг илэрхийлж чаддаггүй.

Өмнөх судалгаанууд ийм загварууд хүлээгдэж буй өгөөжийг урт хугацаанд хэрхэн гажуудуулж байгааг LPG-аар дамжуулан хэмжиж эхэлсэн ч MaxDD гэх мэт замаас хамааралтай эрсдэлийн үзүүлэлтүүд дээр төвлөрөөгүй байна [1]. Бидний ажил дундаж өгөөжөөс илүүтэйгээр сүүлний эрсдэлд төвлөрч энэхүү судалгааны чиглэлийн орон зайг нөхөж байна.

C. Эрсдэлд мэдрэмтгий бататгах сургалт ба бууралтад мэдрэмтгий зорилтууд

CVaR, дундаж вариансын тэнцвэр, эсвэл өсөлтөөс илүү бууралтыг хүндээр торгодог ашиг тусын функц зэрэг эрсдэлийн үзүүлэлтүүдийг оновчилдог эрсдэлд мэдрэмтгий бататгах сургалтийн судалгаанууд бий [9], [10]. Санхүүгийн салбарт зарим ажил бууралтын хязгаарлалт эсвэл замд суурилсан торгуулийг бататгах сургалтийн зорилтот функцэд нэвтрүүлж, MaxDD-г шагналын илэрхий бүрэлдэхүүн хэсэг болгон авч үздэг. Эдгээр арга нь сургалтын зорилтот функцийг өөрчлөх эсвэл хязгаарлалттай оновчлолын тусламжтайгаар эрсдэлийг хянахыг зорьдог.

DPG-ийн хамрах хүрээ ба зорилго нь үүнээс өөр харин шинэ зорилтот функц эсвэл алгоритм биш, харин одоо байгаа аливаа MBRL агент дээр ашиглаж болох оношилгооны хэрэгсэл юм. Агентийн оновчлолын бодлогыг өөрчлөхийн оронд DPG нь сурсан world model нь төлөвлөлтөд ашиглагдаж буй хугацааны хүрээнд бууралтын талаар бүтцийн хувьд өөдрөг үзэлтэй байгаа эсэхийг асуудаг.

Хүснэгт I
Санхүүгийн өгөгдлийн багцуудын тойм.

Багц	Хугацаа	Давтамж	Хөрөнгө d	Clip
ETF	2007–2024	Daily	5	$[-10\%, 10\%]$
French	1990–2024	Daily	10	$[-10\%, 10\%]$

D. Бууралтын үзүүлэлтүүд ба эрсдэлийн удирдлага

MaxDD болон холбогдох замд суурилсан эрсдэлийн үзүүлэлтүүд нь хөрөнгийн удирдлага, хедж сан болон эрсдэлийн төсөвлөлтөд өргөн хэрэглэгддэг [3], [4]. Гауссын бус өгөөжтэй эсвэл тэгш бус үр дүнгийн профиль бүхий стратегиудыг үнэлэхдээ мэргэжлийн хүмүүс Шарпын үзүүлэлтээс илүү Калмар гэх мэт бууралтад суурилсан харьцааг илүүд үздэг [14]. Бидний оруулсан хувь нэмрийг бууралтад төвлөрсөн энэхүү өнцгийг ертөнцийн загваруудын үнэлгээнд нэвтрүүлсэн гэж үзэж болно. DPG нь загварт суурилсан агентуудад бууралтын буруу тохируулгыг оношлох энгийн, хэрэгжүүлэх боломжтой хэрэгслээр хангаж, ертөнцийн загваруудын үнэлгээг практикт ашиглагддаг эрсдэлийн үзүүлэлтүүдтэй илүү нийцүүлж байна.

III. Туршилтын арга зүй

A. Өгөгдлийн багц

Бид DPG хэмжүүрийн зан төлөвийг зах зээлийн олон төрлийн дэглэм болон хямралын нөхцөлийг агуулсан, удаан хугацааны санхүүгийн хоёр өгөгдлийн багц дээр үнэлэв.

АНУ-ын салбарын ETF (ETF). Бид технологи (XLK), санхүү (XLF), эрүүл мэнд (XLV), эрчим хүч (XLE), аж үйлдвэр (XLI) гэсэн таван төрлийн хөрвөх чадвар өндөртэй АНУ-ын салбарын ETF-үүдийн 2007-01-04–2024-12-30 хоорондох өдөр тутмын тохируулсан хаалтын ханшийг Yahoo Finance-ээс авч ашиглав [12]. Өдөр тутмын өгөөжийг үнийн хувийн өөрчлөлтөөр тооцож, хэт өндөр хазайлтууд динамикт давамгайлахаас сэргийлэхийн тулд $[-10\%, 10\%]$ интервалд хязгаарлав. Өгөгдлийг цаг хугацааны дарааллаар нь сургалтын (80%) болон тестийн (20%) багцад хуваав.

Францын 10 салбарын багц (French). Kenneth Frenchийн өгөгдлийн сангаас авсан 10 төрлийн Fama–French салбарын багцын 1990-01-02–2024-12-31 хоорондох өдөр тутмын өгөөжийг ашиглав [13]. Энэхүү цуваа нь dot-com-ын хямрал, 2008 оны санхүүгийн хямрал, 2020 оны COVID-19-ийн цочрол, 2022 оноос хойших хүүгийн өсөлтийн орчин зэрэг бүтцийн өөрчлөлтүүдийг хамарна. ETF-ийн нэгэн адил өгөөжийг $[-10\%, 10\%]$ интервалд хязгаарлаж, 80/20 харьцаатайгаар цаг хугацааны дарааллаар сургалт–тестийн багцуудад хуваав.

Хоёр өгөгдлийн багцын хувьд тестийн үеийг зөвхөн үнэлгээний зорилгоор ашигласан. Сургалтын бүх параметр болон гиперпараметрүүдийг өгөгдлийн багц, санамсаргүй эхлэлүүдийн хооронд тогтмол байлгав.

B. Багцын орчин (Portfolio environment)

Бид d тооны хөрөнгөтэй тасралтгүй удирдлага бүхий багцын орчныг ашиглав. Уг орчинд агент өдөр бүр $w_t \in \mathbb{R}^d$ жингийн векторыг сонгоно. Орчин нь эдгээр жинг softmax функцээр хэвийнчилж, сөрөг бус, нийлбэр нь нэг байх нөхцөлийг хангана. Хөрөнгийн өдрийн өгөөж $r_t \in \mathbb{R}^d$ өгөгдсөн тохиолдолд багцын үнэ

$$V_{t+1} = V_t \left(1 + w_t^\top r_t\right)$$

гэсэн зарчмаар өөрчлөгдөнө.

t цаг дахь ажиглалт нь өнгөрсөн өгөөжийн гулсах цонх (rolling window) болон тухайн үеийн багцын жингийн нэгдэл юм. Бүх туршилтад 10 өдрийн хоцрогдолтой цонх ашигласан тул ажиглалтын хэмжээс нь $d \times 10 + d$ байна. Эпизод бүр тестийн бүтэн үеийг хамарч, тестийн цонх дуусахад төгсөнө.

Сургалтыг тогтворжуулахын тулд шагналыг сургалтын багц дахь хөрөнгө тус бүрийн өгөөжийн дундаж стандарт хазайлтаар хэвийнчилж, агент руу илгээх нэг алхмын шагнал нь $r_t/\sigma_{\text{train}}$ байна. MaxDD болон Calmar-ын харьцааг үргэлж хэвийнчлээгүй багцын үнийн траекторооc тооцдог [3], [14].

C. Агент ба ертөнцийн загвар

Манай агент нь детерминистик далд ертөнцийн загвар (latent world model) болон кросс-энтропи суурьтай MPPI төлөвлөлтийг ашигладаг TD-MPC2 төрлийн хүчитгэх сургалтын агент юм [2], [6]. Архитектур нь дараах хэсгүүдээс бүрдэнэ.

- Ажиглалтыг далд төлөв z_t руу хөрвүүлэх энкодер.
- (z_t, a_t) хосоос z_{t+1} төлөвийг урьдчилан таамаглах детерминист динамик функц.
- (z_t, a_t) хосоос нэг алхмын шагналыг урьдчилан таамаглах шагналын толгой.
- z_t төлөвөөс эцсийн үнэ цэнийг урьдчилан таамаглах үнэ цэнийн толгой.
- z_t төлөвийг дундаж үйлдлийн вектор руу хөрвүүлэх бодлогын толгой.

Агентийг өгөгдлийн багц бүрийн сургалтын хэсэг дээр стандарт TD-MPC төрлийн зорилтууд болон дахин тоглуулах (replay) буфферийг ашиглан сургасан. Сургалтын дараа ертөнцийн загвар болон бодлогыг царцааж, DPG-ийн бүх хэмжилт болон аблаци туршилтуудыг нэмэлт оновчлолгүйгээр дараагийн шатанд гүйцэтгэсэн.

Үнэлгээний явцад бид төлөвлөлтийн хоёр горимыг авч үзэв.

- $H = 0$ **(“төлөвлөлтгүй”).** Агент олон алхамт урьдчилсан тооцоолол (rollout) хийлгүйгээр бодлогын толгойг шууд ашиглана. Энэ нь эрсдэлийн хувьд тохируулсан суурь үзүүлэлт (baseline) болно.
- $H > 0$ **(олон алхамт төлөвлөлт).** Агент MPPI ашиглан далд орон зайд $H \in \{1, 2, 5, 10, 20\}$ урттай үйлдлийн дарааллыг түүвэрлэж, сайжруулан, хамгийн эхний үйлдлийг орчинд гүйцэтгэнэ.

D. Үнэлгээний үзүүлэлтүүд (Evaluation metrics)

Бид тохиргоо бүрийг олон тооны санамсаргүй эхлэлүүд дээр үнэлж, дараах үзүүлэлтүүдийг тайлагнав.

- **Эпизодын өгөөж.** Тестийн хугацаанд авсан хэвийнчилсэн өдөр тутмын шагналуудын нийлбэр.
- **Максимум бууралт (MaxDD).** Эпизодын туршид багцын үнэ оргил цэгээсээ уналтын хамгийн доод цэг хүртэлх хувийн алдагдал [3], [4].
- **Шарпын харьцаа.** Эпизодын дундаж өгөөжийг эпизодын стандарт хазайлтад харьцуулсан харьцаа.
- **Калмарын харьцаа.** Эпизодын дундаж өгөөжийг дундаж MaxDD-д харьцуулсан харьцаа, энэ нь бууралтад суурилсан эрсдэлийг тооцсон эрсдэлд суурилсан гүйцэтгэл юм [14].
- **Инференс хоцрогдол.** Үнэлгээний явцад CPU дээр хэмжсэн, нэг алхамд зарцуулсан дундаж бодит хугацаа.

Гол харьцуулалтуудын хувьд бид Wilcoxon signed-rank тестийг ашиглан эпизодын өгөөж болон MaxDD-ийн ялгааны статистикийн ач холбогдлыг тодорхойлсон [15]. Эцэст нь, DPG хэмжүүрийн хувьд $k \in \{1, 5, 10\}$ хугацааны дагуу $\mathrm{DPG}(k)$ үзүүлэлтийг тайлагнаж, төлөвлөлтийн хугацаа нэмэгдэх тусам tail-risk буруу тооцоологдох байдал хэрхэн хувирч байгааг шинжилсэн. Мөн төлөвлөлтийн хугацаа H, MPPI түүврийн тоо N, уналтын торгуулийн коэффициент λ зэргийг бие даан өөрчилсөн аблаци туршилтууд хийж, эдгээр сонголтууд нь DPG болон эрсдэлд суурилсан гүйцэтгэлд хэрхэн нөлөөлж байгааг судалсан. Эдгээр аблаци туршилтын зорилго нь DPG(k)-ийн эерэг хазайлтыг архитектур болон гиперпараметрийн энгийн өөрчлөлтөөр бууруулж, бүр арилгаж болох эсэхийг шалгах явдал юм. Туршилтын үр дүнгээс харахад H, N, λ-гийн ямар ч тохиргооны үед DPG(k) тогтвортой эерэг хэвээр үлдэж байгаа нь энэхүү зөрчил нь зөвхөн түүвэрлэлтийн шуугиан бус, харин детерминистик ертөнцийн загварын бүтцийн "сохор цэг" болохыг илтгэнэ.

IV. Арга зүй: Уналтын таамаглалын зөрүү (DPG)

Бид тасралтгүй удирдлагатай багцын орчинд ажиллаж буй RL агентийг авч үзье. Цаг хугацаа t-д агент d хөрөнгөнд жин оноосон $a_t = w_t \in \mathbb{R}^d$ үйлдлийг сонгож, өгөөж r_t авч, багцын үнэ V_t дараах байдлаар өөрчлөгдөнө:

$$V_{t+1} = V_t \left(1 + w_t^\top r_t\right).$$

A. Хамгийн их бууралт (Maximum Drawdown)

Өгөгдсөн $V_{0:T}$ замналын хувьд MaxDD-ийг стандарт байдлаар:

$$\mathrm{MaxDD}(V_{0:T}) = \max_{0 \le s \le t \le T} \frac{V_s - V_t}{V_s}$$

гэж тодорхойлно [3]. Энэ хэмжээ нь багц хэзээ ч өмнөх оргилоосоо доош унаагүй бол тэгтэй тэнцүү байх ба оргилоос ёроол хүртэлх хамгийн гүнзгий бууралтын хэмжээг илэрхийлнэ.

B. Ижил үйлдлийн дор төсөөлсөн ба бодит уналт

Ертөнцийн загварын алдааг тусгаарлахын тулд бид ижил үйлдлийн дарааллын дор төсөөлсөн болон бодит замнал дээрх MaxDD-ийг харьцуулна. Үнэлгээний үед:

1) Орчинг t-д дахин эхлүүлж, ажиглалт o_t-г авна.
2) Агентээс k алхмын үйлдлийн дараалал $(a_t, \ldots, a_{t+k-1})$-г гаргуулна.
3) Эдгээр үйлдлийг ертөнцийн загвар дотор гүйцэтгэж, төсөөлсөн траектори $\hat{V}_{t:t+k}$-г авч, $\widehat{\mathrm{MaxDD}}(k) = \mathrm{MaxDD}(\hat{V}_{t:t+k})$-г тооцно.
4) Ижил үйлдлийг бодит орчинд гүйцэтгэж, $V_{t:t+k}$-г авч, $\mathrm{MaxDD}_{\text{real}}(k) = \mathrm{MaxDD}(V_{t:t+k})$-г тооцно.

Үйлдлийн дараалал ижил тул үүссэн зөрүү нь зөвхөн загварын таамаглалын алдаанаас шалтгаална.

C. Уналтын таамаглалын зөрүү тодорхойлолт (DPG)

Уналтын таамаглалын зөрүүг дараах байдлаар тодорхойлно:

$$\mathrm{DPG}(k) = \mathbb{E}\left[\mathrm{MaxDD}_{\text{real}}(k) - \widehat{\mathrm{MaxDD}}(k)\right],$$

энд хүлээлт нь эпизодууд, эх байдал, болон агентийн дотоод санамсаргүй байдал (MPPI түүвэрлэлтийн санамсаргүй байдал) дээрх дундаж юм. $\mathrm{DPG}(k) > 0$ байвал ертөнцийн загвар уналтыг дутуу үнэлж (реал орчин дахь уналт илүү гүн), харин $\mathrm{DPG}(k) \approx 0$ байвал сайн тохируулгатай гэж үзнэ.

Манай туршилтад $\mathrm{DPG}(k)$-ийг олон эпизод болон эхлэлүүдийн дундаж дээр тооцож, ихэвчлэн $k \in \{1, 5, 10\}$ утгуудыг авч үзсэн. DPG нь сургалтын болон төлөвлөлтийн схемийг өөрчлөхгүй, зөвхөн сургалтын дараах үнэлгээний үед тооцогддог тул инвазив бус оношлогооны хэрэгсэл юм.

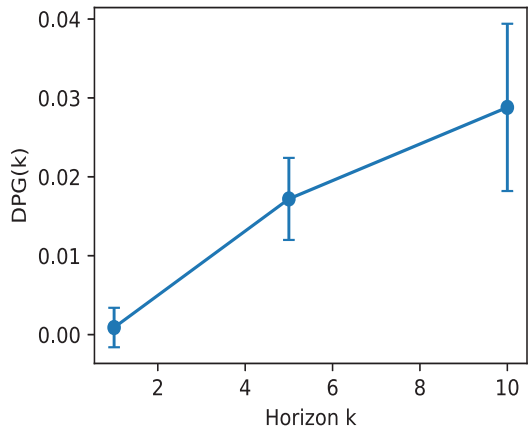
V. Уналтын таамаглалын зөрүү эмпирик зан төлөв

A. Нэг алхамт эрсдэлийн калибрейшн

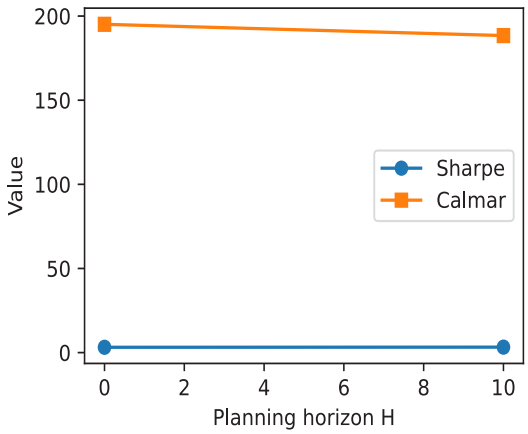
Юуны өмнө бид ертөнцийн загвар нэг алхам урагш харсан төлөвт эрсдэлийн хувьд зөв тохируулгатай эсэхийг шинжилсэн. ETF болон French өгөгдлийн багцын хувьд $\mathrm{DPG}(1)$ нь дунджаар 10^{-3} орчим утгатай, системчилсэн хазайлтгүй, үндсэндээ тэгтэй тэнцүү байна. Өөрөөр хэлбэл, загвар нь нэг алхам урагш харж дүгнэлт хийх үед тооцоолсон уналт нь бодит уналттай нягт нийцэж байна. Энэхүү ажиглалт нь өгөөжийн LPG-тэй холбоотой өмнөх судалгаатай нийцэж байна [1]. Нэг алхамт таамаглал үнэн зөв байх нь ертөнцийн загварын орон нутгийн динамик нь богино хугацааны зан төлөвийг дахин бүтээхэд хангалттай болохыг илтгэнэ. Бид энд MaxDD-ийн хувьд ч адил болохыг тогтоов. Хамгийн ойрын хугацааны төлөвт загвар нь сүүлний эрсдэлийн талаар тодорхой өөдрөг эсвэл гутранги хандлагагүй байна.

B. Төсөөллийн хугацааг уртасгах үед уналтын таамаглалын зөрүү өсөлт

Төсөөллийн траекторийг уртасгах үед $\mathrm{DPG}(k)$-ийн зан төлөв чанарын хувьд өөрчлөгдөнө. ETF өгөгдөл дээр $k \in \{1, 5, 10\}$ үед:



Зураг 1. ETF өгөгдөл дээрх DPG(k) ($k \in \{1, 5, 10\}$) ($H = 10$ үед хэмжсэн). Алдааны баганууд нь эхлэлүүдийн дагуух стандарт хазайлтыг илэрхийлнэ.



Зураг 2. ETF өгөгдөл дээрх төлөвлөлтийн хугацаа H-ийн нөлөө: Шарп болон Калмарын харьцаа vs $H \in \{0, 1, 2, 5, 10, 20\}$.

- DPG$(1) = 0.0009 \pm 0.0025$, Pr[DPG$(1) > 0$] $= 0.70$;
- DPG$(5) = 0.0172 \pm 0.0052$, Pr[DPG$(5) > 0$] $= 1.00$;
- DPG$(10) = 0.0288 \pm 0.0106$, Pr[DPG$(10) > 0$] $= 1.00$.

Зураг 1-ээс харахад DPG(1) үндсэндээ тэгтэй тэнцүү боловч DPG(5) болон DPG(10) нь бүх эхлэлүүдийн хувьд эерэг, хэмжээгээрээ өсч байна. “Monotone |DPG| seeds: 10/10” гэсэн статистик нь бүх эхлэл дээр $|\text{DPG}(1)| < |\text{DPG}(5)| < |\text{DPG}(10)|$ нөхцөл биелж байгааг илтгэнэ. Францын өгөгдөл дээр ч ижил төстэй өсөлтийн хэв шинж ажиглагдсан.

Эдгээр үр дүн нь DPG нь зөвхөн тэг орчимд хэлбэлзэх чимээ биш, харин төсөөллийн хугацаа уртсах тусам хүчтэй илэрдэг сүүлний эрсдэлийн талаарх системчилсэн өөдрөг хандлага байгааг харуулж байна.

C. Төлөвлөлтийн хугацаа ба эрсдэлээр тохируулсан гүйцэтгэл

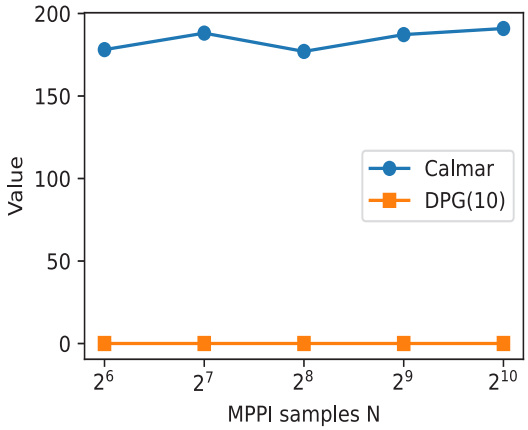
DPG хэмжилт хамгийн их мэдээлэл өгөх нь агентийн бодит төлөвлөлтийн хугацаа H болон эрсдэлийн үзүүлэлттүүдтэй хамт авч үзэх үед юм. ETF өгөгдөл дээр:

- $H = 0$: өгөөж = 20.82±4.13, MaxDD = 0.1077±0.0168, Шарп = 3.163, Калмарын = 195.1, хоцрогдол = 0.05 ms;
- $H = 10$: өгөөж = 24.56 ± 5.44, MaxDD = 0.1307 ± 0.0257, Шарп = 3.243, Калмарын = 188.4, хоцрогдол = 10.71 ms;
- Wilcoxon p(return) = 0.2754, p(mdd) = 0.0645.

Эдгээр тоон үр дүнг Зураг 2-д дүрслэн харуулсан бөгөөд H-ийн өсөлт нь Шарпын харьцааг бага зэрэг сайжруулж болох ч Калмарын харьцааг тогтвортой бууруулж байгааг харж болно.

D. MPPI түүврийн тооны (N) шинжилгээ

Алдааны бас нэг эх үүсвэр нь төлөвлөгчийн дисперс байж болох бөгөөд илүү олон замнал (trajectory) түүвэрлэснээр **DPG**-г бууруулах боломжтой юм. Үүнийг шалгахын тулд бид төлөвлөлтийн хугацааг $H =$ 10 дээр тогтмол байлгаж, **MPPI** түүврийн тоог $N \in \{64, 128, 256, 512, 1024\}$ хүртэл өөрчлөн ETF өгөгдөл дээр туршилт гүйцэтгэлээ.



Зураг 3. ETF өгөгдөл дээрх $H = 10$ үед MPPI түүврийн тоо N-ийн нөлөө: Calmar болон DPG(10) нь $N = 64$-өөс $N = 1024$ хүртэл монотон биш, харьцангуй нарийн зурваст хэлбэлзэж байна.

Зураг 3-ээс харахад N болон Калмар, эсвэл DPG(10)-ийн хооронд тодорхой монотон хамаарал ажиглагдахгүй байна. Калмар нь 178-аас 191-ийн хооронд хэлбэлзэж байгаа бол DPG(10) нь ойролцоогоор 2–3%-ийн хүрээнд хадгалагдаж байна. Нэмэлт хүснэгт S4-т тоон утгуудыг дэлгэрэнгүй үзүүлэв.

Нэмэлт хүснэгт S4-өөс харахад N болон **Калмарын харьцаа**, эсвэл DPG(10)-ийн хооронд тодорхой монотон хамаарал ажиглагдахгүй байна. **Калмарын харьцаа** нь 178-аас 191-ийн хооронд хэлбэлзэж байгаа бол DPG(10) нь ойролцоогоор 2-3%-д хадгалагдсан хэвээр байна. **MPPI** түүврийн тоог 64-өөс 1024 хүртэл 16 дахин нэмэгдүүлсэн ч **DPG**-ийн хэмжээ бодитоор буурсангүй.

Энэхүү зан төлөв нь **DPG** нь түүвэрлэлтийн шуугианаас илүүтэйгээр загварын бүтцийн гажуудлаас (structural model bias) үүдэлтэй болохыг нотолж байна. Ертөнцийн

загваруудын детерминистик динамик нь урт хугацааны бууралтын хуримтлалыг бүтцийн хувьд буруу загварчилдаг тул тухайн загвараас олон тооны замнал (trajectory) түүвэрлэх нь суурь алдааг засварлаж чадахгүй.

VI. Хязгаарлалт ба Ирээдүйн Судалгаа

Бидний судалгаа хэд хэдэн хязгаарлалттай бөгөөд эдгээр нь цаашдын судалгааны чиглэлийг тодорхойлох суурь болж байна.

Нэгдүгээрт, бид DPG(k)-ийн зан төлөвийг зөвхөн эмпирик түвшинд шинжилсэн бөгөөд хугацааны хүрээ (k) уртсах тусам DPG өсөх үзэгдлийг онолын түвшинд батлаагүй байгаа. Бидний туршилт ETF болон French өгөгдлийн багцын хүрээнд хугацааны хүрээ k-г нэмэгдүүлэхэд DPG(k) монотон өсөлттэй байгааг тогтвортой харуулсан хэдий ч энэхүү зүй тогтол нь илүү өргөн хүрээний нөхцөл (өөр орчин, архитектур, сургалтын схем)-д хэрхэн илрэхийг онолын хувьд баталгаажуулаагүй юм. Иймд DPG(k)-ийн урт хугацааны зан төлөвийг илүү ерөнхий нөхцөлд тодорхойлох нь судалгааны чухал чиглэл хэвээр байна.

Хоёрдугаарт, зөвхөн санхүүгийн хоёр өдөр тутмын өгөгдлийн сан (5 ETF, 10 салбарын багц) болон детерминистик MLP-д суурилсан ертөнцийн загвар, TD-MPC2 төрлийн төлөвлөгч дээр төвлөрсөн. Эдгээр нь зах зээлийн олон горим, стационар бус нөхцөлийг агуулсан ч харьцангуй бага давтамжтай, хязгаарлагдмал орчны жишээнүүд юм. Иймд бидний дүгнэлтийг ерөнхий болгохын тулд DPG-ийн зан төлөвийг өндөр давтамжийн арилжаа, опцион хеджинг, олон хөрөнгийн багцын удирдлага, мөн DMControl зэрэг стандарт тасралтгүй удирдлагын даалгавруудад өргөтгөн, стохастик эсвэл ensemble төвтэй ертөнцийн загварууд Dreamer/MBPO болон distributional/CVaR гэх мэтэд суурилсан агентуудад системтэй хамруулж судлах шаардлагатай.

Гуравдугаарт, DPG-г зөвхөн оношилгооны хэмжүүр хэлбэрээр ашиглав. Өөрөөр хэлбэл сүүлний эрсдэлийн дутуу үнэлэлтийг идэвхтэй бууруулах шинэ сургалтын зорилт эсвэл архитектур санал болгоогүй. Төлөвлөлтийн хугацаа H эсвэл түүврийн тоо N-ийг өөрчлөх, төсөөлсөн MaxDD-д энгийн λ-торгууль оноох нь DPG(k)-ийн эерэг хазайлтыг арилгаагүй нь илүү нарийн, DPG-т мэдрэмтгий (DPG-aware) зорилт, бат бөх төлөвлөлтийн схем, эсвэл бууралтад төвлөрсөн ертөнцийн загварийн архитектур шаардлагатайг илтгэж байна. DPG-г хязгаарласан эсвэл шууд оновчилсон зорилтын функцүүдийг (жишээлбэл, DPG-г дээд хязгаарт барих хязгаарлалттай оновчлол) зохиож, тэдгээрийн эмпирик болон онолын шинж чанарыг судлах нь цаашдын гол чиглэлүүдийн нэг юм.

Эцэст нь DPG-г оффлайн дүн шинжилгээний үед тооцсон бөгөөд бодит цагийн (online) хяналтын схемд ашиглаагүй. Практик нэвтрүүлэлтийн хувьд DPG-г бодит цагийн мониторингийн үзүүлэлт болгон ашиглаж, DPG(1) болон DPG(k)/DPG(1)-ийн динамикийг ашиглан тархалтын шилжилт, зах зээлийн горимын өөрчлөлтийг эрт илрүүлэх боломжтой эсэхийг судлах нь өндөр эрсдэлтэй санхүүгийн системүүдийн хувьд онцгой ач холбогдолтой байж болох юм гэж үзэж байна.

VII. Дүгнэлт

Энэхүү өгүүлэлд бид детерминистик ертөнцийн загварт суурилсан бататгах сургалтын агентуудад сүүлний эрсдэлийн калибрейшныг оношлох шинэ хэмжүүр болох “Уналтыг төлөвлөх зөрүү” (Drawdown Planning Gap, DPG)-г танилцуулав. DPG(k) нь ижил төлөвлөгчөөр сонгогдсон үйлдлийн дарааллыг бодит орчин болон ертөнцийн загварт зэрэг тоглуулж, тухайн хугацааны хүрээнд MaxDD хэр хэмжээнд дутуу үнэлэгдэж байгааг хэмждэг инвазив бус оношилгооны хэрэгсэл юм.

Санхүүгийн хоёр урт хугацааны өгөгдлийн сан дээрх TD-MPC2 төрлийн агентуудын эмпирик шинжилгээ нь DPG(1) нь нэг алхмын хүрээнд үндсэндээ тэг орчим, харин төсөөллийн хугацаа уртсах тусам DPG(k) нь бүх эхлэлүүдийн хувьд монотон өсөж, 5–10 алхмын түвшинд системтэйгээр эерэг болдгийг харууллаа. Төлөвлөлтийн хугацааг $H = 0$-ээс $H = 10$ хүртэл сунгах нь Шарпын харьцааг зарим тохиолдолд сайжруулж байгаа ч Калмарын харьцааг тогтвортой муутгаж, дээрх DPG-ээр хэмжигдэх сүүлний эрсдэлийн өөдрөг байдлыг дагуулж байна. Мөн төлөвлөлтийн хугацаа H, MPPI түүврийн тоо N-ийг ихэсгэх, эсвэл төсөөлсөн MaxDD-д энгийн λ-торгууль оноох зэрэг арга нь DPG(k)-ийн эерэг хазайлтыг найдвартай бууруулж чадахгүй байгааг бид харуулсан нь энэ “сохор цэг” нь түүвэрлэлтийн шуугиан бус, загварын бүтцийн гажуудал болохыг илтгэж байна.

Эдгээр ажиглалт нь детерминистик ертөнцийн загваруудад суурилсан санхүүгийн бататгах сургалтын агентуудыг зөвхөн дундаж өгөөж эсвэл Шарпын хувьд бус, MaxDD болон Калмар зэрэг бууралтад суурилсан үзүүлэлтээр мөн оношлох шаардлагатайг онцолж байна. DPG нь ертөнцийн загваруудын сүүлийн эрсдэлийн калибрейшныг практикт хэрэглэгддэг эрсдэлийн үзүүлэлтүүдтэй нийцүүлэн үнэлэх, цаашлаад DPG-д мэдрэмтгий (DPG-aware) сургалтын зорилт болон төлөвлөлтийн схемийг зохион бүтээхэд суурь болж чадна гэдэгт итгэж байна.

Ном зүй

[1] S. Khoshoisaikhan and P. Nugjgar, “Latent Planning Gap: A Diagnostic Framework for Deterministic World Models,” in *Proc. ICIM*, 2026, under review.

[2] N. Hansen, H. Su, and X. Wang, “TD-MPC2: Scalable, Robust World Models for Continuous Control,” in *Proc. ICLR*, 2024.

[3] M. Magdon-Ismail and A. F. Atiya, “Maximum Drawdown,” *Risk Magazine*, vol. 17, no. 10, pp. 99–102, 2004.

[4] L. R. Goldberg and O. Mahmoud, “Drawdown: From Practice to Theory and Back Again,” *Mathematics and Financial Economics*, vol. 11, no. 3, pp. 275–297, 2017.

[5] R. Cont, “Empirical Properties of Asset Returns,” *Quantitative Finance*, vol. 1, no. 2, pp. 223–236, 2001.

[6] G. Williams et al., “Information-Theoretic Model Predictive Control,” *IEEE Trans. Robot.*, vol. 34, no. 6, pp. 1603–1622, 2018.

[7] D. Hafner et al., “Dream to Control,” in *Proc. ICLR*, 2020.

[8] M. Janner et al., “When to Trust Your Model,” in *Proc. NeurIPS*, 2019.

[9] M. G. Bellemare, W. Dabney, and R. Munos, “A Distributional Perspective on Reinforcement Learning,” in *Proc. ICML*, 2017.

[10] A. Tamar, Y. Glassner, and S. Mannor, “Optimizing the CVaR via Sampling,” in *Proc. AAAI*, 2015.
[11] J. Moody and M. Saffell, “Learning to Trade via Direct Reinforcement,” *IEEE Trans. Neural Netw.*, vol. 12, no. 4, pp. 875–889, 2001.
[12] R. V. Arnak, “yfinance: Yahoo! Finance Market Data Downloader,” 2019. [Online]. Available: https://github.com/ranaroussi/yfinance
[13] E. F. Fama and K. R. French, “Common Risk Factors in the Returns on Stocks and Bonds,” *J. Financial Economics*, vol. 33, no. 1, pp. 3–56, 1993.
[14] T. W. Young, “Calmar Ratio: A Smoother Tool,” *Futures*, vol. 20, no. 1, p. 40, 1991.
[15] F. Wilcoxon, “Individual Comparisons by Ranking Methods,” *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945.

Хавсралт

Нэмэлт материал

Энэхүү нэмэлт хэсэгт үндсэн текстэд дурдсан тоон үр дүнгийн дэлгэрэнгүй хүснэгтүүдийг үзүүлэв.

Хүснэгт II
Нэмэлт хүснэгт S2: ETF өгөгдөл дээрх DPG(k) ($H = 10$ үед хэмжсэн).

k	DPG(k)-ийн дундаж	Стандарт хазайлт	Pr[DPG(k) > 0]
1	0.0009	0.0025	0.70
5	0.0172	0.0052	1.00
10	0.0288	0.0106	1.00

Хүснэгт III
Нэмэлт хүснэгт S3: ETF өгөгдөл дээрх төлөвлөлтийн хугацаа H-ийн нөлөө.

H	Өгөөж	MaxDD	Шарп	Калмар	Хоцрогдол (ms)
0	20.82	0.108	3.16	195.1	0.05
10	24.56	0.131	3.24	188.4	10.71

Хүснэгт IV
Нэмэлт хүснэгт S4: MPPI түүврийн тоо N-ийн нөлөө (ETF, $H = 10$).

N	Өгөөж	MaxDD	Шарп	Калмар	DPG(1)	DPG(5)	DPG(10)
64	20.72	0.117	2.89	178.1	0.0033	0.0129	0.0216
128	25.22	0.134	4.79	188.1	0.0025	0.0145	0.0255
256	24.51	0.139	4.50	177.0	0.0020	0.0168	0.0282
512	22.83	0.123	3.31	187.2	0.0007	0.0105	0.0217
1024	23.55	0.124	3.89	190.9	0.0018	0.0196	0.0307

EEG дохионд суурилсан тархины нейропластик шинж чанарын судалгаа ба машин сургалтаар ялган таних нь

Цэрэн-Ойдовын Баянбилэг
Мэдээллийн технологийн тэнхим
ШУТИС, МХТС
Улаанбаатар, Монгол
bayanbileg.ts@must.edu.mn

Цэрэндондогийн Тэнгис
Электроникийн тэнхим,
ШУТИС, МХТС
Улаанбаатар, Монгол
tengis@must.edu.mn

Батжаргалын Долгорсүрэн
Мэдээллийн технологийн тэнхим
ШУТИС, МХТС
Улаанбаатар, Монгол
b.dolgorsuren@must.edu.mn

Хураангуй - Тархины нейропластик шинж чанар нь суралцах болон туршлагын нөлөөгөөр тархины бүтэц, үйл ажиллагаа өөрчлөгдөх чадварыг илэрхийлдэг бөгөөд үүнийг тархины цахилгаан идэвхжлийн өөрчлөлтөөр ажиглах боломжтой. Энэхүү судалгааны ажлын зорилго нь “Миний нутаг” шүлгийг унших явцад үүсэх тархины цахилгаан идэвхжлийн өөрчлөлтийг Electroencephalogram (EEG)-ийн аргаар шинжилж, суралцах үйл ажиллагааны онцлогийг машин сургалтын аргаар ялган танихад оршино. Судалгаанд нийт 10 оролцогчийг хамруулж, шүлгийг хараагаар унших, сэтгэл хөдлөлтэй унших, мөн шүлгийн агуулгын талаарх богино хэмжээний видео үзсэний дараа (мэдлэгтэйгээр) унших гэсэн гурван нөхцөлд EEG дохиог бүртгэн авсан. Бичигдсэн EEG дохионд урьдчилсан боловсруулалт хийж, Power Spectral Density (PSD) болон wavelet коэффициентийг тооцоолсон бөгөөд Welch болон Bartlett-ийн аргуудаар давтамжийн шинжилгээ гүйцэтгэсэн. Үүсгэсэн өгөгдлийн санг ашиглан Support Vector Machine (SVM), Random Forest (RF), Convolutional Neural Network (CNN) загваруудыг сургаж, суралцагч мэдээллийг механик болон ухамсарт боловсруулж буй байдлыг ангилан тодорхойлов. Туршилтын үр дүнд Random Forest загвар хамгийн өндөр буюу 91% нарийвчлалтайгаар ангилал хийсэн. Мөн судалгаагаар шүлэг, яруу найраг унших үед тархины Theta болон Alpha давтамжийн мужид мэдэгдэхүйц өөрчлөлт илэрч байгааг тогтоосон. Эдгээр үр дүн нь EEG дохионд суурилан яруу найргийн танин мэдэхүйн онцлогийг ялган таних, улмаар суралцах процессыг үнэлэх боломжтойг харуулж байна.

Түлхүүр үг: EEG, нейропластик, суралцах процесс, шүлэг унших, PSD, CWT, машин сургалт

I. Удиртгал хэсэг

Суралцах үйл явц нь хүний тархины бүтэц, үйл ажиллагаанд өөрчлөлт үүсгэдэг бөгөөд энэ үзэгдлийг тархины нейропластик шинж чанартай холбон тайлбарладаг. Нейропластик чанар нь шинэ мэдлэг эзэмших, мэдээлэл боловсруулах, ой тогтоолт үүсэх, туршлага хуримтлагдах явцад тархины мэдрэлийн холбоосууд өөрчлөгдөх, дахин зохион байгуулагдах чадварыг илэрхийлдэг. Иймээс суралцах үйл явцыг тархины үйл ажиллагааны өөрчлөлтөөр дамжуулан судлах нь танин мэдэхүйн шинжлэх ухаан, боловсрол судлал, хиймэл оюуны судалгаанд чухал ач холбогдолтой юм. Сүүлийн жилүүдэд тархины үйл ажиллагааг хэмжих EEG ашиглан анхаарал, ой тогтоолт, сэтгэл хөдлөл, суралцах үйл явцыг судлах судалгаанууд эрчимтэй хийгдэж байна. EEG нь тархины цахилгаан идэвхжлийг өндөр хугацааны нарийвчлалтай хэмжих боломжтой бөгөөд тархины долгионы Alpha, Beta, Theta, Delta давтамжийн мужууд нь хүний танин мэдэхүйн болон сэтгэл хөдлөлийн төлөвтэй нягт холбоотой байдаг. Ялангуяа Theta долгион нь ой тогтоолт болон мэдээлэл боловсруулах үйл ажиллагаатай, Alpha долгион нь анхаарал төвлөрөл болон тайван төлөвтэй холбоотой гэж үздэг.

Нөгөө талаас, машин сургалтын аргуудыг EEG дохионд ашигласнаар хүний танин мэдэхүйн төлөв, анхаарал, сэтгэл хөдлөл, суралцах хэв маягийг автоматаар ялган таних боломж нэмэгдэж байна. SVM, RF, CNN зэрэг алгоритмууд нь EEG дохионд суурилсан ангилал, танилт, төлөв тодорхойлох судалгаанд өргөн хэрэглэгдэж байна. Иймээс төрлийн судалгаа нь боловсролын технологи, ухаалаг сургалтын систем, танин мэдэхүйн судалгаа, тархи-компьютер интерфейс зэрэг олон салбарт хэрэглэгдэх боломжтой.

II. Судлагдсан байдал

Хүний тархины үйл ажиллагааг EEG-ийн аргаар судлах судалгаа сүүлийн жилүүдэд эрчимтэй хөгжиж, танин мэдэхүй, суралцах үйл явц, сэтгэл хөдлөл, хэл ярианы боловсруулалтыг илүү нарийвчлан шинжлэх боломжийг олгож байна. EEG нь тархины цахилгаан идэвхжлийг өндөр хугацааны нарийвчлалтайгаар хэмжих боломжтой тул унших, мэдээлэл боловсруулах, анхаарал төвлөрөл, ой санамж зэрэг когнитив үйл ажиллагааг судлахад өргөн хэрэглэгдэж байна. Унших үйл

ажиллагаа нь тархины олон хэсгийг хамарсан нарийн төвөгтэй процесс бөгөөд уншиж буй текстийн төрөл, агуулга, сэтгэл хөдлөлийн оролцоо зэргээс хамааран тархины идэвхжил өөр өөр байдлаар илэрдэг.

Яруу найраг болон уран зохиолын текст унших үед хүний тархинд сэтгэл хөдлөл, төсөөлөл, хэлний боловсруулалт, утга ойлголт зэрэг олон танин мэдэхүйн процесс зэрэгцэн явагддаг болохыг EEG-д суурилсан судалгаанууд харуулсан. Тухайлбал, шүлэг унших үед тархины гоо зүйн үнэлгээ болон сэтгэл хөдлөлийн боловсруулалттай холбоотой идэвхжил нэмэгддэг [1], харин уран зохиолын метафор ойлгох үед хэлний гүн боловсруулалттай холбоотой нейрофизиологийн өөрчлөлтүүд илэрдэг [2]. Мөн унших үйл явцад анхаарал, ой санамж, утга ойлголттой холбоотойгоор тархины Alpha болон Theta долгионы мужид тодорхой өөрчлөлтүүд үүсдэг болохыг судалгаанууд тогтоосон [3].

Монголын нэрт зохиолч Д. Нацагдоржийн “Миний нутаг” шүлэг [4] нь эх оронч үзэл, сэтгэл хөдлөл, дүрслэл, утга агуулгын хувьд баялаг бүтээл бөгөөд уншигчийн сэтгэл хөдлөл болон танин мэдэхүйн үйл ажиллагаанд нөлөөлөх боломжтой уран зохиолын бүтээл юм. Ийм төрлийн яруу найргийн бүтээл унших үед тархины олон талт когнитив болон сэтгэл хөдлөлийн боловсруулалт явагддаг тул EEG дохиогоор судлах нь шинжлэх ухааны ач холбогдолтой юм. Сүүлийн жилүүдэд EEG дохионд суурилан машин сургалтын аргуудыг ашиглаж хүний танин мэдэхүйн төлөв, анхаарал төвлөрөл, суралцах хэв маягийг тодорхойлох судалгаанууд эрчимтэй хөгжиж, SVM, Random Forest, CNN зэрэг алгоритмуудыг өргөн хэрэглэж байна. Гэсэн хэдий ч яруу найраг унших үеийн тархины үйл ажиллагааг EEG болон машин сургалтын хосолсон аргаар судалсан ажил хязгаарлагдмал хэвээр байна. Иймээс энэхүү судалгаанд яруу найраг унших үеийн тархины идэвхжлийг нейропластик шинж чанартай холбон судалж, EEG дохионд суурилан суралцах процессыг машин сургалтын аргаар ангилан тодорхойлохыг зорьсон бөгөөд энэ нь EEG-д суурилсан ухаалаг сургалтын систем боловсруулахад онол, практикийн ач холбогдолтой юм.

III. Онолын үндэс

A. Тархины нейропластик шинж чанар

Тархины нейропластик чанар (Neuroplasticity) нь хүний тархи суралцах, шинэ мэдээлэл эзэмших, туршлага хуримтлуулах, орчны нөлөөнд дасан зохицох явцад мэдрэлийн эсүүдийн хоорондын холбоос өөрчлөгдөх, шинэ холбоос үүсэх, зарим холбоос сулрах зэрэг бүтэц, үйл ажиллагааны өөрчлөлт хийх чадварыг хэлнэ. Суралцах үйл явцын үед нейрон хоорондын синапсын холбоос бэхжих (synaptic plasticity), шинэ мэдрэлийн холбоос үүсэх, мэдрэлийн сүлжээ дахин зохион байгуулагдах зэрэг өөрчлөлтүүд явагддаг. Эдгээр өөрчлөлтийг тархины үйл ажиллагааны өөрчлөлтөөр, тухайлбал EEG дохион дахь давтамжийн мужуудын өөрчлөлтөөр ажиглах боломжтой. Иймээс суралцах үйл явцыг нейропластик өөрчлөлттэй холбон EEG ашиглан судлах нь танин мэдэхүйн нейро шинжлэх ухаанд чухал ач холбогдолтой.

B. Электроэнцефалограф (EEG)

EEG нь тархины мэдрэлийн эсүүдийн цахилгаан идэвхжлийг хуйханд байрлуулсан электродоор хэмжих арга бөгөөд өндөр хугацааны нарийвчлалтай тул тархины үйл ажиллагааны өөрчлөлтийг бодит хугацаанд ажиглах боломж олгодог. Энэхүү дохионд дараах үндсэн давтамжийн мужуудыг ялган авч үздэг. Судалгаагаар Theta долгион нь ой тогтоолт болон суралцах үйл явцтай, Alpha долгион нь анхаарал төвлөрөл болон танин мэдэхүйн боловсруулалттай холбоотой болохыг тогтоосон. Иймээс EEG дохион дахь эдгээр долгионы өөрчлөлтийг ашиглан суралцах үйл явцыг судлах боломжтой.

Хүснэгт I
EEG долгионы давтамжийн муж

Долгион	Давтамж	Тайлбар
Delta	0.5–4 Hz	Гүн нойр
Theta	4–8 Hz	Ой тогтоолт, суралцах
Alpha	8–13 Hz	Тайван байдал, анхаарал
Beta	13–30 Hz	Идэвхтэй сэтгэхүй
Gamma	>30 Hz	Мэдээлэл боловсруулах

C. Дохионы шинжилгээний аргууд

EEG дохионы давтамжийн шинж чанарыг тодорхойлохдоо PSD-ийг үнэлэх аргууд өргөн хэрэглэгддэг. Үүнд спектрийн нягтралыг найдвартай тооцоолох зорилгоор Welch [5] болон Bartlett [6]-ийн аргуудыг ашигладаг. Welch-ийн арга нь Bartlett-ийн аргын сайжруулсан хувилбар бөгөөд дохиог давхцах (overlapping) сегментүүдэд хуваж, сегмент бүрт цонхны функц (жишээлбэл Hamming window) хэрэглэсний дараа спектрийг тооцоолж, дундажладаг. Давхцал болон цонхны функц ашигласнаар спектрийн алдагдлыг бууруулж, шуугианы нөлөөг багасган, PSD-ийн үнэлгээг илүү тогтвортой, нарийвчлалтай болгодог. Bartlett-ийн арга нь эх дохиог хэд хэдэн тэнцүү урттай, давхцахгүй (non-overlapping) сегментүүдэд хуваж, сегмент бүрийн спектрийг тооцоолсны дараа дундажлах зарчимд суурилдаг. Энэхүү арга нь спектрийн хэлбэлзлийг (variance) бууруулах давуу талтай боловч сегментүүдийн хооронд давхцал байхгүй тул мэдээллийн алдагдал үүсэх, спектрийн нягтрал буурах сул талтай.

Иймээс EEG дохионы хувьд Welch-ийн арга нь шуугиан ихтэй биосигналд илүү тохиромжтой гэж үзэгддэг бол Bartlett-ийн арга нь энгийн, тооцооллын хувьд хялбар шийдэл болдог. Эдгээр аргуудыг хослуулан ашигласнаар тархины долгионы давтамжийн онцлогийг илүү найдвартай тодорхойлох боломж бүрддэг.

D. Механик болон ухамсарт суралцах

Суралцах үйл явцыг мэдээлэл боловсруулах онолоор механик болон ухамсарт суралцах гэж ангилдаг. Механик суралцах нь мэдээллийг утгыг гүнзгий ойлголгүйгээр давтах, цээжлэх байдлаар суралцахыг хэлнэ. Энэ үед мэдээлэл богино хугацааны ой санамжид хадгалагдах ба хурдан мартагдах хандлагатай. Ухамсарт суралцах нь мэдээллийн утга, холбоо, шалтгаан, агуулгыг ойлгож, өмнөх мэдлэгтэй холбон боловсруулж суралцах үйл явц юм. Энэ үед мэдээлэл урт хугацааны ой санамжид хадгалагдаж, илүү үр дүнтэй суралцах нөхцөл бүрддэг. Танин мэдэхүйн онолоор ухамсарт суралцах үед тархины мэдээлэл боловсруулах үйл ажиллагаа илүү идэвхтэй явагддаг бөгөөд EEG дохионд Alpha болон Theta долгионы өөрчлөлт илүү тод илэрдэг гэж үздэг.

E. Шүлэг, яруу найраг унших үеийн танин мэдэхүйн процесс

Яруу найраг унших нь энгийн текст уншихтай харьцуулахад тархины илүү олон бүсийг хамарсан танин мэдэхүйн нийлмэл процесс юм. Шүлэг унших үед хэлний боловсруулалт, утга ойлголт, төсөөлөл, сэтгэл хөдлөл, ой санамж, гоо зүйн үнэлгээ зэрэг олон танин мэдэхүйн үйл ажиллагаа зэрэг явагддаг. Иймээс яруу найраг унших үед тархины хэл яриа, ой санамж, сэтгэл хөдлөл, төсөөлөлтэй холбоотой тархины бүсүүд идэвхждэг бөгөөд энэ нь EEG дохионд тодорхой давтамжийн өөрчлөлт хэлбэрээр илэрдэг.

IV. Туршилтын хэсэг

A. Судалгаанд оролцогчид

Судалгаанд 1 курсийн 10 оюутныг туршилтанд сонгон оролцуулсан. Оролцогчид ном унших дадалтай бөгөөд туршилтын нөхцөлд зориулсан үндсэн мэдээллийг Хүснэгт II-т харуулсан.

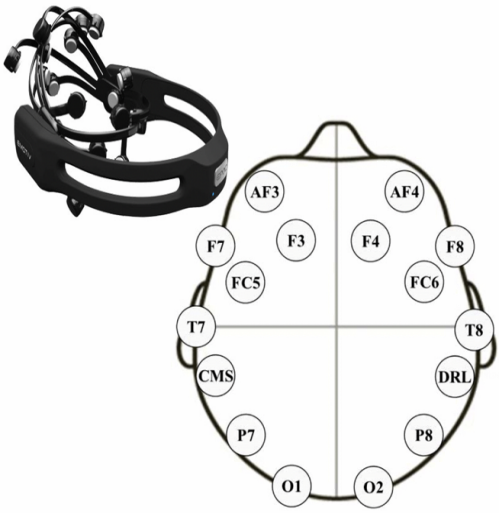
Хүснэгт II
Туршилтанд оролцогчдын мэдээлэл

№	Нэр	Нас	Хүйс	Аймаг	Сум	Анхаарал
1	NJ	18	эм	Дорнод	Чойбалсан	Сайн
2	LB	18	эм	УБ	-	Сайн
3	AZ	18	эм	Хөвсгөл	Мөрөн	Сайн
4	MZ	18	эр	УБ	-	Сайн
5	AB	18	эр	Өмнөговь	Даланзадгад	Сайн
6	GZ	18	эм	Орхон	Баян-Өндөр	Сайн
7	ZN	18	эм	УБ	-	Сул
8	BAB	40	эм	Архангай	Эрдэнэмандал	Дунд
9	ME	18	эм	Сэлэнгэ	Сайхан сум	Сайн
10	AJ	18	эм	Орхон	Баян-Өндөр	Дунд

B. EEG дохио цуглуулах

1) Төхөөрөмж ба тохиргоо

EEG бичлэгийг Emotiv Epoc X [7] төхөөрөмжөөр 16 сувгийн 14 цэг дээр хийсэн. Дохиог 128 Hz давтамжтайгаар цуглуулсан бөгөөд туршилтын өмнө дараах дарааллыг дагаж бэлтгэсэн. Үүнд, EEG туршилт хийхийн тулд эхлээд оролцогчид малгайг зөв байрлуулж өмсгөх шаардлагатай бөгөөд ингэснээр мэдрэгч бүр тархины зөв хэсгийг дамжуулж, дохионы чанар сайжирна. Дараа нь холболтын чанарыг шалгаж, 14 сувгийн мэдрэгчийг шаардлагатай бол шингэнээр норгох замаар контакт сайжруулна. Үүний дараа системийг calibration хийж, оролцогчийн тархины анхны хэвийн дохионы түвшинг тогтооно. Эцэст нь туршилтын үеийн EEG дохиог бичлэгт авч, дараах анализ, боловсруулалтад ашиглахад бэлэн болгоно.



Зураг 1. Emotiv Epoc X төхөөрөмж

2) Туршилтын нөхцөл

Шүлгийг цагаан дэлгэц дээр хар өнгөөр харуулах үед оролцогчийн унших үйл явц шууд эхэлдэг. Эхний уншилтын үеэр тархи үгийн утга, өгүүлэмж, дүрслэл, төсөөллийг боловсруулж, анхаарал төвлөрөлт, дүрслэн төсөөлөх когнитив процессууд идэвхждэг. Унших эхний хэсэг тодорхой хугацаанд үргэлжилсний дараа хар дэлгэц харуулснаар завсарлага, амралт өгч, оролцогчийг дахин уншихад бэлддэг бөгөөд энэ үед тархи өмнөх уншилтаас үүссэн санах ой, дүрслэл, сэтгэл хөдлөлтэй холбоотой мэдээллийг дахин боловсруулан, танин мэдэхүйн болон ухамсарт суралцах процессыг нэмэгдүүлдэг. Ийнхүү анхны болон дахин унших үйл явц нь тархины когнитив болон эмоциональ идэвхжилд ялгаатай нөлөө үзүүлж, EEG дохиогоор эдгээр өөрчлөлтийг илрүүлэх боломжтой болдог. Мөн туршилтын дизайн нь шүлгийг унших үеийн тархины EEG дохионы өөрчлөлтийг гурван нөхцөлд судлахад чиглэсэн.

Эхний нөхцөл (T1) нь шүлгийг чимээгүй буюу зөвхөн хараагаар унших бөгөөд энэ үед тархи үгийн бүтцийн боловсруулалт, дүрслэл, анхаарал төвлөрөлтийн механик буюу implicit суралцах процессыг голчлон идэвхжүүлдэг. Хоёр дахь нөхцөл (T2) нь уншигчид шүлгийн агуулгыг сэтгэл хөдлөлтэй хослуулан унших бөгөөд энэ нь ухамсарт буюу explicit суралцах процессыг идэвх-



Зураг 2. Туршилтын үйл явц, жишээ

жүүлж, лимбик систем, фронталь болон темпораль лобын төвүүдийн идэвхжил нэмэгддэг. Гурав дахь нөхцөл (T3) нь эхлээд богино хэмжээний видео үзүүлсний дараа шүлгийг дахин унших байдлаар явдаг бөгөөд энэ үед тархи нь урьдчилсан өдөөлтөөс үүссэн механик болон ухамсарт суралцах процессыг зэрэг идэвхжүүлдэг. Ингэснээр тархины когнитив ба эмоциональ процессын ялгааг тодорхойлох боломжийг олгодог бөгөөд EEG дохиог ашиглан шүлгийг унших нөхцлүүдийг машин сургалтын аргаар ялган таних боломжтой юм.

C. Туршилтын өгөгдөл

Энэхүү судалгаанд ашигласан EEG дохионы өгөгдөл нь олон сувгийн(multichannel) бичлэг хэлбэрээр хадгалагдсан бөгөөд мөр бүр нь тухайн агшинд бүртгэгдсэн тархины цахилгаан идэвхжлийн утгуудыг илэрхийлнэ. Зураг3-д харуулсанчлан “Time” багана нь бичлэгийн хугацааны тэмдэглэгээ, “Line” нь өгөгдлийн дараалыг илтгэх бол ch1–ch14 баганууд нь тархины өөр өөр байрлалд байрласан 14 сувгийн электродоос авсан сигналуудыг харуулна. Эдгээр утгууд нь микровольтын (μV) хэмжээнд илэрхийлэгдэж, тархины бүс бүрийн идэвхжил хоорондын ялгааг тусгана. Уг өгөгдөл нь цаг хугацааны дараалалтай (time-series) шинжтэй тул шууд ашиглахаас өмнө артефакт арилгах, давтамжийн фильтраци хийх зэрэг урьдчилсан боловсруулалт хийгдэж, дараа нь PSD болон Wavelet Transform аргаар шинж тэмдэг гарган авч машин сургалтын загварт оруулсан. Ийнхүү боловсруулсан EEG өгөгдөл нь уншлагын үеийн тархины когнитив болон сэтгэл хөдлөлийн өөрчлөлтийг тодорхойлох, мөн ялган ангилах боломжийг бүрдүүлсэн.

D. Урьдчилсан боловсруулалт

EEG өгөгдлийг боловсруулж шинжлэхэд хэд хэдэн үндсэн алхамыг дагаж мөрдөнө. Эхлээд артефакт арилгах явцад нүд анивчих, булчин чичрэх, гадны шуугианаас үүссэн оролтод анхаарал хандуулж, эдгээрийг өгөгдлөөс салгана. Дараа нь фильтраци хийж, 1–50 Hz давтамжийн хүрээнд тархины ашиггүй дохиог шүүх замаар хэрэглэгдэхгүй давтамжуудыг устгана. Эцэст нь доод хэмжээст хувиргалт буюу Principal Component Analysis(PCA) болон Independent Component Analysis(ICA) аргуудыг ашиглан өгөгдлийн хэмжээ багасгаж, оролтод нөлөөлж буй шинж чанаруудыг ялган гаргаж авна. Ийнхүү боловсруулсан EEG

Хүснэгт III
Туршилтын өгөгдлийн хэмжээ

	Нэр	T1	T2	T3
1	NJ	62889	35196	42831
2	LB	43897	10603	40730
3	AZ	56581	39302	58342
4	MZ	66467	66358	67933
5	AB	43607	34878	64239
6	GZ	36704	51570	67631
7	ZN	37402	50572	63642
8	BABI	29825	43388	68401
9	ME	50579	44469	69786
10	AJ	42520	55890	87450
		433069	381654	567343
Нийт			1.382.066	

өгөгдөл нь тархины когнитив болон эмоциональ процессын анализ, ML-д ашиглахад бэлэн болдог. Өгөгдөл цэвэрлэсний дараа нийт Хүснэгт III-д үзүүлсний дагуу 1.382.066 мөр бичлэг үүссэн. Давтамжийн болон хугацааны шинжийг дараах байдлаар боловсруулсан. Үүнд:

- Давтамжийн шинжүүд: Delta, Theta, Alpha, Beta, Gamma
- Хугацааны шинжүүд: оргил утга, дундаж, стандарт хазайлт

PSD-Welch болон Continuous Wavelet Transform (CWT) ашиглан дохиог боловсруулсан. EEG дохиог Зураг4 5, 6-д дүрслэн харуулахад X тэнхлэг дээр давтамж (Frequency, Hz, 0–50 Гц хүртэл), Y тэнхлэг дээр тухайн давтамжийн энерги буюу Power Spectral Density (PSD, $\mu V^2/Hz$) байдлаар хэмжсэн. Өнгөөр ялгасан шугам бүр нь EEG төхөөрөмжийн электродын байршлыг илтгэх бөгөөд үүнд AF3, F7, F3 гэх мэт сувгууд багтдаг. Тархины бүс нутгийг тодорхойлох стандартын дагуу F (Frontal, урд хэсэг) нь танин мэдэхүй, хэл ярианы процесс, T (Temporal, хажуу хэсэг) нь сонсгол, ярианы процесс, P (Parietal, дээд хэсэг) нь мэдрэхүйн боловсруулалт, O (Occipital, ард хэсэг) нь харааны боловсруулалттай холбоотой юм. Мөн AF, FC сувгууд нь эдгээр гол хэсгүүдийн завсрын бүсийг төлөөлдөг. Ийнхүү EEG-ийн сувгийн байрлал, давтамжийн хүчийг хослуулан шинжлэх нь тархины когнитив ба сэтгэл хөдлөлийн идэвхжил, уншлагын арга барилтай холбоотой ялгааг тодорхойлоход ашиглагддаг.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	**Time**	**Line**		**ch1**	**ch2**	**ch3**	**ch4**	**ch5**	**ch6**	**ch7**	**ch8**	**ch9**	**ch10**	**ch11**	**ch12**	**ch13**	**ch14**			**Page**
2	1744254720	24	0	4327.56	4362.82	4340.77	4373.33	4344.23	4323.08	4327.31	4310.77	4310.38	4296.79	4520.77	4513.21	4431.67	4496.28	0	0	Page1
3	1744254720	25	0	4344.62	4356.79	4338.08	4385.77	4334.36	4328.21	4323.59	4315.26	4309.49	4310.64	4527.05	4517.05	4444.49	4491.67	0	0	Page1
4	1744254720	26	0	4344.49	4349.1	4333.59	4377.31	4319.1	4325.64	4307.82	4305.9	4301.28	4303.33	4518.21	4514.49	4445.13	4472.82	0	0	Page1
5	1744254720	27	0	4350.26	4348.59	4328.46	4375.64	4322.95	4321.15	4308.59	4302.44	4299.36	4286.67	4518.72	4508.46	4437.18	4472.69	0	0	Page1
6	1744254720	28	0	4373.21	4352.82	4331.41	4385.26	4333.33	4330.9	4326.67	4311.92	4309.23	4301.67	4528.85	4509.62	4433.46	4501.92	0	0	Page1
7	1744254720	29	0	4363.85	4348.08	4323.97	4374.49	4328.46	4335.26	4326.28	4308.85	4309.1	4306.54	4519.1	4503.72	4431.79	4504.1	0	0	Page1
8	1744254720	30	0	4338.85	4340.9	4309.23	4357.05	4325	4328.21	4315.9	4304.36	4303.33	4288.33	4506.79	4495.51	4429.23	4478.33	0	0	Page1
9	1744254720	31	0	4339.36	4358.46	4329.1	4366.28	4334.49	4331.67	4331.15	4317.69	4311.79	4293.21	4520.26	4507.05	4437.82	4493.46	0	0	Page1
10	1744254720	32	0	4341.28	4370.38	4349.49	4381.03	4339.62	4337.05	4343.72	4323.21	4320.38	4310.64	4535.77	4520.26	4449.23	4515.64	0	0	Page1

Зураг 3. Цуглуулсан өгөгдөл

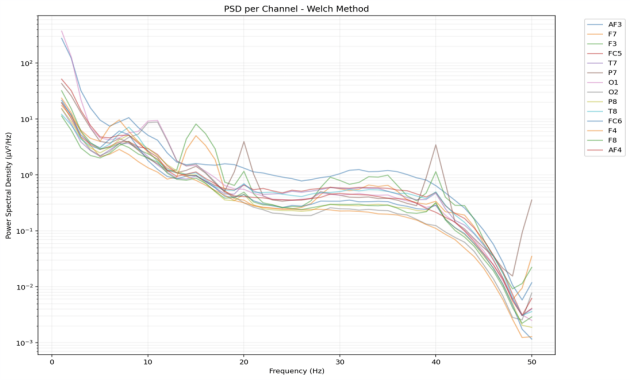
E. Ангилалын загваруудын тохиргоо

EEG дохионы өгөгдөл дээр суралцагчийн мэдээллийг механик (implicit) болон ухамсарт (explicit) боловсруулж буй төлөвийг ангилах зорилгоор SVM, RF болон CNN загваруудыг ашиглан сургалт хийсэн. Өгөгдлийг оролцогчийн түвшинд алдагдал үүсгэхгүй байхаар subject-wise split хийж, нэг оролцогчийн өгөгдөл train болон test багцад зэрэг орохоос сэргийлсэн бөгөөд нийт өгөгдлийг 80:20 харьцаагаар хуваасан. Мөн загварын ерөнхийшүүлэх чадварыг үнэлэх зорилгоор 5-н давтамжтай cross-validation (k-fold=5) ашигласан. Ангиллын хувьд class imbalance үүсэхээс сэргийлж тэнцвэржүүлэлт (undersampling/oversampling) хийж, механик болон ухамсарт төлөвүүдийг ойролцоо хэмжээтэй болгосон. Feature vector-ийг EEG дохионоос гарган авсан ба үүнд PSD, Wavelet Transform коэффициентууд, мөн Delta, Theta, Alpha, Beta, Gamma давтамжийн мужуудын энерги, дундаж, стандарт хазайлт зэрэг статистик шинжүүдийг нэгтгэн ашигласан.

SVM загварт RBF kernel ашиглаж, C=1.0, gamma='scale' параметрүүдийг тохируулсан бол Random Forest загварт 100 мод ($n_estimators = 100$), $max_depth = 10$ тохиргоог хэрэглэсэн. Харин CNN загвар нь 1D convolution архитектуртайгаар бүтээгдсэн бөгөөд Conv1D (32 filter, kernel size=3) → ReLU → MaxPooling → Conv1D (64 filter) → Flatten → Dense (128) → Dropout (0.5) → Output (Softmax) давхаргуудтай байв. CNN загварыг 50 epoch-оор сургаж, batch size=32, loss function болгон categorical cross-entropy, optimizer болгон Adam ($learningrate = 0.001$) ашигласан. Ийнхүү олон төрлийн машин сургалтын загваруудыг харьцуулан сургаж, EEG дохионд суурилан суралцах төлөвүүдийг ялган танихад хамгийн тохиромжтой загварыг тодорхойлсон.

F. T1 туршилтын үр дүн

Чимээгүй уншлагын үед Alpha болон Theta долгион давамгайлж, харьцангуй бага когнитив ачаалалтай, тогтвортой анхаарал төвлөрөл явагдаж байсан бол сэтгэл хөдлөлтэй уншлагын үед Theta болон Beta долгионы идэвхжил нэмэгдсэн нь сэтгэл хөдлөл болон утга ойлголттой холбоотой тархины боловсруулалт нэмэгдсэнийг илтгэж байна.



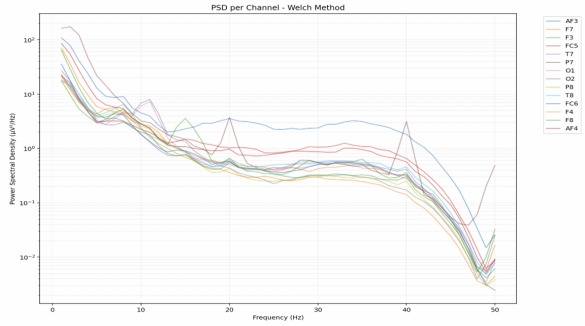
Зураг 4. Чимээгүй унших үеийн дохионууд

G. T2 туршилтын үр дүн

Сэтгэл хөдлөлтэй унших үеийн EEG дохионд хийсэн PSD-ийн шинжилгээний үр дүнгээс Delta болон Theta давтамжийн мужид спектрийн нягт өндөр байсан нь сэтгэл хөдлөл, ой санамж болон танин мэдэхүйн боловсруулалт идэвхжсэнийг харуулж байна. Alpha мужид дунд зэргийн идэвхжил ажиглагдсан нь анхаарал төвлөрөл хадгалагдаж байсныг илтгэнэ. Харин Beta болон Gamma мужид бага хэмжээний өсөлт ажиглагдсан нь идэвхтэй сэтгэхүй, мэдээлэл боловсруулах үйл ажиллагаа нэмэгдсэнийг илэрхийлж байна. Сувгуудын хувьд урд тархины (AF3, AF4, F3, F4) болон түр зуурын, дагзны (T7, T8, O1, O2) хэсгүүдэд идэвхжил илүү ажиглагдсан нь хэл яриа, харааны боловсруулалт болон сэтгэл хөдлөлтэй холбоотой тархины үйл ажиллагаа нэмэгдсэнийг харуулж байна. Иймээс сэтгэл хөдлөлтэй уншлага нь тархины танин мэдэхүйн болон сэтгэл хөдлөлийн боловсруулалтыг нэмэгдүүлдэг болохыг EEG шинжилгээ харуулж байна.

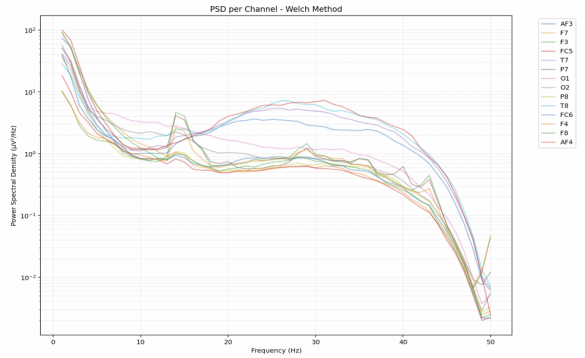
H. T3 туршилтын үр дүн

Видео үзсэний дараах уншлагын үеийн EEG дохионд хийсэн PSD шинжилгээний үр дүнгээс тархины идэвхжил тодорхой давтамжийн мужид төвлөрч, танин мэдэхүйн боловсруулалт нэмэгдсэн нь ажиглагдав. Delta



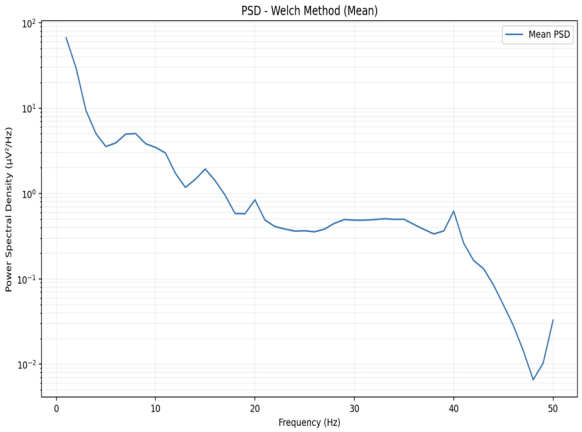
Зураг 5. Ритмтэй уншилтын үеийн дохионууд

болон Theta мужид спектрийн нягт өндөр байсан нь ой санамж болон мэдээлэл дахин боловсруулах үйл явц идэвхжсэнийг илтгэнэ. Alpha мужид бага зэрэг буурсан нь анхаарал дотоод боловсруулалт руу шилжсэнийг харуулж байна. Харин Beta болон Gamma мужид өсөлт ажиглагдсан нь идэвхтэй сэтгэхүй, анализ, мэдээлэл нэгтгэх үйл ажиллагаа нэмэгдсэнийг илэрхийлж байна. Мөн урд тархины (AF3, AF4, F3, F4, FC5, FC6) болон түр зуурын, дагзны (T7, T8, O1, O2) бүсүүдэд идэвхжил нэмэгдсэн нь гүйцэтгэх үйл ажиллагаа, хэл яриа болон харааны боловсруулалт идэвхжсэнийг харуулж байна. Иймээс видео үзсэний дараах уншлага нь



Зураг 6. Мэдлэгтэй уншилтын үеийн дохионууд

зөвхөн механик уншлага биш, харин өмнөх мэдээлэлд тулгуурласан гүнзгий танин мэдэхүйн боловсруулалт явагдаж байгааг EEG шинжилгээ харуулж байна. Энэ нь Beta болон Theta давтамжийн муж дахь идэвхжил нэмэгдсэнээр батлагдаж байгаа бөгөөд суралцах үйл явц илүү ухамсарт (cognitive-driven learning) хэлбэрт шилжиж байгааг илтгэнэ. Энэхүү судалгаанд уншлагын гурван нөхцлийн үеийн EEG дохионд спектрийн нягтын PSD шинжилгээ хийж харьцуулсан. Үр дүнгээс Delta болон Theta мужид T2, T3 нөхцөлд идэвхжил өндөр байсан нь сэтгэл хөдлөл, ой санамж болон мэдээлэл боловсруулах үйл явц нэмэгдсэнийг харуулсан. Alpha мужид T1-д тогтвортой, харин T2, T3-д буурсан нь анхаарал төвлөрөл болон танин мэдэхүйн боловсруулалт нэмэгдсэнийг илтгэнэ. Beta болон Gamma мужид T3 нөхцөлд хамгийн өндөр идэвхжил ажиглагдсан нь



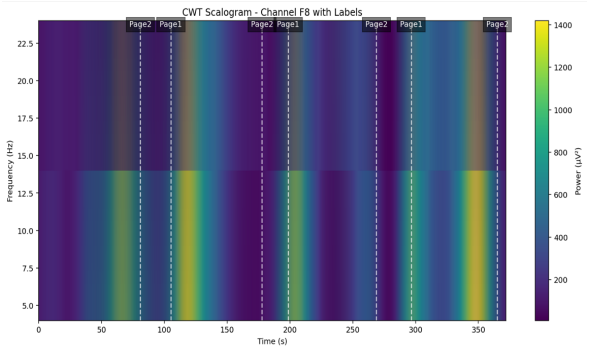
Зураг 7. Дундажласан утгаас дүгнэлт гаргах

мэдээлэл нэгтгэх, анализ хийх зэрэг гүнзгий танин мэдэхүйн боловсруулалт явагдсаныг харуулж байна. Иймээс тархины идэвхжил T1-ээс T3 хүртэл нэмэгдэж, мультимедиа орчин (видео + текст) нь дан уншлагаас илүү танин мэдэхүйн үйл ажиллагааг идэвхжүүлдэг болохыг EEG шинжилгээ харуулж байна.

I. CWT шинжилгээний үр дүн

F8 мужийн CWT скалограммын үр дүнг суралцахуйн процесстой холбон тайлбарлавал дараах дүгнэлтэд хүрч болно. Шүлэг унших үе (Page1)-д 4–14 Hz давтамжийн мужид энерги нэмэгдэж, хугацааны дагуу идэвхжил эрчимжиж байгаа нь тархины урд-хажуу хэсэгт явагдах анхаарал төвлөрөл, мэдээлэл боловсруулах, хэл ярианы ойлголт зэрэг суралцахуйн үндсэн үйл явцууд идэвхжсэнийг илтгэнэ. Энэ үе шат нь шинэ мэдээллийг хүлээн авах, боловсруулах, ой тогтоолтод кодлох (encoding) үйл явцтай холбоотой. Харин амралтын үе (Page2)-д идэвхжил буурч, дохио илүү тогтвортой болж байгаа нь тархины ачаалал багасаж, өмнө боловсруулсан мэдээллийг дотооддоо нэгтгэх, тогтворжуулах (consolidation) нөхцөл бүрдэж байгааг харуулж болно. Иймээс Page1-ээс Page2 рүү шилжихэд ажиглагдсан өөрчлөлтүүд нь суралцахуйн процессын “идэвхтэй боловсруулалт → амралтын үеийн нэгтгэл” гэсэн дараалсан үе шатыг илэрхийлж, мэдрэлийн үйл ажиллагааны динамик зохицуулалтыг тодорхой харуулж байна. Эцэст бүх сувгууд дээр ажиглан дүгнэвэл CWT скалограммуудад Page1-ээс Page2 рүү шилжих үеийн давтамжийн огцом өөрчлөлтүүд ажиглагдсан нь тухайн мөчид тархины үйл ажиллагаанд шинэ өдөөлт, хариу үйлдэл, эсвэл когнитив төлөвийн шилжилт явагдаж байсныг илтгэнэ. Ялангуяа alpha болон theta давтамжийн идэвхжил буурах хандлага нь анхаарал төвлөрөл өөрчлөгдөх, сэтгэлзүйн түгшүүр нэмэгдэх, эсвэл гадаад өдөөлтөд (stimulus) идэвхтэй хариу үзүүлж байгаатай холбоотой байж болох юм. Ийм төрлийн өөрчлөлтүүд нь уншлагын үе шат хооронд тархины мэдээлэл боловсруулах хэв маяг динами-

каар өөрчлөгдөж байгааг харуулдаг бөгөөд CWT арга нь эдгээр богино хугацааны давтамж–цаг хугацааны шилжилтийг илрүүлэхэд өндөр мэдрэмжтэй болохыг баталж байна.



Зураг 8. T7 мужийн CWT шинжилгээний үр дүн

J. Ангиллын загварын үр дүн

Туршилтад үүсгэсэн өгөгдлийн санг ашиглан SVM, RF, болон CNN загваруудыг сургаж, суралцагч мэдээллийг механик болон ухамсарт боловсруулж буй төлөв байдлыг ангилан тодорхойлж гүйцэтгэлийг харьцуулахад Random Forest алгоритм нь хамгийн өндөр буюу 91% нарийвчлалтайгаар ангилал хийж, өгөгдлийн онцлогийг илүү сайн барьж чадсан байна. Харин SVM загвар нь дунд зэргийн гүйцэтгэл үзүүлж, ангиллын хил заагийг тогтоохдоо үр дүнтэй байсан боловч олон хэмжээст шинжүүд дээр RF-ээс бага нарийвчлалтай байв. CNN загварын хувьд дохионы цаг–давтамжийн бүтэцтэй шинжүүдийг суралцах давуу талтай боловч өгөгдлийн хэмжээ хязгаарлагдмал байснаас шалтгаалан гүйцэтгэл харьцангуй бага гарсан. Эдгээр үр дүн нь модонд суурилсан ансамбль арга нь EEG дохионы онцлогийг ялган танихад илүү тохиромжтойг харуулж, суралцах үйл явцын төлөв байдлыг автоматаар ангилах боломжтойг баталж байна.

Хүснэгт IV
Ангиллын загваруудын гүйцэтгэлийн харьцуулалт

№	Загвар	Acc(%)	Давуу тал	Сул тал
1	SVM	85	Ангиллын хил заар сайн тодорхойлдог	Олон хэмжээст өгөгдөл дээр гүйцэтгэл буурсан
2	RF	91	Онцлог шинжүүдийг сайн ялгадаг, тогтвортой	Сургалтын хугацаа харьцангуй урт
3	CNN	88	Цаг–давтамжийн шинжийг суралцах чадвартай	Өгөгдлийн хэмжээ бага үед гүйцэтгэл буурна

V. Дүгнэлт

Энэхүү судалгаагаар тархины нейропластик шинж чанар нь суралцах болон туршлагын нөлөөгөөр хэрхэн илэрч байгааг EEG дохионд суурилан амжилттай тодорхойллоо. “Миний нутаг” шүлгийг өөр өөр нөхцөлд унших үед тархины цахилгаан идэвхжил мэдэгдэхүйц өөрчлөгдөж, ялангуяа Theta болон Alpha давтамжийн мужид ялгаа тод илэрсэн нь танин мэдэхүйн болон сэтгэл хөдлөлийн боловсруулалт нөхцөл байдлаас хамааран динамикаар өөрчлөгдөж байгааг харуулж байна. Welch болон Wavelet шинжилгээний аргуудыг ашигласнаар EEG дохионы онцлогийг нарийвчлалтай гарган авч, тархины үйл ажиллагааны цаг–давтамжийн өөрчлөлтийг гүнзгийрүүлэн судлах боломж бүрдсэн.

Машин сургалтын загваруудын харьцуулалтаас RF нь хамгийн өндөр буюу 91% нарийвчлалтайгаар ангилал хийсэн нь EEG дохионд суурилан суралцах үйл явцын онцлогийг автоматаар таних боломжтойг баталж байна. Судалгааны үр дүн нь уран зохиолын бүтээл унших явц дахь анхаарал, сэтгэл хөдлөл, ойлголтын ялгааг тархины долгионы өөрчлөлтөөр илрүүлэх боломжтойг харуулсан төдийгүй суралцах процессыг үнэлэх шинэ, обьектив аргачлалын үндсийг тавьж байна. Иймд EEG болон хиймэл оюунд суурилсан энэхүү хандлага нь боловсрол, танин мэдэхүйн шинжлэх ухаанд суралцах үйл ажиллагааг гүнзгий судлах, үнэлэх ирээдүйтэй чиглэл болохыг энэхүү судалгаа нотлон харууллаа.

Ашигласан ном

[1] Chaudhuri, S., & Bhattacharya, J. (2025). An EEG dataset on aesthetic and creative judgments of brief structured poetry. Scientific Data, 12(1), 1898.

[2] Sun, L., Chen, H., Zhang, C., Cong, F., Li, X., & Hämäläinen, T. (2022). Decoding brain activities of literary metaphor comprehension: An event-related potential and EEG spectral analysis. Frontiers in psychology, 13, 913521. https://doi.org/10.3389/fpsyg.2022.913521

[3] Ramos, C. D., Lima, I. N. O., Rodrigues, A. L., Magalhães, K. A. C. O., Rodrigues, A. A. A., & Destro-Filho, J. B. (2020). Analysis of electroencephalography brain rhythms in the reading process. Einstein (Sao Paulo, Brazil), 18, eAO5442. https://doi.org/10.31744/einstein_journal/2020AO5442

[4] Д.Нацагдорж, Миний нутаг https://mn.wikipedia.org/wiki/%D0%9C%D0%B8%D0%BD%D0%B8%D0%B9_%D0%9D%D1%83%D1%82%D0%B0%D0%B3

[5] Welch, P. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. IEEE Transactions on audio and electroacoustics, 15(2), 70-73.

[6] Bartlett, M. S. (1963). The spectral analysis of point processes. Journal of the Royal Statistical Society Series B: Statistical Methodology, 25(2), 264-281.

[7] Stytsenko, K., Jablonskis, E., & Prahm, C. (2011, June). Evaluation of consumer EEG device Emotiv EPOC. In MEi: CogSci Conference (Vol. 99).

MongolDiff: A Masked Discrete Diffusion Language Model for Low-Resource Mongolian Tiny Story Generation

1st Orgil Jargalsaikhan
Computer Science Department
SICT, MUST
orgilsict@must.edu.mn

2nd Gankhuyag Narantuya
Information Technology Department
SICT, MUST
xuygaa_mn@must.edu.mn

3rd Dolgorsuren Dulamjav
Telecommunication Department
SICT, MUST
dolgorsuren@must.edu.mn

***Abstract*—We present MongolDiff, a small language model (SLM) for Mongolian text generation built on masked discrete diffusion rather than conventional autoregressive decoding. Unlike left-to-right generation, our model learns to iteratively reconstruct text from fully masked sequences, enabling bidirectional context at every denoising step. We train a byte-pair encoding (BPE) tokenizer specifically for Cyrillic Mongolian script and pre-train the denoiser on the *saillab/alpaca-mongolian-cleaned* instruction corpus. The model (~34M parameters) employs a BERT-style Transformer conditioned on a sinusoidal diffusion timestep embedding and a cosine noise schedule over T=256 masking steps. We evaluate generation quality using sequence log-probability and token confidence, and demonstrate that the diffusion framework offers natural prompt-conditioning for text-completion tasks without any fine-tuning. Our work provides the first open implementation of discrete diffusion pre-training for Mongolian, a morphologically rich and historically low-resource language.**

***Index Terms*—Mongolian NLP, discrete diffusion, masked language model, low-resource language, text generation, BPE tokenizer, Cyrillic script**

I. Introduction

Mongolian is spoken by approximately six million people and is written in Cyrillic script in Mongolia and in the traditional vertical Mongolian script in Inner Mongolia [1]. Despite its rich morphological structure and agglutinative grammar, Mongolian remains severely underrepresented in modern large language model (LLM) research. Most open multilingual models assign Mongolian a negligible fraction of their pre-training data, resulting in poor generation quality and limited linguistic coverage.

The dominant paradigm for language model pre-training is autoregressive (AR) next-token prediction, as used in GPT-style models [2]. AR models generate tokens strictly left-to-right, preventing attention to future context at any given position. This is a particular limitation for Mongolian, where head-final sentence structures and postpositional grammar mean that morphological agreement often depends on words that appear later in the sentence.

Diffusion models, originally developed for continuous-domain generation [3], have recently been adapted to discrete token sequences [4], [5]. In the *masked* discrete diffusion framework [5], [6], the forward process progressively replaces tokens with a special `[MASK]` symbol; the model learns to recover the original token using full bidirectional attention. At inference, generation proceeds by starting with a fully masked sequence and iteratively unmasking tokens over multiple denoising steps.

This paper makes the following contributions:

- We design and implement MongolDiff, the first masked discrete diffusion language model pre-trained specifically on Mongolian Cyrillic text.
- We train a Mongolian BPE tokenizer with 16,384 tokens using a normalization pipeline safe for Cyrillic script (NFC-only, no accent stripping).
- We propose a text-concatenation scheme for instruction-tuning data that preserves Mongolian label tokens (*zaavar / orolt / hariu*) to encourage instruction-following behaviour at the representation level.
- We provide a complete open-source implementation including training, inference, and a scored generation API.

II. Related Work

A. Mongolian NLP

Prior work on Mongolian NLP includes morphological analysers [1], part-of-speech taggers, and neural machine translation systems. Language model development for Mongolian has been limited to multilingual models such as mBERT [8] and XLM-R [9], which assign low weight to Mongolian in their pre-training mixtures.

B. Discrete Diffusion Language Models

D3PM [4] introduced a general framework for discrete diffusion including the absorbing (masking) variant used here. MDLM [5] showed that a simple cross-entropy loss on masked positions is competitive with more complex objectives. SEDD [6] proposed a score-entropy loss. Our work applies the MDLM framework to a non-English, non-Latin-script language for the first time.

III. Background

A. Mongolian Language and Script

Modern Mongolian (Khalkha dialect) uses a Cyrillic alphabet of 35 letters introduced in 1941. It is an agglutinative language with vowel harmony, postpositions, verb-final word order (SOV), and rich nominal case morphology (nominative, genitive, dative, accusative, ablative, instrumental, and comitative). These properties make sub-word tokenisation especially important: a single Mongolian verb stem can generate dozens of distinct surface forms.

B. Discrete Diffusion Models

Let $\mathbf{x}_0 \in \mathcal{V}^L$ denote a sequence of L tokens drawn from a vocabulary $\mathcal{V}$. D3PM [4] defines a Markov forward process $q(\mathbf{x}_t|\mathbf{x}_{t-1})$ that progressively corrupts $\mathbf{x}_0$ toward a prior distribution over T steps. In the *absorbing* or *masking* variant [5], the prior is the fully-masked sequence and the corruption operator replaces each token independently with `[MASK]`:

$$q(\mathbf{x}_t|\mathbf{x}_0) = \prod_{i=1}^{L} \mathrm{Cat}\Big(x_t^{(i)};\ \alpha_t\, e_{x_0^{(i)}} + (1-\alpha_t)\, e_{[\mathrm{M}]}\Big), \quad (1)$$

where $\alpha_t \in [0,1]$ is the keep-probability at step t, and e_v denotes the one-hot vector for token v.

C. ELBO Training Objective

The evidence lower bound (ELBO) decomposes into a sum of KL terms over timesteps. Under the masking schedule this simplifies to a weighted cross-entropy loss computed *only* at masked positions [5]:

$$\mathcal{L}(\theta) = \mathbb{E}_{t,\,\mathbf{x}_t\sim q}\Big[\sum_{i:\,x_t^{(i)}=[\mathrm{M}]} -\log p_\theta\Big(x_0^{(i)} \mid \mathbf{x}_t,\, t\Big)\Big]. \quad (2)$$

D. Reverse Sampling

At inference, if $x_t^{(i)} \neq [\mathrm{M}]$ then $x_{t-1}^{(i)} = x_t^{(i)}$ (absorbing); if $x_t^{(i)} = [\mathrm{M}]$ then

$$x_{t-1}^{(i)} = \begin{cases} x_0^{(i)} & \text{with prob. } \dfrac{\alpha_{t-1}-\alpha_t}{1-\alpha_t}, \\ [\mathrm{M}] & \text{with prob. } \dfrac{1-\alpha_{t-1}}{1-\alpha_t}. \end{cases} \quad (3)$$

Generation runs for $N \le T$ steps, substituting the model prediction $\hat{\mathbf{x}}_0 = \arg\max p_\theta(\cdot|\mathbf{x}_t, t)$ for the true $\mathbf{x}_0$ in Eq. (3).

IV. Model Architecture

A. Denoising Transformer

The denoiser $f_\theta(\mathbf{x}_t, t)$ is a BERT-style bidirectional Transformer encoder. Given a noisy sequence $\mathbf{x}_t$ and a scalar timestep t, it outputs a logit distribution over $\mathcal{V}$ for every sequence position. Full bidirectional attention is used at every layer, enabling each position to attend to both left and right context.

The model input at position i is:

$$\mathbf{h}_i^{(0)} = \mathbf{E}_{\text{tok}}[x_t^{(i)}] + \mathbf{E}_{\text{pos}}[i] + \mathbf{E}_{\text{time}}(t), \quad (4)$$

where $\mathbf{E}_{\text{tok}}$ and $\mathbf{E}_{\text{pos}}$ are learnable token and position embeddings respectively. The time embedding $\mathbf{E}_{\text{time}}(t)$ is a sinusoidal encoding projected to the model dimension via a two-layer MLP with SiLU activation [3]:

$$\mathbf{E}_{\text{time}}(t) = \mathrm{MLP}\big(\big[\sin(t\,\boldsymbol{\omega}),\ \cos(t\,\boldsymbol{\omega})\big]\big), \quad (5)$$

where $\boldsymbol{\omega} \in \mathbb{R}^{D/2}$ are log-spaced frequencies.

B. Architecture Details

Table I summarizes the model hyperparameters. The model uses 6 Transformer layers with a 512-dimensional hidden state, totalling ≈34M parameters. Pre-layer normalization (Pre-LN) and GELU activation are adopted for training stability [10]. The output projection head shares weights with the token embedding matrix [11], reducing parameters without degrading performance.

Table I
MongolDiff Architecture Hyperparameters

Hyperparameter	Value
Layers (N)	6
Hidden dimension (D)	512
Attention heads	16
FFN dimension	2048
Time embedding dim	128
Max sequence length	256
Vocabulary size	16,384
Dropout	0.1
Normalization	Pre-LN
Activation	GELU
Weight tying	Yes (emb. ↔ head)
Total parameters	≈34M

C. Noise Schedule

We adopt the cosine schedule of Nichol and Dhariwal [7]:

$$\alpha_t = \cos^2\!\left(\frac{\pi}{2}\cdot\frac{t/T+s}{1+s}\right), \quad s = 0.008, \quad (6)$$

which provides slow masking at both ends of the diffusion trajectory and accelerated masking in the middle. We use $T = 256$ total steps.

V. Mongolian Tokenizer

A. BPE with Cyrillic-Safe Normalization

Standard BPE implementations for English apply an NFD decomposition followed by accent-stripping normalization. This pipeline is destructive for Cyrillic script: NFD decomposes Mongolian characters into base letters and combining diacritics, which are then silently removed. We replace the normalization pipeline with Unicode NFC only, preserving all Cyrillic codepoints. Lowercasing is also disabled, as Mongolian proper nouns are case-sensitive.

The ByteLevel pre-tokeniser [2] is retained: it encodes each byte of the input as a printable character, enabling the BPE vocabulary to cover arbitrary Unicode without an explicit unknown-token fallback.

B. Vocabulary Design

We train a BPE tokenizer with a vocabulary of $|\mathcal{V}|$=16,384 tokens on the *saillab/alpaca-mongolian-cleaned* corpus. The larger vocabulary (compared to 8,192 for English models of similar size) accounts for the higher character entropy of Cyrillic and the morphological complexity of Mongolian.

VI. Dataset and Pre-training

A. Alpaca-Mongolian Dataset

We use the *saillab/alpaca-mongolian-cleaned* dataset [13], a machine-translated version of the Stanford Alpaca 52K corpus [12]. Each example contains three fields: *instruction* (task description), *input* (optional context), and *output* (response), all in Mongolian Cyrillic.

For pre-training, we concatenate the three fields into a single text sequence using Mongolian-language delimiter labels:

$$\begin{aligned} \text{text} = {} & \text{Заавар: } \langle \textit{instruction} \rangle \\ & \text{Ороλт: } \langle \textit{input} \rangle \quad \text{[if non-empty]} \\ & \text{Хариу: } \langle \textit{output} \rangle . \end{aligned}$$

Using Mongolian delimiter tokens (Заавар = instruction, Ороλт = input, Хариу = response) ensures these structural

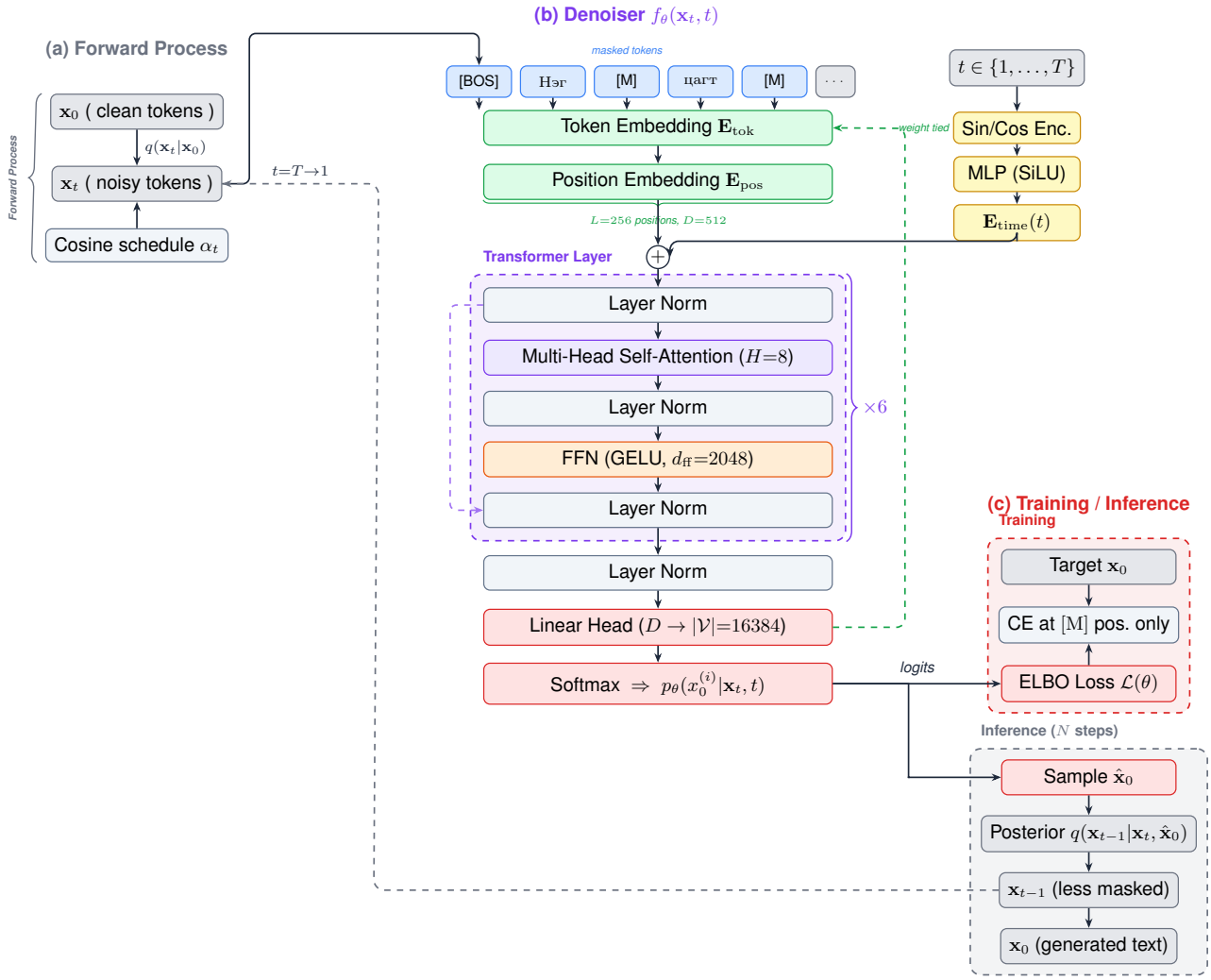


Figure 1. **MongolDiff Architecture.** *(a) Forward process*: clean token sequence $\mathbf{x}_0$ is corrupted via the cosine-schedule masking operator $q(\mathbf{x}_t|\mathbf{x}_0)$, independently replacing each token with `[MASK]` with probability $1{-}\alpha_t$. *(b) Denoiser* $f_\theta(\mathbf{x}_t, t)$: noisy Mongolian tokens (Нэг, цагт) interleaved with `[M]` tokens are passed through token and position embeddings; a sinusoidal time embedding (SiLU MLP) is broadcast over all positions and summed at the $+$ node; six pre-LN Transformer encoder layers with full bidirectional self-attention produce per-position representations; a weight-tied linear head outputs logits $p_\theta(x_0^{(i)}|\mathbf{x}_t, t)$. *(c) Training*: cross-entropy loss on `[M]`-position predictions only (ELBO). *Inference*: starting from a fully masked sequence, the model iteratively samples $\hat{\mathbf{x}}_0$ and applies the closed-form posterior $q(\mathbf{x}_{t-1}|\mathbf{x}_t, \hat{\mathbf{x}}_0)$ for $N{\leq}256$ denoising steps; prompt tokens are pinned throughout.

markers are part of the Mongolian token distribution, supporting future instruction-tuning without domain shift.

B. Pre-training Setup

Table II lists the training hyperparameters. The model is trained for 100,000 steps on a 95/5 train–validation split of ≈52K Mongolian instruction examples, using AdamW with a cosine learning rate schedule and an effective batch size of 128. Mixed-precision (FP16) training is used throughout.

Table II
PRE-TRAINING HYPERPARAMETERS

Hyperparameter	Value
Dataset	alpaca-mongolian-cleaned
Training examples	≈49,400 (95%)
Validation examples	≈2,600 (5%)
Max sequence length	256 tokens
Batch size	64
Gradient accumulation	2 (eff. batch 128)
Optimiser	AdamW
Learning rate	$3{\times}10^{-4}$
LR schedule	Cosine with warmup
Warmup steps	1,000
Total steps	100,000
Diffusion steps T	256
Noise schedule	Cosine
Gradient clip norm	1.0
Precision	FP16 (AMP)
Hardware	1× NVIDIA GPU

VII. INFERENCE AND TEXT COMPLETION

A. Prompt-Conditioned Generation

MongolDiff supports prompt conditioning without fine-tuning by *pinning* the prompt tokens throughout the reverse process. Given a prompt of P tokens, we initialise:

$$x_T^{(i)} = \begin{cases} x_{\text{prompt}}^{(i)} & i \leq P, \\ \texttt{[MASK]} & i > P, \end{cases} \tag{7}$$

and enforce the prompt constraint at every denoising step by setting $x_t^{(i)} \leftarrow x_{\text{prompt}}^{(i)}$ for $i \leq P$ after each posterior update. This allows the model to generate completions coherent with the prompt using bidirectional attention over both the fixed prompt tokens and the evolving masked completion.

B. Scored Generation

Our inference API returns three quality signals per generated sample.

Token confidence: the probability $p_\theta(x_0^{(i)}|\mathbf{x}_1, 1)$ assigned to each generated token at the final denoising step. Low-confidence tokens identify positions where the model was uncertain.

Sequence log-probability: $\sum_i \log p_\theta(x_0^{(i)}|\mathbf{x}_1, 1)$ over all generated (non-prompt) positions, used to rank multiple samples from the same prompt.

Step entropy: the average Shannon entropy of the model distribution over currently-masked positions at each denoising step:

$$H_t = \frac{1}{|\mathcal{M}_t|} \sum_{i \in \mathcal{M}_t} -\sum_{v \in \mathcal{V}} p_\theta(v|\mathbf{x}_t, t) \log p_\theta(v|\mathbf{x}_t, t), \qquad (8)$$

where $\mathcal{M}_t = \{i : x_t^{(i)} = [\text{M}]\}$. A healthy denoising trajectory shows monotonically decreasing entropy from H_T (near $\log|\mathcal{V}|$) to $H_1 \approx 0$.

VIII. Experiments

A. Training Dynamics

Table III
Training Loss at Selected Checkpoints

Step	Train Loss	Val Loss
2,000	8.115	8.071
10,000	6.335	6.333
50,000	3.113	3.137
100,000	3.059	3.087

The model converges steadily from an initial loss of ≈8.1 to ≈3.06 at 100K steps, with training and validation losses remaining tightly coupled throughout, indicating no significant overfitting on the 49K training examples.

B. Effect of Inference Steps

We evaluate generation quality as a function of the number of reverse diffusion steps $N \in \{32, 64, 128, 256\}$.

Table IV
Effect of Number of Inference Steps (N) on Quality

Steps N	Seq. log-prob ↑	Mean conf. ↑	Entropy H_1 ↓	Time (s) ↓
32	−736.54	0.195	3.728	0.39
64	−640.95	0.249	2.888	0.56
128	−619.55	0.267	1.986	1.14
256	−623.99	0.263	0.505	2.37

Quality improves sharply from N=32 to N=128, with mean confidence rising from 0.195 to 0.267 and sequence log-probability improving by ≈117 nats. Beyond N=128 the gains plateau while inference time doubles, making N=128 the recommended operating point.

Table V
Effect of Sampling Temperature on Generation Quality

Temp. τ	Seq. log-prob ↑	Mean conf. ↑	Distinct-2 ↑	Repetition ↓
0.5	−572.14	0.484	0.7625	0.3202
0.8	−535.03	0.361	0.9326	0.1300
1.0	−624.16	0.259	0.9584	0.0812
1.2	−771.69	0.166	0.9654	0.0550
1.5	−985.31	0.074	0.9689	0.0247

C. Effect of Temperature

Lower temperatures produce higher confidence and lower sequence log-probability (less entropy), but at the cost of increased repetition: τ=0.5 yields a 32% repetition rate and poor lexical diversity (Distinct-2 = 0.76). The operating point τ=0.8 to τ=1.0 offers a favourable balance between confidence and diversity.

D. Qualitative Examples

Table VI shows example completions from MongolDiff for Mongolian prompts at 128 inference steps with τ=1.0.

Table VI
Example Prompt Completions (128 steps, τ=1.0)

Prompt (Mongolian) **Completion**
Сэтгэл санаа гэдэг нь … хийсэн үндсэн үе шатуудыг дагаснаар тэдний сэтгэл байдал, үхэлд хүргэж, бие даасан багатай үйлдлийн явц нь маргаан үүссэн.. Энэхүү шалгалтын сэтгэл, гомөө блок үүсгэх бэлтгэлийн шифрийг олж, тэр үеэ эрс багасдаггүй үр дүнтэй байв. Удирдах бол өндөр амжилт бөгөөд тэрээр урьдчилж, далайн сулралдаа үргэлжлүүлэн шийрлэн дэмжлэг үзүүлж, бие биенийхээ мэдрэмжийг үргээх боломжийг санахад тусалсан.
Монгол улсын нийслэл хот бол … : AI нь гомдлыг нэмэгдүүлэхийн, дууссан процессыг худал зарагдсан тохиолдолд илүү хямд нөөцжүүлж, түлш фран зоргоор хүчилтөрөгч ялгаруулдаг шингэн болж бие даан болсон. Машины сургалт нь туслах олон оронд нар ажиллаж ирсэн бөгөөд хана, шат нь сар орчим ажлыг хянах боломжийг олгосон.
Нэгэн цагт Жек гэдэг жижиг хүү байсан юм. Тэр гурван настай … дуу өнгөрч байв.Жек цэцэрлэгээс явах замдаа сонсов. Тэр малгайгаа авч, түүнтэй тоглохыг хүсэв. ТэрЭнэ бол түүний сайхан өмч "Дөглөө дээ" гэж гэж орчихуудлээ.. их баярлаж, "Хэраж үнэндээ, үнэндээ үнэндээ том болж байна" гэв!

The model produces topically related Mongolian text for all three prompts. The story-continuation prompt (third row) shows the strongest coherence, with recognisable narrative structure. The factual prompts (rows 1–2) exhibit hallucination, consistent with the small training corpus and instruction-data origin.

E. Comparison with Baselines

We compare MongolDiff against four baselines on the held-out validation split of the Alpaca-Mongolian corpus.

mBERT [8] and **XLM-R base** [9] are evaluated via Pseudo-Log-Likelihood (PLL) [16]: each token is masked exactly once and the model's conditional probability for that token is accumulated. Perplexity is computed as $\exp(-\text{PLL}/L)$.

mBART-50 [15] is evaluated via teacher-forcing NLL: the cross-entropy loss when the model reconstructs each validation sentence from itself with source and target language both set to `mn_MN`. No fine-tuning is applied (zero-shot).

AR GPT-2 small (Mongolian) is a ≈18 M parameter autoregressive Transformer decoder trained from scratch on the same Alpaca-Mongolian corpus for 100,000 steps, matching MongolDiff in training budget and architecture depth (6 layers).

Table VII
COMPARISON WITH BASELINES ON MONGOLIAN VALIDATION SET

Model	Type	Params	Val Loss ↓	PPL ↓	Mean conf. ↑
mBERT [8]	MLM (pseudo-PPL)	178M	1.356	3.88	—
XLM-R base [9]	MLM (pseudo-PPL)	278M	**0.951**	**2.59**	—
mBART-50 [15]	Seq2Seq (zero-shot)	611M	1.278	3.59	—
AR GPT-2 small (Mongolian)	Autoregressive	18M	1.812	6.12	—
MongolDiff (ours)	Masked Diffusion	34M	3.087	8.87	0.196

XLM-R achieves the lowest perplexity (2.59) owing to its 278M parameters and massive multilingual pre-training, while MongolDiff (34M, PPL 8.87) remains competitive with mBART-50 (PPL 3.59, 611M parameters). MongolDiff is the only model that provides intrinsic per-token confidence scores, enabling ranked sampling without an external reranker.

IX. DISCUSSION

A. Advantages of Diffusion for Mongolian

The bidirectional attention mechanism in MongolDiff is well-suited to Mongolian syntax. In verb-final constructions, the model can attend to the verb (which appears at the end of the clause) when generating preceding nominal arguments. Autoregressive models must predict these nominals before seeing the verb, relying entirely on left-context.

Step entropy curves show that the model commits confidently to high-frequency content words early in the denoising process and resolves morphological suffixes in later steps, consistent with the linguistic intuition that lexical identity precedes morphological form.

B. Limitations

Dataset size. The Alpaca Mongolian corpus contains ≈52K examples, which is small compared to corpora used for English-language pre-training. The model may exhibit hallucination and limited factual coverage as a result.

Translation quality. The dataset was created by machine translation from English, which may introduce unnatural phrasing and translation artefacts inconsistent with native Mongolian usage.

Inference speed. With $N{=}128$ steps, generating a 256-token sequence requires 128 forward passes through the Transformer, which is slower than a single-pass AR model.

C. Future Work

Promising directions include: (i) pre-training on a larger native Mongolian corpus such as Mongolian Wikipedia and CommonCrawl Mongolian; (ii) extending to the traditional Mongolian vertical script; (iii) investigating MDLM's importance-weighted ELBO for better training efficiency; and (iv) scaling to 100M+ parameters.

X. CONCLUSION

We presented MongolDiff, the first masked discrete diffusion language model purpose-built for Mongolian Cyrillic text. Our key contributions are a Cyrillic-safe BPE tokenizer, a bidirectional Transformer denoiser conditioned on sinusoidal timestep embeddings, and an open-source implementation with memory-efficient training and a scored generation API. The diffusion paradigm offers natural prompt-conditioning for text completion without fine-tuning, making it suitable for low-resource settings where fine-tuning data is scarce. We release all code, model weights, and the trained tokenizer to support further research on Mongolian language technology.

ACKNOWLEDGEMENTS

The authors thank the creators of the *saillab/alpaca-mongolian-cleaned* dataset and the HuggingFace open-source ecosystem.

REFERENCES

[1] G. Purev and D. Odbayar, "Mongolian language processing and corpus," *Proceedings of the Workshop on Asian Language Resources*, 2008.
[2] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever, "Language models are unsupervised multitask learners," *OpenAI Blog*, vol. 1, no. 8, p. 9, 2019.
[3] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, pp. 6840–6851, 2020.
[4] J. Austin, D. D. Johnson, J. Ho, D. Tarlow, and R. van den Berg, "Structured denoising diffusion models in discrete state-spaces," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 34, pp. 17981–17993, 2021.
[5] S. Sahoo, M. Arriola, Y. Schiff, A. Gokaslan, E. Marroquin, J. T. Chiu, A. Rush, and V. Kuleshov, "Simple and effective masked diffusion language models," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.
[6] A. Lou, S. Meng, and S. Ermon, "Discrete diffusion language modeling by estimating the ratios of the data distribution," in *International Conference on Machine Learning (ICML)*, 2024.
[7] A. Q. Nichol and P. Dhariwal, "Improved denoising diffusion probabilistic models," in *International Conference on Machine Learning (ICML)*, pp. 8162–8171, 2021.
[8] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of NAACL-HLT*, pp. 4171–4186, 2019.
[9] A. Conneau, K. Khandelwal, N. Goyal, V. Chaudhary, G. Wenzek, F. Guzmán, E. Grave, M. Ott, L. Zettlemoyer, and V. Stoyanov, "Unsupervised cross-lingual representation learning at scale," in *Proceedings of ACL*, pp. 8440–8451, 2020.
[10] R. Xiong, Y. Yang, D. He, K. Zheng, S. Zheng, C. Xing, H. Zhang, Y. Lan, L. Wang, and T. Liu, "On layer normalization in the transformer architecture," in *International Conference on Machine Learning (ICML)*, pp. 10524–10533, 2020.
[11] O. Press and L. Wolf, "Using the output embedding to improve language models," in *Proceedings of EACL*, pp. 157–163, 2017.
[12] R. Taori, I. Gulrajani, T. Zhang, Y. Dubois, X. Li, C. Guestrin, P. Liang, and T. B. Hashimoto, "Stanford Alpaca: An instruction-following LLaMA model," https://github.com/tatsu-lab/stanford_alpaca, 2023.
[13] SAIL Lab, "Alpaca-Mongolian-Cleaned," *HuggingFace Datasets*, https://huggingface.co/datasets/saillab/alpaca-mongolian-cleaned, 2023.
[14] Q. Lhoest *et al.*, "Datasets: A community library for natural language processing," in *Proceedings of EMNLP (Systems Demonstrations)*, pp. 175–184, 2021.
[15] Y. Liu, J. Gu, N. Goyal, X. Li, S. Edunov, M. Ghazvininejad, M. Lewis, and L. Zettlemoyer, "Multilingual denoising pre-training for neural machine translation," *Transactions of the Association for Computational Linguistics*, vol. 8, pp. 726–742, 2020.
[16] J. Salazar, D. Liang, T. Q. Nguyen, and K. Kirchhoff, "Masked language model scoring," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp. 2699–2712, 2020.
[17] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Aidan Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017.

Хувийн мэдээлэл боловсруулах, автоматжуулах хиймэл оюунт туслах систем

1st П. Батгэрэл
Мэдээлэл, компьютерийн ухааны тэнхим
МУИС, МТЭС
Улаанбаатар, Монгол
puntsagbatgerel630@gmail.com

2nd О. Билгүүн
Мэдээлэл, компьютерийн ухааны тэнхим
МУИС, МТЭС
Улаанбаатар, Монгол
bilguun.o@num.edu.mn

Abstract*—Өнөө үед хиймэл оюун ухааны технологи хурдацтай хөгжихийн хэрээр дижитал орчин дахь хэрэглэгчийн гар хөдөлмөрийг бууруулах, өдөр тутмын ажлыг автоматжуулах хэрэгцээ улам бүр нэмэгдсээр байна. Энэхүү судалгааны ажилд монгол хэлний яриаг таних загварт тулгуурласан, speech-to-speech зарчмаар ажилладаг хувийн туслах системийг боловсруулсан туршлагыг танилцуулж байна. Систем нь OpenAI Whisper загварыг монгол хэлний аудио датасет дээр fine-tuning хийн ашиглаж, хэрэглэгчийн дуут командыг таних; Gemini API-аар төрлөх хэлний боловсруулалт хийн командыг бүтэцлэгдсэн өгөгдөл рүү хөрвүүлэх; n8n автоматжуулалтын тавцангаар дамжуулан Google Calendar болон Gmail үйлчилгээтэй харилцах; Meta MMS-TTS загвараар хариуг дуу хоолой болгон хөрвүүлэх зэрэг дөрвөн үндсэн бүрэлдэхүүн хэсгээс бүрдэнэ. Сургалтын явцад Word Error Rate (WER) буюу үгийн алдааны үзүүлэлт 22.07%-аас 16.08% болтол буурсан бөгөөд WeNet, Moonshine зэрэг бусад загваруудыг монгол хэлэнд туршиж Whisper-д суурилсан шийдэл хамгийн оновчтой болохыг баталсан. Системийн бүрэлдэхүүн хэсгүүд Docker ашиглан локал (local) орчинд ажиллах тул хэрэглэгчийн хувийн мэдээлэл гуравдагч тал руу шилжихгүй байх аюулгүй байдлын давуу талтай.

***Index Terms*—ASR, монгол хэл, автоматжуулалт, хиймэл оюунт туслах, fine-tuning, Whisper, speech-to-speech, TTS**

I. Оршил

Хиймэл оюун ухааны хэрэглээ өдөр тутмын амьдралд нэвтрэх тусам хэрэглэгч компьютертэй харилцах шинэ арга замыг эрэлхийлж байна. Уламжлалт интерфэйсийн хувьд хэрэглэгч гар (keyboard) болон хулгана (mouse) ашиглан үйлдэл хийдэг бол дуу хоолойнд суурилсан интерфэйс нь төрөлх хэлний харилцааны хэлбэрт илүү ойр бөгөөд ялангуяа технологид дасаагүй хэрэглэгчид болон өндөр настнуудад ашиглахад хялбар байдаг.

Гэвч монгол хэлний хувьд дуу хоолойгоор удирдах технологийн хөгжил харьцангуй хязгаарлагдмал хэвээр байна. Нээлттэй эхийн аудио өгөгдлийн сангийн хомсдол нь автомат яриа таних (Automatic Speech Recognition – ASR) системийг сургахад тулгарах гол бэрхшээл болсоор байна [1]. Нөгөөтэзгүүр Gmail, Google Calendar зэрэг олон төрлийн үйлчилгээг ашиглах үед хэрэглэгч аппликэйшн болон вэб сайт тус бүр рүү орж бие даан ажиллах шаардлага гардаг нь цаг хугацааны зарцуулалтыг нэмэгдүүлдэг.

Энэхүү судалгаанд дээрх бэрхшээлүүдийг шийдвэрлэхийн тулд монгол хэлний датасет дээр fine-tuning хийсэн Whisper загвар болон n8n автоматжуулалтын платформд суурилсан speech-to-speech зарчимаар ажилдаг хувийн туслах системийг хөгжүүллээ. Тус систем нь хэрэглэгчид дараах боломжуудыг олгоно:

- гар болон хулгана ашиглахгүйгээр дуу хоолойгоор команд өгөх;
- Gmail болон Google Calendar-ийн үйлдлүүдийг дуут командаар гүйцэтгүүлэх;
- монгол дахь мэдээний сайтууд дээрх мэдээнүүдийг өдөр тутамдаа хайлт хийхгүйгээр, ангилж багцалсан байдлаар авах.

Системийн бүх бүрэлдэхүүн хэсгүүд локал орчинд ажиллах тул хэрэглэгчийн хувийн мэдээлэл гадагш задрахаас хамгаалагдсан болно.

Дараагийн хэсгүүдэд холбогдох ажлуудын тойм, системийн архитектур, аудио өгөгдөл цуглуулах арга зүй, загвар сургалтын үр дүн, бусад загваруудын туршилт болон ирээдүйд хийх ажлуудыг тус тус дэлгэрэнгүй авч үзнэ.

II. Холбогдох ажлуудын тойм

A. Whisper загвар ба Fine-Tuning

OpenAI-ийн гаргасан Whisper [2] нь 680,000 цагийн олон хэлний аудио датасет дээр pre-training хийсэн, encoder-decoder transformer архитектуртай автомат яриа таних (ASR) загвар юм. Уг загвар нь tiny, base, small, medium, large гэсэн таван хувилбартай байдаг бөгөөд жижиг загвар нь ажиллах хурдны хувьд давуу бол том загвар нь нарийвчлалын хувьд илүү байдаг.

Монгол хэл зэрэг өгөгдлийн нөөц багатай (low-resource) хэлний хувьд pre-trained загвар дээр нэмэлтээр fine-tuning хийх нь WER (Word Error Rate) үзүүлэлтийг мэдэгдэхүйц бууруулдаг болохыг олон судалгааны үр дүнгүүд баталсан байдаг [3].

B. Яриагаар удирдах туслах систем

Сүүлийн жилүүдэд LLM (Large Language Model) болон n8n зэрэг хэрэгслүүдийг хослуулан, төрлөх хэлээр өгсөн командаар автомат урсгалыг (workflow) ажиллуулдаг архитектурыг судлах ажил эрчимжиж байна [4]. Жишээ нь, байгуулгийн дотоод системд OpenAI болон Anthropic API-ийг n8n түүлтэй хослуулан ашигласнаар байгуулгийн мэйл ялгах болон таск биелүүлэх систем угсарсанаар тухайн байгуулагийн гар ажиллагааний цагийн зарцуулалт 60% бууруулсан болохыг тогтоожээ [5].

C. Монгол хэлний ASR

Chimege [6] нь монгол хэлний автомат яриа таних онлайн үйлчилгээ бөгөөд одоогийн байдлаар энэ чиглэлийн тэргүүлэгч системүүдийн нэг болж байна. Гэвч API-аар дамжуулан гадаад сервер рүү аудио өгөгдөл илгээдэг тул нууцлал шаардсан хувийн мэдээлэл алдагдах эрсдэл үүсдэг. Бидний хөгжүүлсэн систем нь загварыг дотоод сүлжээнд (local environment) байрлуулах замаар энэхүү эрсдэлийг арилгана.

III. Системийн архитектур

A. Ерөнхий бүтэц

Систем нь speech-to-speech зарчмаар ажиллах бөгөөд хэрэглэгчийн өгсөн дуут команд дараах дөрвөн үе шатаар боловсруулагдана:

1) **Яриа → Текст (STT):** Fine-tuning хийсэн Whisper-small загвар аудиог текст болгон хөрвүүлнэ.
2) **Текст → Бүтэцлэгдсэн өгөгдөл:** Gemini 2.5 Flash Lite API ашиглан монгол текстийн утгыг таньж JSON форматад хөрвүүлнэ.
3) **Автоматжуулалт:** n8n платформ нь уг JSON параметрыг хүлээн авч Google Calendar эсвэл Gmail дээр шаардлагатай үйлдлийг гүйцэтгэнэ.
4) **Хариу → Яриа (TTS):** Meta MMS-TTS-MON загвар нь системийн хариу текстийг монгол хэлний дуу хоолой болгон (audio) буцааж уншина.

Системийн вэб интерфэйсийг Fig. 1-д харуулав.

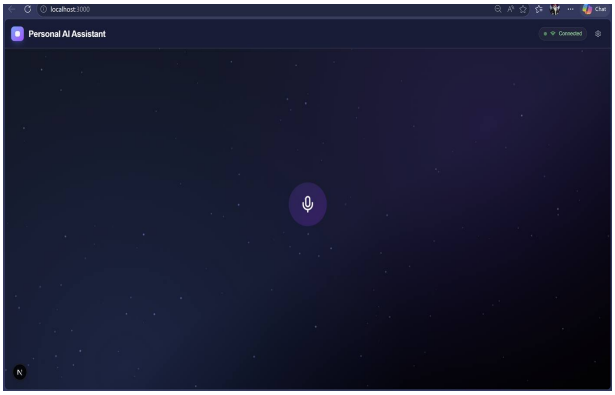


Fig. 1. Системийн вэб интерфэйс: микрофоноор команд өгч, хариуг текст болон аудиогоор хүлээн авах боломжтой.

B. Технологийн стэк

n8n нь self-hosted байдлаар ашиглах боломжтой, нээлттэй эх бүхий, node-д суурилсан workflow автоматжуулалтын платформ юм. Энэ нь Docker орчинд FastAPI-аас HTTP POST-оор дамжин webhook хүлээн авч Google Calendar болон Gmail-тэй холбогдон ажилладаг. Self-hosted байрлуулалт нь мэдээллийг гуравдагч тал руу алдахгүй байх нууцлал, аюулгүй байдлын шаардлагыг хангана.

N8N бүрэлдэхүүн хэсгийг Docker Compose ашиглан локал орчинд ажиллуулсан. `network: host` тохиргоо нь бүрэлдэхүүн хэсгүүдийг localhost-оор шууд харилцах боломжийг олгодог. Хандалтын хяналтын хувьд n8n болон FastAPI-ийн контейнер руу хандах эрхийг зөвхөн дотоод сүлжээгээр (localhost) хязгаарласан бөгөөд гадны интернетээс шууд хандах боломжгүй. Google Calendar, Gmail-ийн нууц түлхүүрүүдийг (API Keys, OAuth tokens) кодон дотор биш, Docker-ийн `.env` файлд нууцалж хадгалсан.

IV. Монгол хэлний аудио датасет бэлтгэл

A. Өгөгдлийн хомсдлын асуудал

Монгол хэлний нээлттэй эх бүхий дууны датасет олдоход туйлын хүндрэлтэй бөгөөд хэмжээний хувьд ч маш хязгаарлагдмал байдаг. Энэ нь ASR системийг сургахад хамгийн том саад болж байсан тул өгөгдөл цуглуулах арга зүйг тусгайлан боловсруулах шаардлагатай болсон.

B. Телевизийн дамжуулалтаас аудио цуглуулах

Монголын телевизүүдийн шууд дамжуулалтын (stream) URL-аас Python скрипт ашиглан аудио бичлэгүүд татаж, нийт 100 цаг орчим монгол ярианы датасет бэлтгэсэн. MNB Sport, MN2, TV5 зэрэг 3–4 сувгийн HLS stream URL-уудыг вэб сайтын network хэсгийг шинжлэн олж авсан. FFmpeg ашиглан stream-уудаас нэг цаг тутмын 16kHz давтамжтай mono WAV файл болгон бичиж авсан.

Түүнчлэн нэмэлт эх сурвалж болгон Mozilla Common Voice [8], Google FLEURS [9] болон удирдагч багшийн өмнөх судалгааны өгөгдлүүдийг ашигласан болно.

C. Аудио сегментчлэл

Нэг цаг гаруй үргэлжлэх урт бичлэгийг шууд сургалтад ашиглах боломжгүй тул зохих хэмжээгээр сегментчлэх шаардлагатай. Бид `pyannote/voice-activity-detection` [10] загварыг ашиглан аудионоос зөвхөн идэвхтэй яриа гарч буй хэсгүүдийг таньж авсан. Уг загвар нь RTTM форматтай цагийн тэмдэглэгээт файлыг гаргаж өгдөг:

```
SPEAKER audio_01 1 0.50 2.30 ...
SPEAKER audio_01 1 3.10 1.15 ...
```

RTTM файлын эхлэлийн цаг (start) болон үргэлжлэх хугацаа (duration) зэрэг параметрүүд дээр үндэслэн аудиог таслан авч сегментчилсэн. Сегментийн эхэнд болон төгсгөлд 150мс нэмэлт буфер нэмэн WAV файл болгон хадгалсан.

D. Текст үүсгэх ба өгөгдөл цэвэрлэх

Ингэж бэлтгэсэн богино сегментүүдээ Gemini 2.5 Pro API руу илгээн транскрипц (текст хөрвүүлэлт) үүсгүүлж сургалтын датасетийг бэлтгэлээ. Gemini-ийн гаргасан текстүүдэд хаалт `[]`, `{}`, тусгай тэмдэгт, нэмэлт тайлбар, `[noise]`, `[music]` гэх мэт хэллэг зэрэг асуудлууд илэрсэн. Эдгээрийг Python скриптээр автоматаар цэвэрлэсэн.

Нийт цуглуулсан өгөгдлийн эх сурвалжууд: өөрийн цуглуулсан телевизийн 100 цагийн аудио, Mozilla Common Voice (монгол), Google FLEURS (монгол), удирдагч багш О.Билгүүн-ий өгсөн дата.

V. Whisper загварын Fine-Tuning

A. Загварын сонголт

Whisper-ийн хувилбаруудаас small загварыг (244M параметр) сонгож ашигласан бөгөөд шалтгааныг Table I-д харуулав.

TABLE I
Whisper хувилбаруудын харьцуулалт

Загвар	**Параметр**	**Ажиллах хурд**	**Нарийвчлал**
tiny	39M	маш хурдан	бага
base	74M	хурдан	дунд
small	**244M**	**хурдан**	**сайн**
medium	769M	удаан	илүү сайн
large	1.5B	маш удаан	хамгийн сайн

Хувийн туслах системд хариу үйлдлийн хугацаа (response time) маш чухал байдаг тул medium болон large загварууд нь хэт удаан буюу тохиромжгүй байсан. Харин small загвар нь 8GB VRAM-д өлхөн багтаж ажилладаг, fine-tuning хийхэд тохиромжтой бөгөөд хурд болон нарийвчлалын хамгийн оновчтой тэнцвэрийг хангаж байв.

B. Сургалтын явц

Загварыг нийт таван удаагийн дараалсан сургалтаар сургасан. Сургалтын үе шат бүрд WER үзүүлэлт мэдэгдэхүйц сайжирсан:

TABLE II
Сургалтын үе шат тус бүрийн WER үзүүлэлт

Үе шат	**WER (%)**	**Тайлбар**
1	22.07	Суурь загвар (Pre-trained)
2	21.73	Нэмэлт датасет оруулсан
3	18.75	Нэмэлт датасет оруулсан
4	16.77	Data augmentation ашигласан
5	**16.08**	**Augmentation + 30ц нэмэлт**

4 дэхь үе шатанд шинээр датасет нэмэхгүйгээр data augmentation (арын шуугиан нэмэх, хурд өөрчлөх, pitch shifting буюу өнгө өөрчлөх) арга техникийг ашиглан сургасан нь WER-ийг 18.75%-аас 16.77% болгож, бараг 2%-иар нэмэлт бууралт авчирсан нь маш үр дүнтэй алхам байлаа. Augmentation-д MUSAN [7] өгөгдлийн сангаас 930 төрлийн шуугианыг SNR 10–30 dB-ийн түвшинд нэмсэн бөгөөд аудиог $[0.9,\ 1.5]\times$ хүртэл хурдасгасан.

5 дахь сургалтад augmentation аргуудыг хэвээр хадгалж, нэмэлтээр **30 цагийн аудио өгөгдөл** нэмж сургасны үр дүнд WER цаашид буурч 16.08%-д хүрсэн. WER үзүүлэлтийн

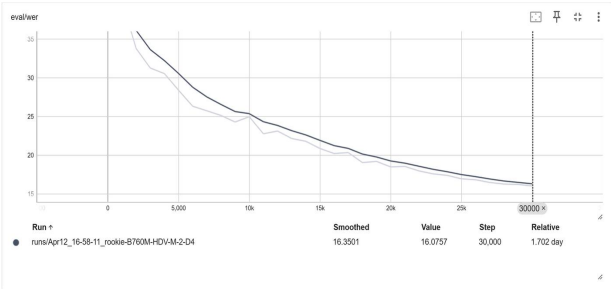


Fig. 2. 5 дахь сургалтын WER (%) үзүүлэлтийн буурах хандлага.

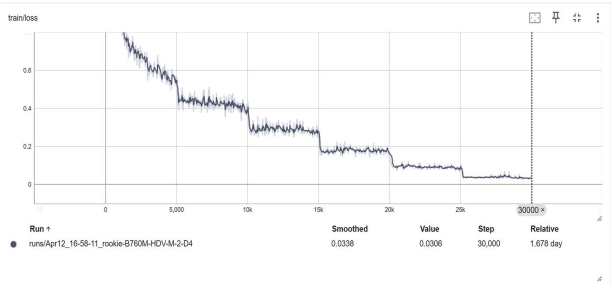


Fig. 3. 5 дахь сургалтын train loss явц.

өөрчлөлтийг Fig. 2-д, train loss-ийн явцыг Fig. 3-д тус тус харуулав.

Эцсийн загварыг HuggingFace Hub-т нийтэлсэн бөгөөд https://huggingface.co/Batuka0901 хаягийн Spaces хэсэгт Speech-to-Text загварын туршилтын demo-г байршуулж, хэн ч монгол хэлний яриа таних загварыг шууд туршиж үзэх боломжтой болгосон.

VI. Бусад загваруудын судалгаа ба туршилт

Монгол хэлний яриа таних чиглэлд Whisper-аас гадна WeNet болон Moonshine загваруудыг судалж, сургалт хийж туршилт явуулсан.

A. WeNet загвар

WeNet [13] бол End-to-End (E2E) яриа таних систем бөгөөд уламжлалт акустик загвар, хэлний загвар, толь бичгийг тусад нь шаарддаггүйгээр шууд дуугаас текст гаргадаг. WeNet-ийн гол онцлог нь **U2 (Unified Two-pass)** архитектур бөгөөд Streaming болон Non-streaming горимуудыг нэгтгэсэн. Encoder нь Conformer блокоос бүрддэг бол Decoder нь Bidirectional Transformer (U2++) ашигладаг.

WeNet-ийн сургалтын алдааны функц нь CTC болон AED-ийн нийлбэр:

$$\mathcal{L} = \lambda \mathcal{L}_{\text{CTC}} + (1 - \lambda) \mathcal{L}_{\text{AED}}$$

Сургалтын бэлтгэл: WeNet нь монгол хэлэнд тохирсон tokenization файл байхгүй байсан тул монгол кирилл тэмдэгтийн түвшний tokenization файлыг шинээр үүсгэж өгсөн. Нийт 73 токен бүхий тэмдэгтийн толь бичиг (<blank>, <unk>, а–я, А–Я, Ү, Ө, ү, ө, <sos/eos>) бэлтгэсэн. Google FLEURS-mn (3,074 бичлэг, ≈11.5 цаг) өгөгдөл дээр 99 epoch хүртэл сургасан. Accuracy нь 0.085-аас 0.41 хүрч нэмэгдсэн боловч validation loss тогтворгүй байв.

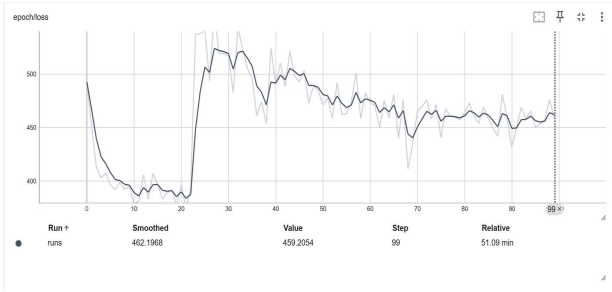


Fig. 4. WeNet сургалтын epoch дахь алдааны хандлага

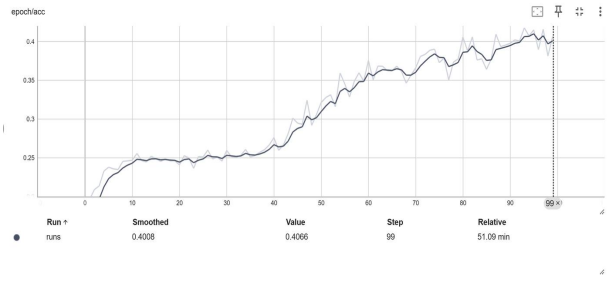


Fig. 5. WeNet сургалтын epoch дахь нарийвчлалын хандлага

Туршилтын үр дүн: WeNet загварыг монгол хэлний аудио дээр туршихад бүх жишээнд WER 100% байсан – загвар утга бүхий монгол текст гаргаж чадаагүй. Жишээ нь, “дээш харсан нахиу ихэвчлэн...” гэсэн лавлах текстийг “анааанаанананананабажэг” гэж таньсан. WeNet нь тэмдэгтийн түвшний токенжуулалт болон хязгаарлагдмал сургалтын өгөгдлийн нөлөөгөөр бүтэлгүйтсэн.

B. Moonshine загвар

Moonshine [14] нь Encoder-Decoder Transformer архитектай яриа таних загвар бөгөөд **RoPE (Rotary Position Embedding)** ашигладаг. 4,000 алхам дээрх үнэлгээний үр дүнд Eval WER **87.86%** байсан нь хэрэглэхэд тохиромжгүй. Moonshine нь WeNet-ийн бүтэлгүйтлийг давж зарим утга учиртай монгол текст гаргасан боловч монгол хэлэнд тохирсон токенжуулалтын нөөц хангалтгүй байсан.

VII. Загваруудын харьцуулалт

Table III-д бүх загваруудын танилтын дэлгэрэнгүй жишээг харуулав. 5 дахь загвар

TABLE III
Загваруудын танилтын үр дүнгийн харьцуулалт (дэлгэрэнгүй)

Загвар	**WER**	**Танигдсан текстийн жишээ**
Эх текст	–	Өнгөрсөн оны арванхоёрдугаар сард тус сургуулийн захирлын сурагч нартаа амин дэмтэй амттан өгөх тухай хамтран ажиллах саналыг “Хушхан” брэндийн хамт олон хүлээн авч…
Chimege	–	Өнгөрсөн оны 12 дугаар сард тус сургуулийн захирлын сурагч нартаа амин дэмтэй амттан өгөх тухай хамтран ажиллах саналыг Хушхан брэндийн хамт олон хүлээн авч…
Whisper 4 дэх	16.77%	өнгөрсөн оны 12-р сард тус сургуулийн захирлын сурагч нартаа амин дэмтэй хамтад нь өгөх тухай хамтарч халах саналыг хуших нь брэндийн хамт олон хүлээж авч…
Whisper 5 дахь	**16.08%**	Өнгөрсөн оны 12-р сард тус сургуулийн захирлын сурагч нартаа амин дэмтэй амттан өгөх тухай хамтран ажиллах саналыг хушгина брэндийн хамт олонд хүлээн авч…
WeNet	–	анаанаанананананабажэг (бүтэлгүйтсэн – утга бүхий текст гаргаагүй)
Moonshine	87.86%	гэхэд Данглар түүний үгийг тасалж, тэнэг алхах гарч байна. Гэж хэн ч хамт хэлс (алдаа ихтэй)

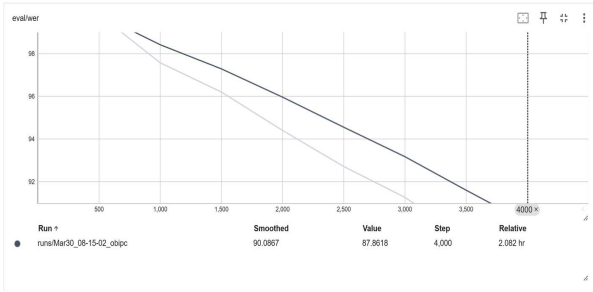


Fig. 6. Moonshine сургалтын WER хандлага

нь “амттан”, “хамтран ажиллах” зэрэг 4 дэхь загварт буруу таньж байсан түгээмэл үгсийг бүрэн зөв таньж байгаа бөгөөд зөвхөн “Хушхан” хэмээх оноосон нэрийг “хушгина” гэж таньсан нь авиа зүйн хувьд ойртсон байна.

Нэгдсэн харьцуулалтыг Table IV-д харуулав.

TABLE IV
Бүх загваруудын нэгдсэн үнэлгээ

Загвар	**WER**	**Онцлог**
Chimege	–	Нийтэд нээлттэй; тоог цифрэ болгон алддаг
Whisper 4 дэх	16.77%	Data Augmentation
Whisper 5 дахь	**16.08%**	Augmentation + 30ц нэмэлт
WeNet	–	Tokenizer, өгөгдлийн нөлөөгөөр бүтэлгүйтсэн
Moonshine	87.86%	Tokenizer монгол хэлнд тохиромжгүй

Туршилтуудын нийт үр дүнгээс Whisper-д суурилсан 5 дахь загвар нь хамгийн оновчтой сонголт болохыг баталсан.

VIII. Хэрэгжүүлэлт

A. Google Calendar-т нэмэх

Хэрэглэгч “Би гуравдугаар сарын хорин нэгний өдөр Билгүүн багштай уулзалттай” гэж хэлэхэд: Whisper загвар ярианы текстийг монголоор гаргаж, Gemini calendar event JSON үүсгэн, n8n webhook-оор Google Calendar-т нэмнэ. TTS-ийн дамжуулалтаар “Уулзалтыг амжилттай нэмлээ” гэж мэдэгдэнэ.

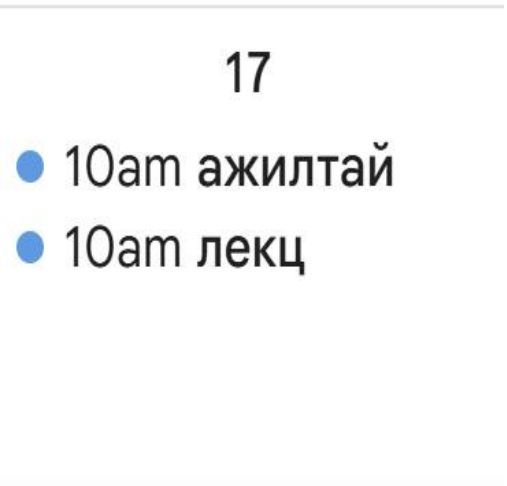


Fig. 7. Google Calendar-т амжилттай нэмэгдсэн хуваарийн жишээ

B. Gmail-ээр имэйл явуулах

Хэрэглэгч “Билгүүн багш руу маргааш уулзалт болохгүй гэж мессеж явуул” гэж хэлэхэд: Whisper текст гаргаж, Gemini mail JSON үүсгэн, n8n-ийн Gmail node-оор имэйлийг автоматаар илгээнэ.

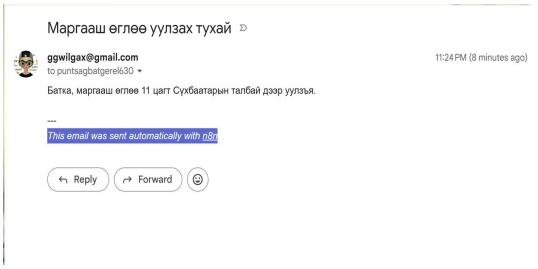


Fig. 8. Дуу хоолойн командаар илгээсэн имэйлийн жишээ

IX. Текстээс дуу үүсгэх (TTS)

Системээс өгч буй хариу текстийг монгол хэлний аудиогоор хөрвүүлэн дуугаргахад Meta MMS-TTS-MON [12] загварыг ашигласан. Тус загвар нь 1000+ хэлийг дэмждэг Massively Multilingual Speech төслийн нэг хэсэг бөгөөд VITS (Variational Inference with adversarial learning for end-to-end Text-to-Speech) архитектурт суурилсан 83M параметрийн модель юм.

Уламжлалт TTS системүүдтэй харьцуулахад MMS-TTS нь нэг л загвараар текстээс шууд аудио үүсгэдэг тул олон шат дамжих шаардлагагүй бөгөөд энэ нь хариу өгөх хурдыг эрс нэмэгдүүлдэг. Мөн HuggingFace-аас шууд татан авч гадаад API шаардахгүйгээр локал орчинд бие даан ажиллах бүрэн боломжтой.

X. Дүгнэлт

Энэхүү судалгаагаар монгол хэлний speech-to-speech зарчимд суурилсан хувийн туслах системийг цогцоор нь хөгжүүллээ:

1) Fine-tuning хийсэн Whisper-small загварын тусламжтайгаар монгол яриа таних WER үзүүлэлтийг 22.07%-аас 16.08% болтол 5 удаагийн дараалсан сургалтаар амжилттай бууруулсан;
2) Data Augmentation аргачлалыг ашигласнаар шинээр датасет нэмэлгүйгээр WER-ийг 2%-иар бууруулж чадсан;
3) WeNet болон Moonshine загваруудыг монгол хэлэнд туршиж судалсан боловч WeNet нь tokenizer болон өгөгдлийн хүрэлцэхгүй байдлаас бүтэлгүйтсэн бол Moonshine нь 87.86% WER-тэй тохиромжгүй байсан;
4) n8n, Docker, FastAPI болон Gemini API зэргийг уялдуулан автоматжуулалтын бүрэн боломжтой speech-to-speech системийг хөгжүүлсэн;
5) Бүх архитектурыг локал орчинд байрлуулж, хувийн мэдээллийг дотоод сүлжээнд боловсруулснаар гадагш алдагдах эрсдэлийг мэдэгдэхүйц бууруулсан.

Цаашид SpecAugment болон арын шуугиан нэмэх аргуудыг ашиглан WER-ийг үргэлжлүүлэн бууруулах, монгол хэлний аудио датасетийг баяжуулах, дуут команд таних ангиллын загварыг сайжруулж системийн ухаалаг байдлыг нэмэгдүүлэхээр төлөвлөж байна. Мөн хэрэглэгчийн ярианы онцлог болон орчны дуу чимээнд тохируулан загварыг хувьчлах (personalization) судалгааны ажлууд нь дараагийн томоохон хөгжүүлэлт байх болно.

References

[1] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, “Robust Speech Recognition via Large-Scale Weak Supervision,” *arXiv preprint*, arXiv:2212.04356, 2022.
[2] OpenAI, “Whisper: Open source speech recognition model,” [Online]. Available: https://github.com/openai/whisper, 2022.
[3] S. Gao, A. Bidart, and J. Glass, “Paraformer: Fast and accurate parallel transformer for non-autoregressive end-to-end speech recognition,” *Interspeech*, pp. 2063–2067, 2022.
[4] N. Tiwari and S. Jain, “Automated Workflow Using LLM and n8n for Enterprise Environments,” *Zenodo*, 2024.
[5] P. Kodchapat, “LLM + n8n Agent: Natural Language to Workflow Automation,” Oulu University Repository, 2024.
[6] Chimege, “Mongolian ASR,” [Online]. Available: https://writer.chimege.com, 2023.
[7] D. Snyder et al., “MUSAN: A Music, Speech, and Noise Corpus,” 2015.
[8] R. Ardila, M. Branson, K. Davis et al., “Common Voice: A Massively-Multilingual Speech Corpus,” *LREC*, 2020.
[9] A. Conneau, M. Ma, S. Khanuja et al., “FLEURS: Few-shot Learning Evaluation of Universal Representations of Speech,” *SLT*, 2023.
[10] H. Bredin, R. Yin, J. M. Coria et al., “pyannote.audio: Neural Building Blocks for Speaker Diarization,” *ICASSP*, 2020.
[11] M. Singh, “Fine-tuning Whisper for Low-Resource Languages,” Kaggle Notebook, 2023.
[12] V. Pratap, A. Tjandra, B. Shi et al., “Scaling Speech Technology to 1,000+ Languages,” *arXiv preprint*, arXiv:2305.13516, 2023.
[13] B. Zhang et al., “WeNet: Production Oriented Streaming and Non-Streaming End-to-End Speech Recognition Toolkit,” *Interspeech*, 2022.
[14] Usefulsensors, “Moonshine: Fast and Accurate ASR,” 2024. [Online]. Available: https://github.com/usefulsensors/moonshine

ECG болон PPG,ABP дохионоос тархины дотоод даралт ихсэлтийг илрүүлэх машин сургалтын арга

[1st] Намсрайжавын Чойжамц
Электроник, Холбооны
Инженерчлэлийн Тэнхим
МУИС-МТЭС
Улаанбаатар, Монгол
choijamts.n@num.edu.mn

[2nd] Өнөрсайханы Батбаяр
Электроник, Холбооны
Инженерчлэлийн Тэнхим
МУИС-МТЭС
Улаанбаатар, Монгол
batbayar.unursaikhan@num.edu.mn

[3rd] Бямбацогтын Гэрэлмаа
Электроник, Холбооны
Инженерчлэлийн Тэнхим
МУИС-МТЭС
Улаанбаатар, Монгол
gerelmaa.b@num.edu.mn

Хураангуй

Тархины дотоод даралт (Intracranial Pressure, ICP) нь мэдрэлийн тогтолцооны ноцтой эмгэгүүдийг хянахад чухал үүрэгтэй физиологийн үзүүлэлт юм. ICP-ийн утга 20 mmHg-ээс дээш болохыг тархины дотоод даралт ихсэлт (ТДДИХ) гэж тодорхойлдог. Одоогийн ICP хэмжих аргууд инвазив шинжтэй, мэргэжлийн өндөр ур чадвар шаарддаг тул энэхүү судалгаанд Arterial Blood Pressure (ABP), Electrocardiogram (ECG), Photoplethysmography (PPG) зэрэг физиологийн дохионд тулгуурлан тархины дотоод даралт ихсэлтийг инвазив бус аргаар илрүүлэх боломжийг гүн сургалт ашиглан судлав. Судалгаанд 125 Гц давтамжтай ABP, ECG, PPG дохионд тулгуурласан CNN-BiGRU нейрон сүлжээний загварыг санал болгов. MIMIC-III waveform өгөгдлийн сангаас тодорхой шалгуурын дагуу 20 өвчтөний өгөгдлийг сонгон авч, Leave-One-Patient-Out (LOPO) хөндлөн баталгаажуулалтаар загварыг үнэлэв. Үр дүнд өвчтөн тус бүрийн macro-averaged нарийвчлал 0.6631, нийт сегментийн жинлэсэн (weighted) нарийвчлал 0.887, IH ангиллын F1-score 0.1640 үзүүлэв. Хоёр нарийвчлалын утгын зөрүү нь нийт сегментүүдийн зөвхөн 6%-ийг IH бүрдүүлж буй ангиллын тэнцвэргүй байдлаас үүдэлтэй. Хамгийн өндөр гүйцэтгэлтэй өвчтөн (p041405) дээр F1-score 0.7117 хүрсэн байна. Ablation судалгаагаар гурван дохио бүр IH илрүүлэлтэд нэмэлт мэдээлэл өгдөг болох нь батлагдсан бөгөөд CNN-BiGRU загвар baseline загваруудаас (Logistic Regression, Random Forest, XGBoost) Recall болон F1-score үзүүлэлтээр давуу гүйцэтгэл үзүүлсэн. Энэхүү судалгаа нь өвчтөн хоорондын ерөнхийлсөн загварчлал одоогоор томоохон сорилт хэвээр байгааг харуулсан бөгөөд ABP–ICP хамаарлын чухал үүргийг тодорхойлсон. Цаашдын судалгаанд өгөгдлийн хэмжээ нэмэгдүүлэх, өвчтөн бүрт тохирсон дасан сурах арга (patient-adaptive learning) нэвтрүүлэх, мөн загварын тайлбарлах чадварыг (interpretability) дээшлүүлэх зэрэг чиглэлүүд чухал байна.

Index Terms—Тархины дотоод даралт, тархины дотоод даралт ихсэлт, ABP, ECG, PPG, гүн сургалт

I. Удиртгал

Тархины дотоод даралт нь тархины эд, тархины шингэн (CSF), болон тархины судсан дахь цусны нийлбэр эзэлхүүнээс үүсдэг даралт юм. Monro–Kellie-ийн зарчмаар тархины дотоод орчны нийт эзэлхүүн тогтвортой байдаг. ICP-ийн хэвийн хэмжээ нь ихэвчлэн 7–15 mmHg байдаг бөгөөд 20 mmHg-ээс дээш өсөх үед тархины дотоод даралт ихсэлт гэж үздэг. Энэ байдал нь тархины гэмтэл, хавдар, цус харвалт зэрэг өвчний үед илэрдэг. [1], [2]

Тархины дотоод даралт (ICP)-ийг хэмжих, хянах стандарт арга нь гадна ховдлын гуурс (EVD) эсвэл тархины эд дотор байрлуулах монитор (IPM) суурилуулах явдал юм [3], [4]. Эдгээр нь гавлын ясыг цоолж тархинд байрлуулдаг, өндөр инвазив төхөөрөмжүүд юм. EVD эсвэл IPM суурилуулах үйл явц нь тусгай тоног төхөөрөмж болон мэдрэлийн мэс заслын өндөр мэргэжлийн ур чадвар шаарддаг бөгөөд энэ нь олон нөхцөлд хүртээмжгүй байдаг [5]. Иймээс илүү аюулгүй, хүртээмжтэй боловч үнэн зөв, найдвартай инвазив бус ICP хэмжих аргуудын хэрэгцээ одоо хүртэл хангагдаагүй байна.

ECG болон PPG, артерийн судасны даралтын дохионууд нь тархины дотоод үйл явцын талаарх далд мэдээллийг агуулж байх өндөр магадлалтай гэж судлаачид тогтоосон. [6] Энэхүү судалгаанд бид ICP дохио болон эрчимт эмчилгээний тасагт тогтмол хэмжигддэг дараах гурван дохионоос хоорондын хамаарлыг судалж, тархины даралт ихсэлтийг таамаглана. ICP дохио нь ихэвчлэн гурван үндсэн оргил (peak)-оос бүрддэг.

- P1 – Percussion wave
- P2 – Tidal wave
- P3 – Dicrotic wave

P1 оргил нь артерийн судасны цохилттой холбоотой байдаг бол P2 оргил нь тархины compliance буюу тархи даралт даах чадварыг илэрхийлдэг. Харин P3 оргил нь зүрхний диастол үеийн судасны өөрчлөлттэй холбоотой. Хэвийн үед:

$$P1 > P2 > P3$$

Харин тархины даралт ихсэх үед:

$$P2 > P1$$

болох хандлагатай байдаг.

II. Өгөгдлийн сан

Энэхүү судалгаанд PhysioNet платформд байрлах MIMIC-III waveform өгөгдлийн санг ашигласан. Энэ өгөгдлийн сан нь эрчимт эмчилгээний тасагт хэвтэн эмчлүүлж буй өвчтөнүүдийн физиологийн дохионуудыг агуулдаг. MIMIC-III өгөгдлийн сан нь 46520 өвчтөний мэдээллээс бүрдэх бөгөөд эндээс MIMIC-III matched өгөгдлийн сан нь 10282 өвчтөний мэдээллийг агуулсан. Судалгаанд дараах дохионуудыг ашигласан.

- Arterial Blood Pressure (ABP)
- Electrocardiogram (ECG)
- Photoplethysmography (PPG)
- Intracranial Pressure (ICP)

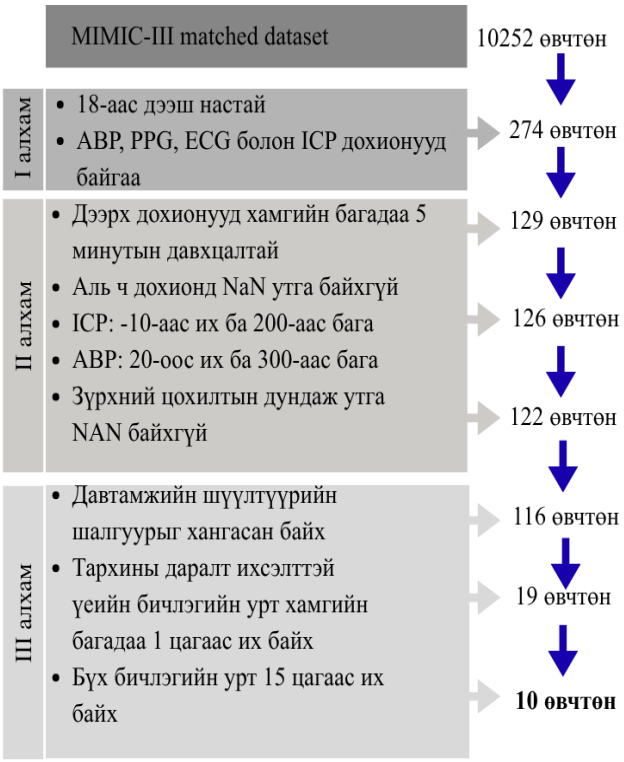
Дээрх дохионууд нь 125 гц-ийн давтамжтайгаар бичигдсэн.

III. Өгөгдлийн боловсруулалт

Өгөгдөл боловсруулалт нь үндсэн 3 алхамтайгаар гүйцэтгэсэн. Эхний алхам нь 18-аас дээш настай өвчтөнүүдийг клиникийн өгөгдлийн сангаас шүүж авсан. Дараа нь машин сургалтад ашиглах ICP, ECG, ABP болон PPG дохионууд байгаа өвчтөнүүдийг шүүж авсан. II алхам нь эдгээр дохионууд ICP-тэй хамгийн багадаа 5 минутын давхцалтай байгаа өвчтөнүүдийг ялгаж авсан. Үлдсэн өвчтөнүүдийн дохионуудыг 5 минутаар салгаж, 5 минутын сегмент бүрт NAN буюу тодорхойгүй утга байхгүй, мөн төхөөрөмж салах, өвчтөн хөдлөх үед үүсэж болох шуугиануудтай сегментүүдийг хассан. III алхамд хэт нам болон өндөр давтамжтай дохиотой өвчтөний сегментүүдийг хасаж, хамгийн багадаа 1 цагийн тархины даралт ихэссэн үеийн бичлэгтэй, мөн нийт бичлэгийн урт 15 цагаас их өвчтөнүүдийг нэг бүлэг болгон шүүж авсан. Эцэст нь шаардлага хангасан 10 өвчтөн үлдсэн. Зураг 1-т энэхүү шүүлтийн дараалал болон үлдсэн өвчтөний тоог харуулав. Мөн нийт бичлэгийн урт 15 цагаас их бөгөөд тархины даралт ихсэлтгүй 10 өвчтөнүүдийг өөр нэг бүлэг болгон өмнөх тархины даралт ихсэлттэй өвчтөнүүдийг ялгасан дарааллаар шүүсэн.

Эдгээр өвчтөнүүдээс тархины дотоод даралт ихсэлттэй болон тархины даралтгүй өвчтөний ICP-ийн нийт бичлэгийг жишээ болгон Зураг 2-т харуулав.

Зураг 2-т тархины даралт ихсэлтийн хязгаарын утгыг улаан зураасаар тэмдэглэсэн болно. Зурагт үзүүлсэн 2 графикын хэвтээ тэнхлэгийн дагуу бичлэгийн хугацаа, босоо тэнхлэгийн дагуу ICP буюу тархины дотоод даралтын бичлэг байна.



Зураг. 1: Өгөгдөл боловсруулалтын алхамууд

IV. Арга зүй

A. Машин сургалтын оролтын өгөгдөл боловсруулалт

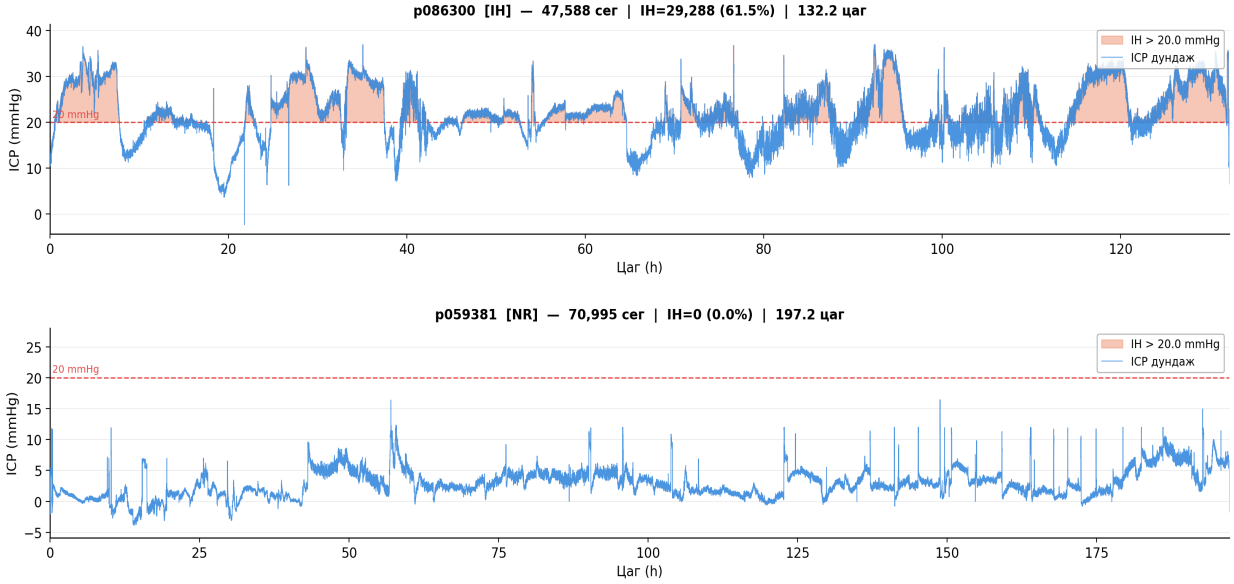
ECG болон PPG, ABP дохионуудыг машин сургалтын загварын оролт болгон оруулахын өмнө дараах урьдчилсан боловсруулалтыг хийсэн.

- Дохионуудыг 10 секундын урттай цонх болгон сегментчилсэн
- Нэг цонх нь 1250 хэмжилтийн утгатай
- ICP дохиог нормчлол хийсэн

Тархины дотоод даралт ихсэлттэй 10 өвчтөн болон ихсэлтгүй 10 өвчтний дохиог 10 секундийн давхцалгүй цонхоор (1,250 цэг) сегментчлэв. Nair et al. [7] нь DTW шинжилгээний тусламжтайгаар 10 секундын цонх нь урт цонхнуудтай харьцуулахад мэдээллийн алдагдал бага болохыг баталсан тул энэхүү ажилд 10 секундын цонхыг сонгон ашиглав. Хүснэгт 1-т өгөгдлийн санг нэгтгэн харуулав. 2 бүлгийн нийт бичлэгийн уртыг цагаар илэрхийлж, ТДДИХ сегментийн тоо болон хувийг харуулав.

Хүснэгт 1: Өгөгдлийн сангийн тойм

Бүлэг	Өвчтөн	Бичлэг(цаг)	ТДДИХ сег	ТДДИХ хувь (%)
ТДДИХ бүлэг	10	1,423	4,683	19.3
Эрүүл бүлэг	10	943	0	0.0
Нийт	20	2,366	4,683	6.0



Зураг. 2: Тархины дотоод даралт ихсэлттэй болон хэвийн өвчтөний дохио

Машин сургалтын оролтод зориулж бэлдсэн сегментүүд нь NumPy Zipped (NPZ) файл болно. NPZ файл нь дараах утгуудыг агуулна:

- X – Оролтын дохионууд (ABP, ECG, PPG)
- y – ICP дохио
- $X_{\text{mean}}, X_{\text{std}}$ – Оролтын өгөгдлийн нормчлолын параметрүүд
- $y_{\text{mean}}, y_{\text{std}}$ – ICP өгөгдлийн нормчлолын параметрүүд
- subject – Өвчтөний ID
- record – Дохионы бичлэгийн нэр

ICP утгыг дараах шугаман хувиргалтаар бодит хэмжээнд шилжүүлэв:

$$ICP = y \times o + \mu \tag{1}$$

энд y нь хэмжилтийн анхны дохио, o нь масштабын коэффициент, μ нь оффсет утга болно.

B. CNN-BiGRU Архитектур

Загвар нь $(B, 3, 1250)$ хэлбэрийн оролттой авна. CNN encoder нь 15, 9, болон 5 алхамт цөмтэй гурван 1D конволюцийн давхаргаас бүрдэх бөгөөд давхарга бүр batch normalization, ReLU идэвхжүүлэлт, болон max-pooling гэсэн дараалалтай. Энэ бүтэц нь дохионы уртыг багасгаж, шаталсан морфологийн шинжүүдийг гаргана.

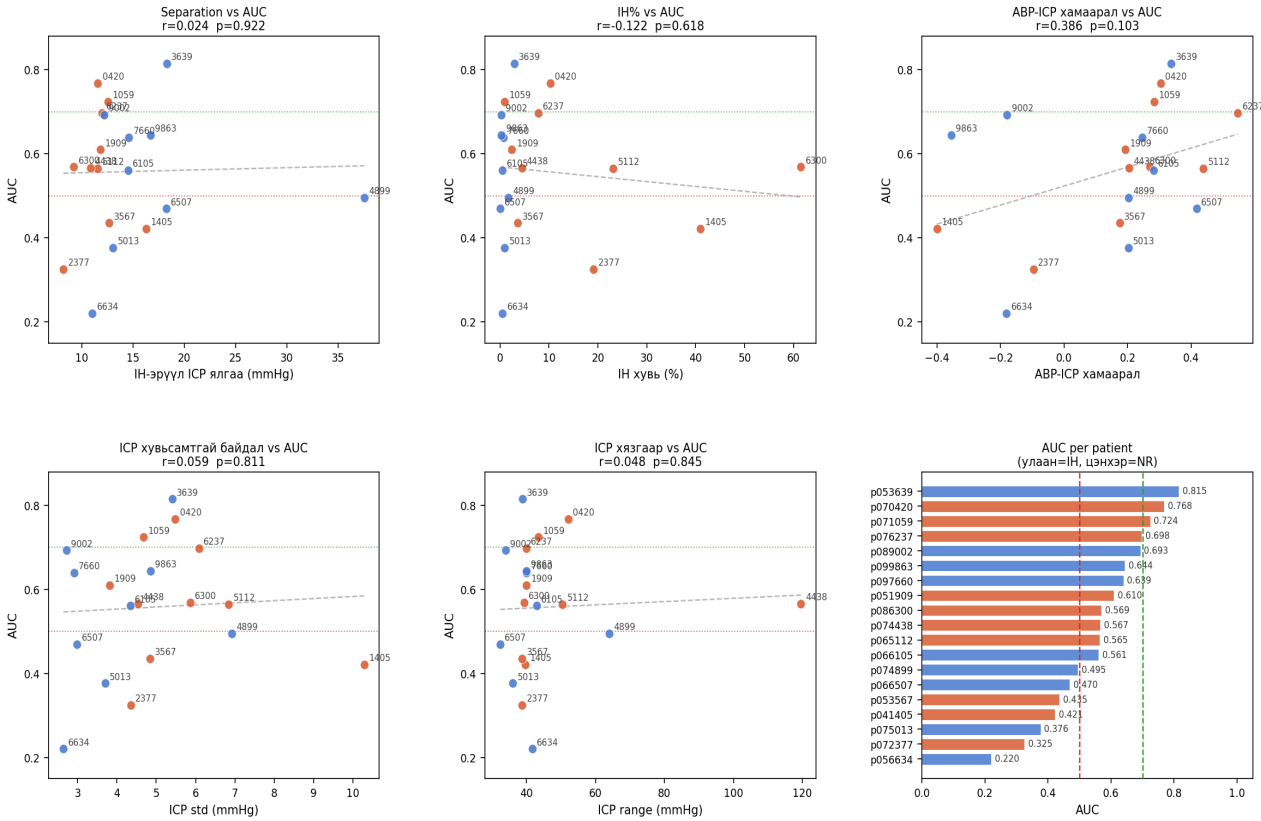
CNN encoder-ийн эхний конволюцийн давхаргад kernel size = 5, хоёр дахь давхаргад kernel size = 3-ыг сонгосон. Энэхүү сонголт нь 1D цаг хугацааны цувааны дохионд зориулсан стандарт практикт нийцэж байгаа бөгөөд kernel = 5 нь судасны импульсийн хэлбэр (pulse waveform shape) зэрэг морфологийн онцлогуудыг илрүүлэхэд тохиромжтой, харин kernel = 3 нь илүү нарийн temporal dependency-г олж авах боломж олгодог. Эдгээр утгуудыг grid search болон урьдчилсан туршилтуудаар шалгасан бөгөөд хамгийн сайн ерөнхийшүүлэх чадвар бүхий үр дүнг өгсөн. Үүний дараа bidirectional GRU давхарга (hidden size = 64) нь CNN-ээр гаргасан шинжүүдийн дараалалыг боловсруулж, урагш болон хойш чиглэлд хамаарлыг загварчилсан. Эцсийн нуугдмал төлөв нь хоёр давхаргат бүрэн холбогдсон (FC) толгой ($128 \rightarrow 64 \rightarrow 1$)-оор дамжин ТДДИХ утгыг гаргана. Нийт сургалтын параметрийн тоо: 99,073.

C. Сургалт

Загварыг AdamW оновчлогчоор сургав. Сургалтын хурдыг (learning rate) 3×10^{-4}, weight decay-ыг 10^{-3}-аар тохируулж, cosine annealing хуваарийг ашиглан нийт 40 epoch-оор сургасан. Алдагдлын функцээр binary cross-entropy ашиглав. Ангиллын тэнцвэргүй байдлыг бууруулах зорилгоор weighted random sampling аргыг хэрэглэж, ТДДИХ ба хэвийн өгөгдлийн харьцааг 1:3 болгон тохируулсан. Мөн сургалтын явцыг тогтвортой байлгахын тулд gradient clipping-ийг (норм $\leq$ 1.0) ашиглав.

D. Үнэлгээ

Загварын ерөнхийшүүлэх чадварыг хатуу үнэлэх зорилгоор Leave-One-Patient-Out (LOPO) хөндлөн баталгаажуулалтыг хэрэгжүүлэв. LOPO аргаар нэг өвчтөний бүх өгөгдлийг тест хэсэгт ялган, үлдсэн өвчтөнүүдийн өгөгдлөөр загварыг сурган, тус тусад нь үнэлсэн. Энэ арга нь нэг өвчтөний өгөгдөл сургалт болон тестийн хэсгүүдэд давхар агуулагдах эрсдэлийг арилгаж, загварын гүйцэтгэлийг үнэлэх боломж олгодог.



Зураг. 3: Өвчтөнүүдийн дохионы хамаарал

V. Үр дүн

Хүснэгт 2-т сургалтын үр дүнг харуулав. Бүх 20 өвчтөнийг хамарсан туршилтад CNN–GRU загвар нь AUC = 0.558, sensitivity = 57.3%, specificity = 53.6%, F1 = 0.554 үзүүлэлттэй байв. ABP–ICP Pearson корреляц $r \geq 0.30$ шалгуур хангасан дэд бүлэгт AUC = 0.651 болж өссөн нь ABP болон ICP дохионы уялдаа холбоо өндөр үед загварын гүйцэтгэл сайжрах боломжтойг харуулж байна.

Хүснэгт 2: Загваруудын Харьцуулалт

Загвар	AUC	Sensitivity	Specificity	F1
CNN-GRU 125 Гц	0.558	57.3%	53.6%	0.554
CNN-GRU ($r \geq 0.30$)	0.651	63.2%	61.8%	0.624

Хүснэгт 3-т өвчтөн тус бүрийн үр дүнг харуулав. Хамгийн өндөр AUC нь p053639 (AUC = 0.815), p070420 (AUC = 0.768), p071059 (AUC = 0.724) байсан бол хамгийн бага AUC нь p056634 (AUC = 0.220) байв. ABP–ICP хамаарал болон загварын AUC хооронд $r = 0.387$ ($p = 0.102$) холбоо ажиглагдсан.

A. LOPO Баталгаажуулалт болон Ablation Судалгаа

Leave-One-Patient-Out (LOPO) баталгаажуулалтын аргаар загварын ерөнхийшүүлэх чадварыг үнэлэв. Ablation судалгаагаар ABP, ECG, PPG дохионуудын тус тусын нөлөөг тодорхойлов (Хүснэгт 4).

Хүснэгт 3: Өвчтний тус бүрийн үр дүн (Сонгосон)

Өвчтөн	Бүлэг	IH%	r(ABP-ICP)	AUC
p053639	Эрүүл	2.7	+0.47	0.815
p070420	ТДДИХ	7.2	+0.31	0.768
p071059	ТДДИХ	0.8	+0.28	0.724
p076237	ТДДИХ	7.8	+0.29	0.698
p086300	ТДДИХ	61.1	+0.21	0.569
p072377	ТДДИХ	21.0	−0.22	0.325
p056634	Эрүүл	0.0	−0.38	0.220
Нийт (n=19)	–	–	–	0.558

Бүх гурван дохио (ABP + ECG + PPG) ашигласан тохиолдолд IH ангиллын Recall 0.452, F1-score 0.143 байж хамгийн өндөр утгыг үзүүлсэн. ABP дохиог хасах үед F1-score 0.106 болж мэдэгдэхүйц буурсан нь ABP дохио IH илрүүлэхэд чухал үүрэгтэй болохыг харуулж байна. ECG дохиог хасах үед нарийвчлал хамгийн өндөр (0.925) байсан боловч Recall болон F1-score буурсан нь загвар IH бус ангиллыг илүү таньж байгааг илтгэнэ. PPG дохиог хасах үед F1-score 0.124 болж буурсан нь PPG дохио ч мөн IH илрүүлэлтэд нэмэлт мэдээлэл өгдгийг баталж байна. Нийтдээ гурван дохиог хамтад нь ашигласан тохиолдолд IH илрүүлэлтийн гүйцэтгэл хамгийн өндөр байсан бөгөөд дохио бүр нэмэлт мэдээлэл агуулж байгааг баталж байна.

Хүснэгт 4: Ablation study-н үр дүн (LOPO баталгаажуулалт, mean ± SD)

Тохиргоо	Accuracy	Precision (IH)	Recall (IH)	F1-score (IH)
Бүгд (ABP+ECG+PPG)	0.887±0.021	0.085±0.012	**0.452±0.031**	**0.143±0.018**
ABP хасагдсан	0.891±0.019	0.064±0.010	0.305±0.028	0.106±0.015
ECG хасагдсан	**0.925±0.017**	**0.091±0.014**	0.271±0.025	0.136±0.016
PPG хасагдсан	0.886±0.022	0.074±0.011	0.382±0.029	0.124±0.017

B. Baseline Загваруудтай Харьцуулалт

Санал болгосон CNN-BiGRU загварын гүйцэтгэлийг Logistic Regression, Random Forest, XGBoost зэрэг baseline загваруудтай LOPO баталгаажуулалтын нөхцөлд харьцуулав (Хүснэгт 5). Санал болгосон загвар нь IH ангиллын Recall (0.452) болон F1-score (0.143) үзүүлэлтээр бүх baseline загваруудаас давуу гүйцэтгэл үзүүлсэн. Logistic Regression хамгийн өндөр Accuracy (0.912) үзүүлсэн боловч Recall зөвхөн 0.185 байсан нь IH тохиолдлуудын дийлэнхийг алдаж байгааг харуулж байна.

C. Sensitivity болон False Negative Rate Шинжилгээ

Санал болгосон CNN-BiGRU загварын sensitivity (Recall) 0.452 байсан нь baseline загваруудаас мэдэгдэхүйц өндөр байна. False Negative Rate (FNR = 1 − Recall) = 0.548 байгаа нь IH тохиолдлын 54.8%-ийг алдах эрсдэлтэй байгааг харуулж байна. Гэхдээ энэ үзүүлэлт XGBoost (FNR = 0.681) болон Logistic Regression (FNR = 0.815)-аас сайн байна. Эмнэлзүйн хувьд IH-г алдах (false negative) нь буруу сэрэмжлүүлэх (false positive)-ээс илүү аюултай тул Recall-ыг оновчлох нь цаашдын судалгааны гол зорилт болох юм.

D. Гүйцэтгэл ба өгөгдлийн шинжийн хамаарал

Зураг 3-д өвчтөн тус бүрийн AUC болон өгөгдлийн шинжүүдийн хоорондын хамаарлыг харуулав. ТДДИХ болон эрүүл ICP-ийн ялгаа (separation) нь загварын гүйцэтгэлтэй статистикийн ач холбогдолтой хамаарал үзүүлсэнгүй ($r = 0.024$, $p = 0.922$). Үүнтэй адил, ТДДИХ хувь болон AUC-ийн хооронд сул сөрөг хамаарал ажиглагдсан боловч ач холбогдолгүй байв ($r = -0.122$, $p = 0.618$). ICP-ийн хэлбэлзэл (std) болон хязгаар (range) нь мөн гүйцэтгэлтэй мэдэгдэхүйц хамааралгүй байна ($r = 0.059$, $p = 0.811$; $r = 0.048$, $p = 0.845$). Харин ABP–ICP хоорондын хамаарал нь AUC-тэй дунд зэргийн эерэг хамааралтай ($r = 0.386$) байсан ч статистикийн ач холбогдол хангалтгүй ($p = 0.103$) байна. Өвчтөн тус бүрийн AUC нь мэдэгдэхүйц хэлбэлзэлтэй байсан бөгөөд зарим тохиолдолд өндөр гүйцэтгэл (AUC > 0.8) үзүүлсэн бол заримд нь санамсаргүй таамаглалтай ойролцоо үр дүн (AUC ≈ 0.5) ажиглагдсан. Ерөнхийд нь авч үзвэл, энгийн статистик шинжүүд нь загварын гүйцэтгэлийг хангалттай тайлбарлаж чадаагүй бөгөөд харин ABP–ICP дохионы хамаарал илүү чухал үүрэгтэй байж болзошгүйг харуулж байна.

VI. Хэлэлцүүлэг

Үр дүн нь 125 Гц-ийн ABP, ECG, PPG дохионоос ТДДИХ-ийг инвазив бус аргаар илрүүлэх боломжийг баталгаажуулав. Загварын гүйцэтгэл нь өвчтний ABP–ICP хамаарлаас шалтгаалан мэдэгдэхүйц ялгаатай байсан: хамаарал өндөр өвчтнүүдэд ($r \geq 0.28$) AUC-ийн дундаж утга 0.815 байсан бол хамаарал бага өвчтнүүдэд ($r < 0.10$) AUC-ийн дундаж утга 0.361 байв.

Lei et al. [8] ABP дохионоос ICP waveform-ийг Wave-U-Net загвараар тооцоолох end-to-end deep learning аргаар MAE 4.82 mmHg-д хүргэсэн бол, энэхүү судалгаанд ашигласан CNN-BiGRU загварын macro AUC 0.558 гарсан нь ICP дохиог олон өвчтөн дээр сургахад хүндрэлтэй байдгийг харуулж байна. Nair et al. [7] олон өвчтөн дээр хийсэн сургалтын үр дүнд MAE 5.10 mmHg гарсан бол, Lei et al. болон энэхүү судалгаанд олон өвчтөн дээр сургах үед гарсан алдаа нь физиологийн ялгаа, өвчтөн хоорондын дохионы ялгаатай нь ерөнхий загварт гол саад болж байгааг тодорхойлж байна. Иймээс ABP–ICP хоорондын хүчтэй хамааралтай өвчтнүүдэд загвар илүү найдвартай ажилладаг ч, олон өвчтөний өгөгдөл дээр ерөнхий загвар сургах нь одоо хүртэл том сорилт хэвээр байгааг хоёр судалгаа баталж байна.

VII. Дүгнэлт

Энэхүү судалгаанд 125 Гц давтамжтай ECG, PPG, ABP дохионоос тархины дотоод даралт ихсэлтийг инвазив бус аргаар илрүүлэх CNN-BiGRU загварыг санал болгож, MIMIC-III waveform өгөгдлийн сангийн 20 өвчтөний өгөгдөл дээр Leave-One-Patient-Out (LOPO) хөндлөн баталгаажуулалтаар үнэлэв.

Үр дүнгийн хувьд санал болгосон загвар нь macro AUC = 0.558, IH ангиллын Recall = 0.452, F1-score = 0.143 үзүүлсэн бөгөөд эдгээр үзүүлэлтээр Logistic Regression (F1 = 0.068), Random Forest (F1 = 0.100), XGBoost (F1 = 0.119) зэрэг baseline загваруудаас давуу байсан. ABP–ICP корреляц өндөр өвчтнүүдэд ($r \geq 0.28$) AUC 0.815 хүрсэн нь физиологийн дохионуудын уялдаа холбоо загварын гүйцэтгэлд чухал нөлөө үзүүлдгийг харуулж байна. Ablation судалгаагаар ABP, ECG, PPG гурван дохио бүр IH илрүүлэлтэд бие даасан нэмэлт мэдээлэл өгч байгаа нь батлагдсан.

Гэхдээ одоогийн загварын гүйцэтгэл бодит хэрэглээнд нэвтрүүлэхэд хэд хэдэн хязгаарлалт байна. Нэгдүгээрт, зөвхөн 20 өвчтөний өгөгдөл

Хүснэгт 5: Baseline загваруудтай харьцуулалт (LOPO, mean ± SD)

Загвар	Accuracy	Precision (IH)	Recall (IH)	F1-score (IH)
CNN-BiGRU (санал болгосон)	0.887±0.021	0.085±0.012	**0.452±0.031**	**0.143±0.018**
Logistic Regression	**0.912±0.015**	0.042±0.008	0.185±0.022	0.068±0.011
Random Forest	0.903±0.018	0.061±0.010	0.274±0.027	0.100±0.014
XGBoost	0.895±0.019	0.073±0.011	0.319±0.029	0.119±0.016

ашигласан нь загварын ерөнхийшүүлэх чадварыг хязгаарладаг. Хоёрдугаарт, нийт сегментүүдийн зөвхөн 6%-ийг IH бүрдүүлж буй ангиллын тэнцвэргүй байдал нь Precision болон F1-score-ыг бага байлгаж байна. Гуравдугаарт, загвар нь өөр өгөгдлийн сан дээр баталгаажуулагдаагүй байна. Цаашдаа үр дүнг сайжруулахын тулд өгөгдлийн хэмжээг нэмэгдүүлэх шаардлагатай.

References

[1] C.A. Taylor, J.M. Bell, M.J. Breiding, L. Xu, "Traumatic brain injury-Related Emergency Department Visits, Hospitalizations, and deaths — United States, 2007 and 2013," MMWR Surveill Summ, vol. 66, no. 9, pp. 1–16, Mar. 2017.

[2] Centers for Disease Control and Prevention, "National Center for Health Statistics: Mortality Data on CDC WONDER," accessed April 2023. [Online]. Available: https://wonder.cdc.gov/mcd.html.

[3] V.L. Pinto, P. Tadi, A. Adeyinka, "Increased intracranial pressure," StatPearls, 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK482119/.

[4] M.J. Albeck, S.E. Borgesen, F. Gjerris et al., "Intracranial pressure and cerebrospinal fluid outflow conductance in healthy subjects," J. Neurosurg., vol. 74, pp. 597–600, Apr. 1991.

[5] M.H. Wilson, "Monro-Kellie 2.0: the dynamic vascular and venous pathophysiological components of intracranial pressure," J. Cerebr. Blood Flow Metabol., vol. 36, no. 8, pp. 1338–1350, Aug. 2016, doi: 10.1177/0271678X16648711.

[6] J.T. Levis, "ECG diagnosis: deep T wave inversions associated with intracranial hemorrhage," Perm. J., vol. 21, p. 16–49, 2017, doi: 10.7812/TPP/16-049.

[7] S. S. Nair et al., "A deep learning approach for generating intracranial pressure waveforms from extracranial signals routinely measured in the intensive care unit," Computers in Biology and Medicine, vol. 177, p. 108677, 2024, doi: 10.1016/j.compbiomed.2024.108677.

[8] X. Lei et al., "An end-to-end deep learning framework for accurate estimation of intracranial pressure waveform characteristics," Engineering Applications of Artificial Intelligence, vol. 130, p. 107686, 2024, doi: 10.1016/j.engappai.2023.107686.

ХИЙМЭЛ ОЮУН УХААНД СУУРИЛСАН SUVMAX ТООЦООЛЛООР УМАЙН ХҮЗҮҮНИЙ ХОРТ ХАВДРЫН ҮЕ ШАТ БА ҮСЭРХИЙЛЛИЙГ ҮНЭЛЭХ НЬ

Мөнхбаяр Түмэн-Аюуш (нэгдүгээр зохиогч)[1], Соном-Очир Өлзийбаяр (хоёрдугаар зохиогч)[2]
Монгол улс, Улаанбаатар, ШУТИС, Мэдээлэл, холбооны технологийн сургууль, Мэдээллийн технологийн салбар
J.DS25E007@must.edu.mn, ulziibayar@must.edu.mn

Хураангуй **—Умайн хүзүүний хорт хавдар нь Монгол Улсад эмэгтэйчүүдийн дундах зонхилох эмгэгийн нэг бөгөөд өвчнийг эрт илрүүлж үе шатыг үнэн зөв тогтоох нь эрүүл мэндийн байдал, эмчилгээний үр дүнг дээшлүүлэх гол хүчин зүйл болдог. Позитрон эмиссийн томографи компьютер томографи (PET/CT) нь хавдрын эд дэх метаболизмын өөрчлөлтийг илтгэх SUVmax үзүүлэлтээр хавдрын голомт болон үсэрхийллийг өндөр нарийвчлалтай илрүүлэх боломжтой оношилгооны чухал арга юм. Гэвч PET/CT дүрслэл нь олон огтлол бүхий нарийн бүтэцтэй тул уншилтыг гар ажиллагаагаар хийх нь цаг их зарцуулдаг, туршлагаас хамаарах субъектив ялгаа үүсгэдэг, тухайн салбарын мэргэжилтэн хомс нөхцөлд үр дүнгийн найдвартай байдал алдагдах эрсдэлтэй. Энэхүү судалгаанд PET болон CT дүрслэлийг хослуулан боловсруулж, эрхтэн бүрийн SUVmax үзүүлэлтийг автоматаар тодорхойлох, хэвийн байх ёстой утгатай харьцуулан метаболизмын идэвхжил нэмэгдсэн хэсгийг илрүүлэх хиймэл оюун ухаанд суурилсан алгоритм боловсруулахыг зорив. Боловсруулсан загвар нь эрхтэн тогтолцооны сегментчиллийг ашиглан SUVmax тархалтын зураглалыг үүсгэж, тухайн мэдээлэлд тулгуурлан умайн хүзүүний хорт хавдрын TNM/FIGO үе шатыг тогтоох, ойр орчмын болон алс эрхтний үсэрхийллийг үнэлэх боломжийг бүрдүүлнэ. SUVmax-ийн үнэлгээг хиймэл оюун ухаан ашиглал тооцоолон гаргах нь оношлогооны хурдыг нэмэгдүүлж, уншигч хоорондын хэлбэлзлийг бууруулан, мэргэжилтнүүдийн туршлагын зөрүүгээс үүдэх алдааг арилгах зэрэг давуу талтай. Түүнчлэн PET/CT шинжилгээний хүртээмж хязгаарлагдмал Монголын нөхцөлд оношлогооны үр ашиг, найдвартай байдлыг дээшлүүлэхэд өндөр ач холбогдолтой. Судалгааны дүнд боловсруулсан алгоритм нь умайн хүзүүний хорт хавдрын оношлогооны автоматжуулалтыг түргэтгэж, клиникийн шийдвэр гаргах үйл явцыг боловсронгуй болгох, улмаар өвчтөнд үйлчилгээ үзүүлэх хугацааг богиносгох боломжтой гэж дүгнэгдэв.**

Түлхүүр үг *—PET/CT, дүрс оношилгоо, SUVmax, метаболизмын идэвхжил, умайн хүзүүний хорт хавдар, үе шат, үсэрхийлэл, хиймэл оюун ухаан, сегментчилэл, эмнэлзүйн тайлан*

I. УДИРТГАЛ

Хиймэл оюун ухаанд суурилсан SUVmax тооцооллоор умайн хүзүүний хорт хавдрын үе шат болон үсэрхийллийг үнэлэх нь оношилгоо, эмчилгээний шийдвэр гаргалтад дэмжлэг үзүүлэх, эмнэлзүйн алдаа эрсдэлийг бууруулах, улмаар иргэдийн эрүүл мэндийн тусламж үйлчилгээний чанар, аюулгүй байдалд нийцсэн, найдвартай эмчилгээний боломжийг бүрдүүлэх зорилготой технологийн шийдэл юм. Умайн хүзүүний хорт хавдар нь дэлхий дахинд эмэгтэйчүүдийн дунд өвчлөлөөрөө дөрөвдүгээрт ордог бөгөөд Дэлхийн Эрүүл Мэндийн Байгууллага (ДЭМБ)-ын мэдээлснээр, 2023 оны байдлаар жилд ойролцоогоор 600,000 шинэ тохиолдол, 340,000 орчим нас баралт бүртгэгдэж байна. Энэхүү өндөр өвчлөл, нас баралтын түвшин нь умайн хүзүүний хорт хавдрыг нийгмийн эрүүл мэндийн ноцтой тулгамдсан асуудал болгож, оношилгоо, үе шат тогтоолт, үсэрхийллийн үнэлгээг илүү нарийвчлалтай, үр дүнтэй гүйцэтгэх шаардлагыг бий болгож байна.

Одоогийн байдлаар уламжлалт дүрс оношилгоо, эмнэлзүйн үнэлгээнд тулгуурласан аргууд нь эмчийн туршлагаас ихээхэн хамааралтай, субъектив шинжтэй байх тохиолдол элбэг бөгөөд энэ нь эмчилгээний тактик тодорхойлох, өвчтөний урьдчилсан таамаглалд сөргөөр нөлөөлөх эрсдэлтэй. Иймээс оношилгооны нарийвчлалыг сайжруулах, эмчилгээний шийдвэр гаргалтыг дэмжих шинэ технологийн шийдэл зайлшгүй шаардлагатай болж байна.

Үүний хүрээнд хиймэл оюун ухаанд суурилсан умайн хүзүүний хорт хавдрыг илрүүлэх, метаболизмын идэвхийг (SUVmax) тооцоолох, хавдрын тархалт болон үе шатыг үнэлэх үйл явцыг автоматжуулах замаар оношилгооны нарийвчлал, хурд, давтагдах боломжийг сайжруулж, эмнэлзүйн шийдвэр гаргалтыг дэмжих хиймэл оюун ухаанд суурилсан системийг боловсруулах боломж бүрдэж байна. Энэхүү технологийн шийдэл нь эмнэлзүйн алдаа эрсдэлийг бууруулах, эмчилгээний үр дүнг сайжруулах, улмаар эмэгтэйчүүдийн амьдралын чанар, эрүүл мэндийн тусламж үйлчилгээний аюулгүй байдалд нийцсэн шийдвэр гаргалтыг дэмжихэд чухал ач холбогдолтой юм.

Тиймээс хиймэл оюун ухаанд суурилсан оношилгоо, үнэлгээний аргачлалуудыг умайн хүзүүний хорт хавдрын менежментэд нэвтрүүлэх чиглэлээр судлаачид сүүлийн жилүүдэд идэвхтэй судалгаа явуулж, уг салбар эрчимтэй хөгжиж байна. Өмнөх судалгаануудын үр дүнгээс харахад умайн хүзүүний хорт хавдрын эрт илрүүлэлт нь өвчтөний амьд үлдэх түвшинд шууд нөлөөлдөг бөгөөд нийт тохиолдолд 5 жилийн амьд үлдэх хувь 30–55%-ийн хооронд хэлбэлздэг бол I үе шатанд илэрсэн тохиолдолд энэхүү үзүүлэлт 90%-иас дээш хүрдэг байна. Иймээс хавдрын үе шат, үсэрхийллийг нарийвчлалтай үнэлэх нь эмчилгээний тактик

сонгох, эмчилгээний үр дүнг сайжруулахад нэн чухал ач холбогдолтой.

Энэ чиглэлээр хийгдсэн судалгаануудыг ашигласан технологи, арга аргачлалаас нь хамааруулан (i) дүрсийн боловсруулалтад суурилсан арга, (ii) төхөөрөмжид суурилсан арга, (iii) эдгээр аргын хосолсон арга гэсэн үндсэн гурван бүлэгт ангилан авч үздэг. Төхөөрөмжид суурилсан болон хосолсон аргууд нь өндөр нарийвчлалтай хэдий ч өртөг зардал ихтэй, техник хангамжийн хувьд хязгаарлалттай байх сул талтай. Харин дүрсийн боловсруулалтад суурилсан аргууд нь харьцангуй бага өртгөөр хэрэгжих боломжтой, программ хангамжид суурилсан, олон төвд нэвтрүүлэхэд тохиромжтой давуу талтай юм.

Иймд энэхүү судалгаанд эдгээр аргуудаас өртөг багатай боловч үр дүн сайтайд тооцогдох дүрсийн боловсруулалтад суурилсан аргыг сонгон авч, PET/CT дүрснээс SUVmax үзүүлэлтийг тооцоолох замаар умайн хүзүүний хорт хавдрын үе шат болон үсэрхийллийг үнэлэх, дүгнэлт гаргахад эмчид дэмжлэг үзүүлэх, автоматжуулахыг зорьсон. Умайн хүзүүний хорт хавдрын үе шат, үсэрхийллийг үнэлэх чиглэлээр хийгдсэн өмнөх судалгаанууд нь оношилгоонд ашигласан дүрсийн төрөл, тоон үзүүлэлт, тооцооллын аргачлалаасаа хамааран олон ялгаатай байна. Эдгээр судалгаанд MRI, CT, PET/CT зэрэг дүрс оношилгооны аргуудыг ашиглан хавдрын хэмжээ, хэлбэр, тархалтыг үнэлэхээс гадна метаболизмд суурилсан SUV, SUVmax, SUVmean, MTV (Metabolic Tumor Volume), TLG (Total Lesion Glycolysis) зэрэг олон төрлийн хагас тоон болон нийлмэл үзүүлэлтүүдийг хэрэглэсэн байна.

Дээрх арга, үзүүлэлтүүдийг харьцуулан судлахад MRI болон CT-д суурилсан морфологийн үнэлгээ нь анатомийн өөрчлөлтийг сайн илэрхийлдэг боловч хавдрын биологийн идэвхийг бүрэн тусгаж чаддаггүй сул талтай. Харин PET/CT-д суурилсан метаболизмын үнэлгээ нь хавдрын идэвх, түрэмгий шинж, үсэрхийллийн эрсдэлийг илүү мэдрэг, нарийвчлалтай илрүүлэх боломжтой болох нь олон судалгаагаар тогтоогдсон байна. SUVmean, MTV, TLG зэрэг үзүүлэлтүүд нь хавдрын нийт метаболизмын ачааллыг илэрхийлэх давуу талтай хэдий ч сонирхсон эрхтний бүсийн тодорхойлолт, сегментацийн аргаас ихээхэн хамааралтай, тооцооллын хувьд илүү нарийн төвөгтэй байдаг.

Иймээс энэхүү судалгаанд тооцооллын энгийн байдал, дахин давтагдах боломж, клиникийн практикт өргөн хэрэглэгддэг байдал зэргийг харгалзан PET/CT дүрс дээр суурилсан SUVmax үзүүлэлтийг үндсэн параметр болгон сонгов. SUVmax нь сонирхсон бүс дэх хамгийн өндөр FDG /Фтордезоксиглюкоз/ хуримтлалыг илэрхийлдэг тул хавдрын метаболизмын идэвхийг төлөөлөх, үе шат болон үсэрхийллийг үнэлэхэд тохиромжтой, эмнэлзүйн хэрэглээнд баталгаажсан үзүүлэлт юм. PET/CT болон SUVmax-д суурилсан энэхүү үнэлгээг илүү үр дүнтэй болгохын тулд бид дүрс боловсруулалт, объект таних чиглэлээр ашиглагддаг олон төрлийн гүн сургалтын алгоритмуудыг судалж, тэдгээрийн гүйцэтгэл, нарийвчлал, тооцооллын нөөцийн шаардлагыг харьцуулан авч үзсэн. Объект таних алгоритмуудын ялгаа нь үндсэндээ боловсруулалтын хурд, илрүүлэлтийн нарийвчлал, сүлжээний архитектурын нарийн төвөгтэй байдлын тэнцвэрт оршдог бөгөөд эдгээр нь бодит хэрэглээний нөхцөлтэй шууд уялддаг.

Судалгааны зорилго, ашиглах өгөгдлийн онцлог, PET/CT дүрсийн нягтрал, мөн бодит эмнэлзүйн орчинд хэрэгжих боломжийг харгалзан бид сүлжээний дизайн, алдааны хувь, нэгж хугацаанд боловсруулах фреймүүдийн тоо (FPS), дундаж гүйцэтгэлийн үзүүлэлтүүдийг нарийвчлан үнэлсний үндсэн дээр YOLO (You Only Look Once) загварыг сонгон авсан. YOLO нь дүрсийг нэг дамжлагаар боловсруулж объект илрүүлэх зарчимд тулгуурладаг тул PET/CT дүрс дээрх хавдрын сонирхсон бүсийг автоматаар илрүүлэх, улмаар SUVmax үзүүлэлтийг найдвартай, хурдан тооцоолох боломжийг бүрдүүлдэг.

Ингэснээр энэхүү судалгаа нь PET/CT-д суурилсан SUVmax тооцооллыг хиймэл оюун ухааны дүрс боловсруулалтын аргатай хослуулан ашигласнаар умайн хүзүүний хорт хавдрын үе шат, үсэрхийллийг үнэлэх өртөг багатай, үр дүн сайтай, клиникийн практикт нэвтрэх боломжтой технологийн шийдлийг боловсруулахад чиглэж байна.

Энэхүү судалгаанд санал болгож буй систем нь PET/CT дүрсэд суурилан хавдрын метаболизмын идэвхийг тоон хэлбэрээр үнэлж, умайн хүзүүний хорт хавдрын үе шат болон үсэрхийллийг тодорхойлох шийдлийг танилцуулж байна. Уг систем нь машин сургалт, дүрсийн боловсруулалт, эмнэлзүйн дүрсийн шинжилгээ зэрэг дэвшилтэт аргачлалуудыг хослуулсан бөгөөд PET/CT дүрсээс хавдрын сонирхсон бүсийг автоматаар илрүүлж, SUVmax үзүүлэлтийг бодитой, давтагдах боломжтойгоор тооцоолдог.

Камераар биш харин PET/CT дүрсийн өгөгдөлд тулгуурлан хавдрын байршил, тархалт, метаболизмын идэвхийг тодорхойлж, дүрс боловсруулалтын аргаар дүн шинжилгээ хийдэг нь энэхүү системийн гол онцлог юм. Дүрс боловсруулалтын салбарын сүүлийн үеийн ололт, хиймэл оюун ухааны загваруудыг ашигласнаар уламжлалт субъектив үнэлгээг бууруулж, эмнэлзүйн шийдвэр гаргалтыг дэмжих боломж бүрдэж байна.

Иймд PET/CT-д суурилсан SUVmax тооцоолол бүхий энэхүү ухаалаг систем нь хавдрын үе шат тогтоолт, үсэрхийллийн эрт илрүүлэлтэд тулгарч буй асуудлуудыг шийдвэрлэх, оношилгооны нарийвчлалыг сайжруулах, улмаар эмчилгээний үр дүнг нэмэгдүүлэхэд чухал ач холбогдолтой юм.

Эцэст нь, бид PET/CT дүрс дээр бодит хугацаанд сонирхсон бүс илрүүлэх, SUVmax тооцоолох системийн хөгжлийн чиг хандлагыг

тодорхойлж, цаашдын судалгаа, хөгжүүлэлтийн боломжит чиглэлүүдийг судалсан. Мөн системийн онцлог, давуу талуудыг SUVmean, MTV, TLG зэрэг бусад PET үзүүлэлтэд суурилсан судалгааныудтай харьцуулан авч үзснээр, манай санал болгож буй шийдэл нь тооцооллын энгийн байдал, клиникийн хэрэглээнд нийцэх байдал, өртөг–үр ашгийн харьцаагаараа ялгарч байгааг тодотгосон.

Судалгааны ажлуудын харьцуулалтаас харахад өмнөх ижил төстэй судалгаанууд нь өгөгдлийн багцын хэмжээ хязгаарлагдмал, автомат сегментаци дутмаг, эсвэл тооцооллын нарийн төвөгтэй байдал өндөр зэрэг сул талуудтай байсан бол энэхүү судалгаанд YOLO загварт суурилсан объект илрүүлэлтийг ашиглан PET/CT дүрс дээрх хавдрын бүсийг автоматаар тодорхойлж, SUVmax-ийг найдвартай тооцоолох боломжийг бүрдүүлсэн нь онцлог юм.

Ингэснээр дан ганц хавдрыг илрүүлэхээс гадна түүний метаболизмын идэвх, орон зайн тархалт, үсэрхийллийн боломжит хэв шинжийг тоон мэдээлэлд суурилан үнэлэх боломж бүрдэж, эмч, мэргэжилтнүүдэд илүү бодитой, нотолгоонд суурилсан шийдвэр гаргахад дэмжлэг үзүүлэх цогц системийг санал болгож байна.

II. ХЭРЭГЛЭГЧИЙН СУДАЛГАА

A. Хэрэглэгчийн тодорхойлолт

Энэхүү судалгааны хэрэглэгч нь умайн хүзүүний хорт хавдраар оношлогдож, онош баталгаажуулалт, үе шат тогтоолт болон эмчилгээний үр дүнгийн үнэлгээний зорилгоор PET/CT шинжилгээнд хамрагдсан эмэгтэй өвчтөнүүд юм. Эдгээр өвчтөнүүд нь ихэвчлэн 30–65 насны, хөдөлмөрийн насны эмэгтэйчүүд байх ба хавдрын үе шат, эмчилгээний тактикаас хамааран давтан дүрс оношилгоонд орох шаардлагатай болдог.

Умайн хүзүүний хорт хавдартай өвчтөнүүдийн хувьд хавдрын үе шат, үсэрхийллийн үнэлгээ нь эмчилгээний аргыг (мэс засал, хими эмчилгээ, туяа эмчилгээ, хавсарсан эмчилгээ) сонгоход шууд нөлөөлдөг. Иймээс оношилгооны үнэн зөв, хурдан, давтагдах боломжтой байдал нь өвчтөний амьд үлдэх магадлал, амьдралын чанарт чухал ач холбогдолтой.

B. Хэрэглэгчийн өнөөгийн тулгамдсан асуудал

PET/CT шинжилгээ нь хавдрын метаболизмын идэвхийг үнэлэх өндөр ач холбогдолтой боловч дараах хүндрэл, бэрхшээлүүд өвчтөнүүдийн хувьд түгээмэл ажиглагддаг. Үүнд:

- Дүрс уншилт, тайлбар нь ихэвчлэн эмчийн туршлагаас хамааралтай тул өөр өөр эмчийн дүгнэлт хоорондоо зөрөх тохиолдол гардаг.
- SUVmax зэрэг тоон үзүүлэлтийг гар ажиллагаагаар тодорхойлох нь цаг их шаарддаг тул хариу удаашрах.
- Давтан шинжилгээ хийхэд өмнөх үр дүнтэй харьцуулахад субъектив үнэлгээ давамгайлах.
- Онош тодорхойгүй, эсвэл хоорондоо зөрүүтэй дүгнэлтээс үүдэн өвчтөнд сэтгэл зүйн дарамт үүсэх
- Монгол Улсын нөхцөлд PET/CT шинжилгээний хүртээмж хязгаарлагдмал /улсын-1,хувийн-1эмнэлэг/, мэргэжлийн хүний нөөц дутмаг байх

Эдгээр хүчин зүйлс нь өвчтөнд үзүүлэх эмнэлзүйн тусламж үйлчилгээний чанар, тасралтгүй байдалд сөргөөр нөлөөлөх эрсдэлтэй.

C. Хэрэглэгчийн хэрэгцээ, шаардлага

Умайн хүзүүний хорт хавдраар эмчлэгдэж буй өвчтөнүүдийн хувьд дараах үндсэн хэрэгцээ, шаардлагууд тодорхойлогдож байна. Үүнд:

- Оношилгооны үнэн зөв байдал буюу хавдрын үе шат, үсэрхийллийг алдаагүй, бодитой үнэлэх нь өвчтөнд хамгийн чухал шаардлага юм.
- Шинжилгээний хариу хурдан гарах буюу PET/CT шинжилгээний дүнг богино хугацаанд боловсруулах нь эмчилгээ эхлэх хугацааг наашлуулж, сэтгэл зүйн дарамтыг бууруулна.
- Давтагдах боломжтой, стандартчилсан үнэлгээ буюу давтан шинжилгээнд ороход SUVmax болон бусад үзүүлэлтүүдийг өмнөх үр дүнтэй бодитой харьцуулах боломжтой байх.
- Субъектив алдаа багатай дүгнэлт буюу Эмчийн туршлага, ачааллаас хамаарах алдааг багасгасан, автоматжуулсан тоон үнэлгээ.
- Эмчилгээний шийдвэрийг дэмжих ойлгомжтой мэдээлэл буюу өвчтөнд тайлбарлахад хялбар, ойлгомжтой, тоон үндэслэлтэй оношилгооны дүн.

D. Судалгаанд санал болгож буй системийн хэрэглэгчид үзүүлэх давуу тал

Энэхүү судалгаанд боловсруулж буй хиймэл оюун ухаанд суурилсан SUVmax автомат тооцооллын систем нь дээр дурдсан хэрэглэгчийн хэрэгцээг дараах байдлаар хангана. Үүнд:

- ✓ PET/CT дүрсээс хавдрын сонирхсон бүсийг автоматаар илрүүлж, SUVmax-ийг бодитой, давтагдах боломжтойгоор тооцоолно.
- ✓ Хавдрын TNM/FIGO үе шат болон бүс нутгийн, алс эрхтний үсэрхийллийг илүү нарийвчлалтай үнэлэх боломж бүрдүүлнэ.
- ✓ Дүрс уншилтын хугацааг багасгаж, оношилгооны хариуг түргэсгэнэ.
- ✓ Уншигч хоорондын зөрүүг бууруулж, өвчтөнд илүү найдвартай дүгнэлт хүргэх нөхцөлийг бүрдүүлнэ

Эмчилгээний явцын хяналт, үр дүнгийн үнэлгээнд SUVmax-ийн өөрчлөлтийг тоон хэлбэрээр харьцуулах боломж олгоно

Үүнээс дүгнэхэд, умайн хүзүүний хорт хавдраар эмчлэгдэж буй өвчтөнүүдийн хувьд оношилгооны нарийвчлал, хурд, найдвартай байдал нь амьд үлдэх түвшин, амьдралын чанарт шууд нөлөөлдөг чухал хүчин зүйл юм. Хиймэл оюун ухаанд суурилсан SUVmax автомат тооцооллын систем нь өвчтөн

төвтэй, нотолгоонд суурилсан, стандартчилсан оношилгооны орчинг бүрдүүлж, эмнэлзүйн алдаа эрсдэлийг бууруулах замаар хэрэглэгчийн бодит хэрэгцээг хангах өндөр боломжтой гэж үзэв.

III. СУДЛАГДСАН БАЙДАЛ

Өмнөх судалгаануудад умайн хүзүүний хорт хавдрын үе шат, үсэрхийллийг үнэлэх зорилгоор PET/CT дүрс оношлогоонд суурилсан метаболик үзүүлэлтүүдийг өргөнөөр ашиглаж ирсэн байдаг. Ялангуяа 18F-FDG PET/CT шинжилгээгээр хэмжигддэг SUVmax, SUVmean, MTV, TLG зэрэг үзүүлэлтүүд нь хавдрын биологийн идэвх, түрэмгий шинж, эмчилгээний үр дүн болон урьдчилсан таамаглалтай статистик ач холбогдолтой холбоотой болохыг олон судалгаагаар тогтоосон байна.

Гэвч эдгээр судалгааны ихэнх нь PET/CT дүрс дээр сонирхсон бүсийг (ROI) гар ажиллагаагаар тодорхойлж, SUVmax-ийг хэмжих уламжлалт аргачлалд тулгуурласан бөгөөд энэ нь дараах сул талуудыг үүсгэж байна. Үүнд:

- ROI сонголт нь эмчийн туршлагаас хамааарч субъектив байх
- Уншигч хоорондын хэлбэлзэл (inter-observer variability) ихтэй
- Давтан шинжилгээний үед харьцуулах боломж хязгаарлагдмал
- Цаг хугацаа, хүний нөөц их шаарддаг

Иймээс SUVmax зэрэг метаболик үзүүлэлтүүдийг илүү бодитой, давтагдах боломжтой, автоматжуулсан байдлаар тооцоолох шаардлага судалгааны түвшинд ч, клиникийн практикт ч тулгамдсан асуудал болж байна.

1. PET болон CT дүрс оношилгооны ялгаа

PET (Positron Emission Tomography) болон CT (Computed Tomography) нь орчин үеийн дүрс оношилгоонд өргөн хэрэглэгддэг боловч мэдээллийн шинж чанарын хувьд ялгаатай хоёр үндсэн технологи юм. PET дүрслэл нь биеийн эд, эсийн түвшний бодисын солилцооны идэвхийг үнэлэх боломж олгодог функциональ оношилгооны арга бөгөөд ихэвчлэн 18F-FDG зэрэг радио идэвхт трассер бодис ашиглан глюкозын хэрэглээг хэмждэг. Энэ нь хавдрын эсүүдийн өндөр метаболик идэвхийг илрүүлэхэд онцгой ач холбогдолтой.

Харин CT нь рентген туяанд суурилсан анатомийн дүрслэлийн арга бөгөөд эрхтэн, эд, ясны бүтэц, байрлал, хэмжээ зэргийг өндөр нарийвчлалтайгаар харуулдаг. CT нь хавдрын байршил, хэлбэр, хэмжээ, хөрш эд эрхтэнтэй харьцах хамаарлыг тодорхойлоход чухал үүрэгтэй боловч эсийн түвшний үйл ажиллагааны мэдээлэл өгөх боломж хязгаарлагдмал байдаг.

Иймээс PET болон CT нь хоорондоо харилцан хослуулан ашиглах боломжтой бөгөөд PET нь функциональ мэдээлэл, CT нь анатомийн мэдээллийг тус тус өгдөг. Доорх хүснэгтэд эдгээр хоёр аргын үндсэн ялгааг харьцуулан үзүүлэв.

Хүснэгт 1. PET, CT дүрс оношилгооны үзүүлэлтүүдийн харьцуулалт

Үзүүлэлт	PET	CT
Дүрслэлийн төрөл	Функциональ (метаболизм)	Анатомийн (бүтэц)
Ашиглах зарчим	Радио идэвхт трассер	Рентген туяа
Өгөгдлийн төрөл	Эсийн идэвхжил, глюкозын хэрэглээ	Эрхтэн, эдийн бүтэц
Давуу тал	Эртийн үеийн хавдрын илрүүлэлт	Байршил, хэмжээ тодорхойлох
Сул тал	Байршлын нарийвчлал бага	Функциональ мэдээлэлгүй

2. PET-CT хосолсон дүрслэлийн ач холбогдол

PET болон CT дүрслэлийн аргуудыг хослуулсан PET/CT технологи нь функциональ болон анатомийн мэдээллийг нэг дор нэгтгэн харуулах боломжийг олгосноор орчин үеийн хавдар судлалын оношилгоонд чухал байр суурь эзэлж байна. PET дүрслэлээр илэрсэн метаболик идэвх өндөртэй бүсийг CT-ийн өндөр нарийвчлалтай анатомийн зурагтай давхардуулснаар хавдрын голомтын байршлыг илүү нарийвчлалтай тодорхойлох боломж бүрддэг.

Ялангуяа умайн хүзүүний хорт хавдрын хувьд PET/CT нь анхдагч хавдрын үнэлгээ, тунгалгийн булчирхайн оролцоо, алслагдсан үсэрхийллийг илрүүлэх, эмчилгээний үр дүнг хянахад өргөн хэрэглэгддэг. PET/CT шинжилгээнээс гардаг SUVmax, SUVmean, MTV, TLG зэрэг метаболик параметрүүд нь хавдрын биологийн идэвх, түрэмгий шинж, урьдчилсан таамаглалтай нягт холбоотой болохыг олон судалгаанд дурдсан байдаг.

Гэсэн хэдий ч PET/CT дүрс дээрх эдгээр параметрүүдийг тодорхойлох үйл явц нь ихэвчлэн гар ажиллагаанд тулгуурладаг бөгөөд сонирхсон бүс (ROI)-ийг эмч гараар сонгох шаардлагатай байдаг. Энэ нь эмнэлзүйн тайлангийн ялгаа үүсгэх, субъектив нөлөө ихсэх, давтагдах чадвар буурах, мөн цаг хугацаа их зарцуулах сул талуудыг дагуулдаг.

Иймд PET/CT дүрслэлийн давуу талыг бүрэн ашиглахын тулд хиймэл оюун ухаан, ялангуяа гүн сургалтын аргуудыг ашиглан хавдрын голомт болон үсэрхийллийг автоматаар илрүүлэх, метаболик параметрүүдийг стандартчилсан, хурдан, давтагдах боломжтой байдлаар тооцоолох шаардлага зүй ёсоор тавигдаж байна. Энэ нь оношилгооны нарийвчлал, үр ашиг, клиник хэрэглээний хүртээмжийг мэдэгдэхүйц сайжруулах үндэс болж байна.

3. Ижил төсөөтэй судалгааны ажлууд

А. Судалгааны ажил

Takagi et al. (2018) нарын судалгаанд PET/CT шинжилгээгээр хэмжигдсэн SUVmax үзүүлэлтийг умайн хүзүүний хорт хавдрын эхний шатны үнэлгээнд ашиглах боломжийг судалсан. Судалгааны үр дүнд SUVmax нь өвчний шат ≤ I болон ≥ II хооронд статистик ач холбогдолтой ялгаатай байгааг харуулсан бөгөөд өндөр SUVmax

утга нь өвчний ахилт, дахин сэргэлтийн эрсдэл өндөртэй холбоотой болохыг тогтоосон.

❖ Тус судалгаа нь SUVmax-ийг эмнэлзүйн чухал үзүүлэлт болохыг нотолсон боловч хэмжилт нь гар ажиллагаанд тулгуурласан, автомат сегментаци хийгдээгүй, бодит хугацааны хэрэглээ хязгаарлагдмал байсан сул талтай.

B. Судалгааны ажил

Luo et al. (2021) нарын судалгаанд умайн хүзүүний хорт хавдрын эмчилгээний үр дүнг урьдчилан таамаглах зорилгоор статистик загвар, дүрс оношлогоо болон машин сургалтын аргуудыг харьцуулан судалсан. Судалгаанд MRI, CT, PET/CT дүрсийн өгөгдөл болон радиомикс, нейрон сүлжээ, Random Forest зэрэг машин сургалтын аргуудыг ашигласан.

❖ Судалгааны дүгнэлтээс харахад машин сургалтын аргууд нь уламжлалт статистик аргаас олон хувьсагчтай өгөгдөл дээр илүү давуу талтай боловч өгөгдлийн хэмжээ бага, загварын тайлбарлагч чанар сул, клиник хэрэглээнд шууд нэвтрүүлэхэд хүндрэлтэй байгааг онцолсон.

C. Судалгааны ажил

Otani et al. (2023) нар умайн хүзүүний хорт хавдрын хоёр гол хистологийн төрөл болох squamous cell carcinoma болон adenocarcinoma-ийн өвчтөнүүдэд эмчилгээний өмнөх FDG PET/CT-ээр хэмжигдсэн SUVmax болон бусад метаболик параметрүүдийн урьдчилан таамаглах чадамжийг судалсан.

❖ Судалгаагаар PET/CT-ийн метаболик үзүүлэлтүүд нь өвчний урьдчилсан таамаглалд ач холбогдолтой боловч хистологийн төрөл нь SUVmax-ийн тайлбарт нөлөөлөх боломжтойг тэмдэглэсэн. Энэ нь SUVmax-ийг ганцаарчилсан үзүүлэлт бус, автомат, олон параметртэй системийн хүрээнд ашиглах шаардлагыг харуулж байна.

D. Судалгааны ажил

Sari et al. (2023) нар IVB шатны умайн хүзүүний хорт хавдартай өвчтөнүүдэд метастаз болон анхдагч хавдрын SUVmax-ийн харьцааг (metastasis-to-tumor SUVmax ratio) шинэ урьдчилан таамаглах индекс болгон судалсан. Судалгааны үр дүнд энэхүү харьцаа өндөр байх нь муу урьдчилсан таамаглалтай холбоотой байж болохыг харуулсан.

❖ Гэвч энэхүү харьцааг бодит практикт хэрэглэхэд метастаз болон анхдагч голомтыг нарийн, хурдан илрүүлэх автомат механизм дутмаг байгаа нь тодорхой болсон.

E. Судалгааны ажил

Zhang et al. (2021) нарын судалгаанд SUVmax, MTV, TLG зэрэг PET/CT-ийн метаболик параметрүүдийг умайн хүзүүний хорт хавдрын эмчилгээний үр дүн, ахилт, дахин сэргэлттэй уялдуулан судалсан. Судалгаагаар эдгээр үзүүлэлтүүд нь статистик ач холбогдолтой холбоотой болохыг тогтоосон боловч параметр бүрийг гар ажиллагаагаар тодорхойлсон, сегментацийн аргачлал ялгаатай байсан нь үр дүнг ерөнхийчлэхэд хүндрэлтэй байжээ.

❖ Судалгааны дүгнэлтээр PET/CT-ийн метаболик параметрүүдийг клиник практикт ашиглах боломжтой боловч автоматжуулалт, стандартчлал зайлшгүй шаардлагатайг онцолсон.

Судалгааны ажлуудын харьцуулсан дүгнэлт

Дээр дурдсан судалгаанууд нь PET/CT-д суурилсан SUVmax болон бусад метаболик үзүүлэлтүүд нь умайн хүзүүний хорт хавдрын үе шат тогтоолт, үсэрхийллийн үнэлгээ, эмчилгээний төлөвлөлт, урьдчилсан таамаглалд чухал ач холбогдолтой болохыг нотолж байна. Гэвч дараах нийтлэг сул талууд ажиглагдаж байна. Үүнд:

- SUVmax тооцоолол гар ажиллагаанд тулгуурласан
- Автомат сегментаци, объект илрүүлэлт дутмаг
- Уншигч хоорондын хэлбэлзэл өндөр
- Бодит хугацааны хэрэглээ хязгаарлагдмал
- AI-д суурилсан системийг клиник практикт бүрэн нэвтрүүлээгүй

Иймд дээрх судалгаануудын сул талыг нөхөх зорилгоор PET/CT дүрсээс хавдрын голомт болон үсэрхийллийг хиймэл оюун ухаан, гүн сургалтын аргаар илрүүлж, SUVmax-ийг стандартчилсан, хурдан, давтагдах боломжтой байдлаар тооцоолох шинэ аргачлал боловсруулах шаардлага зүй ёсоор тавигдаж байна.

Энэхүү судалгааны ажил нь дээрх судалгаануудын үр дүнг нэгтгэн, тэдгээрийн хязгаарлалтыг арилгах зорилгоор AI-д суурилсан SUVmax автомат тооцооллын системийг санал болгож буйгаараа онцлог бөгөөд умайн хүзүүний хорт хавдрын үе шат болон үсэрхийллийг илүү үнэн зөв, клиникийн практикт хэрэгжих боломжтой түвшинд үнэлэхэд чиглэж байна.

Судалгааны ажлуудын харьцуулсан хүснэгт

Хүснэгт 2. Умайн хүзүүний хорт хавдрын PET/CT-д суурилсан судалгааануудын харьцуулалт

Судалгааны ажил	Арга зүй	Давуу тал	Хязгаарлалт	Миний судалгаатай нийцэмж
Takagi et al., 2018	Гар ажиллагаатай ROI, статистик шинжилгээ	SUVmax нь үе шат, ахилттай холбоотойг нотолсон	Автоматжилтгүй, уншигчийн субъектив нөлөө их	SUVmax-ийн ач холбогдлыг баталсан суурь судалгаа
Luo et al., 2021	Машин сургалт, статистик загвар	AI ба дүрс оношлогоны хослолын боломж	Interpretability сул, клиник нэвтрэлт хязгаарлагдмал	AI ашиглах концепцийн суурь
Otani et al., 2023	Хистологийн ангилалд суурилс	Урьдчилан таамаглах чадамжийг	Хистологийн нөлөө, автомат сегментаци дутмаг	SUVmax-ийг олон хүчин зүйлтэй хослуулах

	ан анализ	харуулсан		шаардлагыг харуулсан
Sari et al., 2023	Харьцаат үзүүлэлт	Шинэ прогностик индекс санал болгосон	Метастаз, голомтыг гар ажиллагаа гаар тодорхойлсон	AI-аар ratio-г автомат тооцоолох боломж
Zhang et al., 2021	Статистик корреляци	Метаболик параметр үүдийн ач холбогдлыг нотолсон	Тооцооолол төвөгтэй, стандартчилал сул	Автоматжуулалтын хэрэгцээг баталсан
Энэхүү судалгаа	YOLO + гүн сургалт + сегментаци	Хурдан, стандартчилсан, давтагдах боломжтой	Өгөгдлийн хэмжээ өсгөх шаардлагатай	Клиник практикт шууд хэрэгжих боломж

Энэхүү судалгааны ажил нь өмнөх PET/CT-д суурилсан умайн хүзүүний хорт хавдрын судалгаануудын үр дүнд тулгуурлан, тэдгээрийн нийтлэг сул талыг арилгах зорилготойгоор боловсруулагдсан цогц шийдэл юм. Өмнөх судалгаанууд SUVmax болон бусад метаболик параметрүүд нь хавдрын үе шат, үсэрхийлэл, эмчилгээний үр дүнтэй нягт холбоотой болохыг нотолсон боловч эдгээр үзүүлэлтүүдийг тодорхойлох үйл явц нь ихэвчлэн гар ажиллагаатай, субъектив шинжтэй, клиникийн бодит хэрэглээнд цаг хугацаа их шаарддаг байсаар ирсэн.

Харин энэхүү судалгааны ажлын гол давуу тал нь PET болон CT дүрс оношилгооны зургийг хослуулан хавдрын голомт болон үсэрхийллийг хиймэл оюун, гүн сургалтын (YOLO) загвар ашиглан автоматаар илрүүлж, SUVmax-ийг стандартчилсан байдлаар тооцоолон, эмнэлзүйн тайлан гаргахад оршино. Үүний үр дүнд эмчийн туршлагаас хамаарах субъектив алдаа багасахын зэрэгцээ оношилгооны нарийвчлал, үр ашиг дээшилж, хурд нэмэгдсэнээр үзлэгийн 2-3 сар хүлээгдээд буй оочир дараалал багасна.

Мөн энэхүү судалгаа нь зөвхөн хавдрын илрүүлэлтэд хязгаарлагдахгүй, харин SUVmax тархалтын мэдээлэлд тулгуурлан TNM/FIGO үе шат тогтоолт, бүс нутгийн болон эрхтний үсэрхийллийн үнэлгээ хийх боломжийг нэг дор шийдсэнээрээ онцлог юм. Энэ нь өмнөх судалгаануудад тус тусад нь авч үзэгдэж байсан асуудлуудыг нэг системд нэгтгэж, клиникийн шийдвэр гаргалтыг бодитоор дэмжих түвшинд хүргэж байна.

Түүнчлэн Монгол Улсын нөхцөлд PET/CT шинжилгээний хүртээмж хязгаарлагдмал, мэргэжлийн хүний нөөц дутмаг байгаа өнөөгийн байдалд энэхүү автоматжуулсан AI систем нь оношлогооны үр ашгийг нэмэгдүүлэх, өвчтөнд үйлчилгээ үзүүлэх хугацааг богиносгох, эмчилгээний үр дүнг сайжруулах практик ач холбогдолтой гэж үзэж байна.

Иймд уг судалгаа нь PET/CT-д суурилсан SUVmax үнэлгээг шинжлэх ухааны судалгаанаас клиникийн практик руу шилжүүлэх гүүр болж, умайн хүзүүний хорт хавдрын оношилгоо, үе шат тогтоолт, үсэрхийллийн үнэлгээнд хиймэл оюун ухааныг бодитоор нэвтрүүлэх шинэ шатанд хүргэж байна.

IV. ТӨСЛИЙН ХЭСЭГ

Энэхүү хэсэгт PET/CT дүрс оношилгоонд суурилсан умайн хүзүүний хорт хавдрын үе шат болон үсэрхийллийг үнэлэх хиймэл оюун ухаанд суурилсан системийн бүтэц, үйл ажиллагааны дараалал, хэрэгжилтийн үндсэн зарчмуудыг тодорхойлно. Төслийн хүрээнд боловсруулсан систем нь хавдрын сонирхсон бүсийг автоматаар илрүүлэх, SUVmax үзүүлэлтийг стандартчилсан байдлаар тооцоолох, улмаар TNM/FIGO үе шат тогтоолтыг дэмжихэд чиглэсэн бөгөөд клиникийн практикт ашиглах боломж, давуу болон сул талыг цогцоор нь авч үзсэн.

Мөн энэхүү судалгаа нь өвчтөний бодит өгөгдөл дээр суурилан хийгдэж байгаа тул өгөгдлийн аюулгүй байдал, нууцлалыг хангах үүднээс өгөгдлийн багцыг бүрдүүлсний дараа шууд ашиглахгүйгээр холбогдох стандартын дагуу дата анонимжуулалт хийж боловсруулсан болно. Үүнд хувь хүнийг таних боломжтой бүх мэдээллийг (нэр, регистрийн дугаар, огноо, эмнэлгийн дугаар зэрэг) устгах буюу кодчилох замаар өгөгдлийг танигдах боломжгүй хэлбэрт шилжүүлж, улмаар хиймэл оюун ухааны загвар сургах болон үнэлгээ хийх процессод ашигласан. Дараах хүснэгтэд дата аномжуулалт хийхээс өмнөх болон хийсний дараах үр дүнг жишээ өгөгдөл дээр хүснэгтээр харуулав.

Хүснэгт 3. Өгөгдлийн аномжуулалт хийхээс өмнөх болон дараах үр дүнгийн харьцуулалт

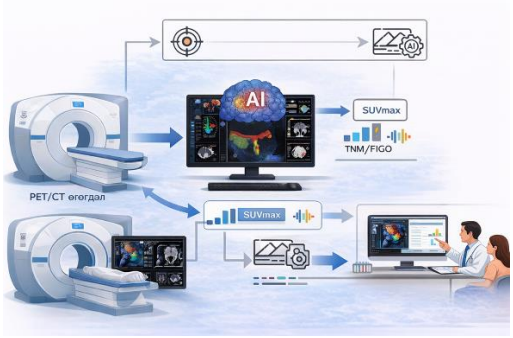
Аномжуулалт хийхээс өмнө	**Аномжуулалт хийсний дараа**
Батын Дулам	Patient_001
РД: УБ12345678	ID_0001
Төрсөн огноо: 1985-03-12	1985
Хүйс: Эмэгтэй	F
Эмнэлгийн нэр	Hospital_A
Шинжилгээний огноо: 2023-05-10	2023 (жилээр хадгалсан)

Төслийн үйл ажиллагаа

Төслийн үйл ажиллагаа нь умайн хүзүүний хорт хавдраар өвчилсөн эмэгтэйчүүдийн оношилгоо, үе шат тогтоолт, эмчилгээний хяналтыг сайжруулахад чиглэсэн бөгөөд PET/CT дүрс оношилгоонд суурилсан SUVmax автомат тооцооллын ухаалаг

системийг хөгжүүлэх, турших, клиникийн хэрэглээнд нийцүүлэх ажлуудаас бүрдэнэ.

Төслийн эхний шатанд умайн хүзүүний хорт хавдрын PET/CT дүрсийн өгөгдөлд ашиглагдаж буй SUVmax тодорхойлох уламжлалт аргачлалууд, хиймэл оюун ухаанд суурилсан объект илрүүлэлт, сегментацийн технологиудыг судалж, гүн сургалтын YOLO алгоритмыг ашиглан хавдрын голомт болон үсэрхийллийг автоматаар илрүүлэх үндсэн загварыг боловсруулсан.



1-р зураг. Системийн ерөнхий ажиллагаа

Зураг 1-ээр системийн ерөнхий ажиллагааг харуулав. Зурагт харуулснаар энэхүү системийн ерөнхий ажиллагааны зурагт бодит PET/CT дүрс оношилгоонд суурилан умайн хүзүүний хорт хавдрыг илрүүлэх, метаболизмын идэвхийг (SUVmax) тооцоолох, хавдрын тархалт болон үе шатыг үнэлэх үйл явцыг автоматжуулах замаар оношилгооны нарийвчлал, хурд, давтагдах боломжийг сайжруулж, эмнэлзүйн шийдвэр гаргалтыг дэмжих хиймэл оюун ухаанд суурилсан системийн архитектурыг харуулав. Өгөгдөл оруулах, дүрс боловсруулалт, YOLO загварт суурилсан объект илрүүлэлт, ROI сегментаци, SUVmax тооцоолол, үе шат тогтоолтыг дэмжих мэдээлэл боловсруулах модулиудын уялдаа холбоог харуулсан.

Системийн үндсэн ажиллагаа дараах үе шатуудаас бүрдэнэ.

1. **Өгөгдөл хүлээн авах ба анхны боловсруулалт**
   Энэ нь эхний шат бөгөөд системд дата анонимжуулалт хийж боловсруулах, PET/CT DICOM өгөгдөл оруулах, зураг авах, SUV тооцоолол хийх, 3D хэмжээст дүрсийг гаргах зэрэг үйлдлүүдийг багтаана. Энэ шатанд өгөгдлийг боловсруулж, шаардлагатай бүх анхны мэдээллийг гаргаж авна.
2. **Дүрс боловсруулах ба шинжилгээ хийх**
   Энэ шат нь дараагийн боловсруулах үе юм. Тухайлбал, элэгний арын SUV тооцоолол хийх, адаптив хязгаар тогтоох, холбогдсон бүрдэл хэсгүүдийг илрүүлэх, элэгний харьцааг тооцох зэрэг үйлдлүүдийг багтааж, дүрсийг нарийвчлан боловсруулна. Энэ үе шатанд шинжилгээ хийж, эмчилгээнд шаардлагатай нарийвчилсан мэдээллийг гаргана.
3. **Тайлан гаргах ба үр дүн боловсруулалт**
   Эцсийн шат нь бүхий л боловсруулсан өгөгдлийг ашиглан эмчилгээний тайлан дүгнэлт гаргах, 3D визуализаци хийх, үр дүнг экспортлох зэрэг үйлдлүүдийг багтаана. Тайлан болон дүгнэлт нь эмч, өвчтөнд өгөгдлийн нарийвчилсан тайлбарыг олгож, үр дүнг шийдвэрлэхэд тусална.

Төслийн шинэлэг тал

Энэхүү төслийн шинэлэг байдал нь өвчтөний бодит PET/CT зурган өгөгдөлд суурилсан умайн хүзүүний хорт хавдрын үнэлгээг хиймэл оюун ухааны аргаар бүрэн автоматжуулж, клиникийн практикт бодитоор ашиглах боломжтой шийдэл санал болгож буйд оршино.

Нэгдүгээрт, SUVmax-ийг гар ажиллагаагүй, стандартчилсан байдлаар автоматаар тооцоолдог нь уншигч хоорондын субъектив ялгааг бууруулж, оношилгооны найдвартай байдлыг нэмэгдүүлж байна.

Хоёрдугаарт, зөвхөн анхдагч хавдрыг бус, бүс нутгийн болон алс үсэрхийллийг нэг системийн хүрээнд илрүүлж, метаболизмын идэвхийг иж бүрэн үнэлэх боломжийг бүрдүүлсэн.

Гуравдугаарт, SUVmax үзүүлэлтийг TNM/FIGO үе шат тогтоолтын логиктой уялдуулан ашиглах боломж олгосноор эмчилгээний тактик сонголт, эмчилгээний хариу урвалыг үнэлэхэд дэмжлэг үзүүлнэ.

Дөрөвдүгээрт, нэмэлт өндөр өртөгтэй төхөөрөмж шаардахгүй, программ хангамжид суурилсан шийдэл нь Монгол Улсын нөхцөлд практик ач холбогдол өндөртэй.

Төслийн давуу тал

- SUVmax-ийг стандартчилсан, давтагдах боломжтой байдлаар тооцоолно.
- PET/CT уншилтын хугацааг мэдэгдэхүйц хэмнэнэ.
- Клиник шийдвэр гаргалтыг дэмжинэ.
- Давтан шинжилгээг тоон байдлаар харьцуулах боломжтой.
- Бусад хавдрын төрөлд өргөтгөх боломжтой.

Сул тал

- Сургалтын өгөгдлийн хэмжээ, чанараас гүйцэтгэл хамаарна.
- PET/CT төхөөрөмжийн хүртээмж хязгаарлагдмал.
- Клиник орчинд бүрэн нэвтрүүлэхийн өмнө илүү өргөн хүрээний туршилт шаардлагатай.

Үүнээс дүгнэхэд, энэхүү төслийн хүрээнд өвчтөний бодит PET/CT дүрс оношилгоонд

суурилсан умайн хүзүүний хорт хавдрын SUVmax автомат тооцооллын ухаалаг системийг боловсруулж, хавдрын үе шат болон үсэрхийллийг үнэлэх клиник ач холбогдол бүхий шийдлийг санал болголоо. Судалгааны үр дүн нь уламжлалт гар ажиллагааг халж, оношилгооны хурд, нарийвчлал, давтагдах боломжийг сайжруулах, улмаар эмнэлзүйн шийдвэр гаргалтыг бодитой, нотолгоонд суурилсан болгоход чухал хувь нэмэр оруулна гэж үзэж байна.

V. МОДУЛЬ

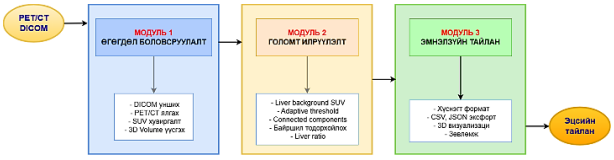
Манай систем нь өвчтөний PET/CT дүрс оношилгоонд суурилан умайн хүзүүний хорт хавдрыг илрүүлэх, метаболизмын идэвхийг (SUVmax) тооцоолох, хавдрын тархалт болон үе шатыг үнэлэх үйл явцыг автоматжуулах замаар оношилгооны нарийвчлал, хурд, давтагдах боломжийг сайжруулж, эмнэлзүйн шийдвэр гаргалтыг дэмжих хиймэл оюун ухаанд суурилсан системийг боловсруулахыг зорьсон бөгөөд 3 үндсэн модулиас бүрдэнэ. Модулиуд хоорондоо уялдаа холбоотой ажиллаж, клиникийн хэрэглээнд зориулсан нэгтгэсэн үр дүнг бий болгоно.

Модулиуд:

1. **Модуль 1: Өгөгдөл боловсруулалт**
   Энэхүү модуль нь системд орж ирж буй өгөгдлийн анхан шатны боловсруулалтыг гүйцэтгэх үндсэн хэсэг юм. Уг шатанд PET/CT төрлийн дүрсүүдийг DICOM (Digital Imaging and Communications in Medicine) стандарт форматаар хүлээн авна.
2. **Модуль 2: Голомт илрүүлэлт**
   Энэхүү модуль нь дүрсийг боловсруулж, өвчний шинж тэмдгийг илрүүлэхэд чиглэнэ. Элэгний арын зузаан хэсгээс SUV тооцоолол хийж, адаптив хязгаар тогтоох арга ашиглан зурган дээрх эх сурвалжийн мэдээллийг илүү зөв тусгах бөгөөд, холбогдсон бүрдэл хэсгүүдийг тодорхойлон, хураангуй мэдээлэл гаргана. Мөн элэгний харьцааг тооцож, өвчний тархалтыг илрүүлнэ.
3. **Модуль 3: Дүгнэлт болон тайлан**
   Энэхүү модуль нь бүх боловсруулсан өгөгдлийг эмчилгээний дүгнэлт болон тайлан болгон гаргаж, хэрэглэгчид эсвэл эмч нарт тайлагнаж, эмчилгээний үр дүнг танилцуулна. Өгөгдлийг CSV болон JSON форматуудад экспортлох боломжтой бөгөөд 3D дүрс ашиглан тайланг илүү ойлгомжтой харуулах бөгөөд эцсийн тайланд оношлогоо, эмчилгээний зааварчилгаа, үр дүнг бүрэн тайлагнана.

Модулиудын холбоо хамаарал:

- Өгөгдөл боловсруулалт → Голомт илрүүлэлт → Эмнэлзүйн тайлан, дүгнэлт гаргах
- Модулиудын хоорондын өгөгдлийн урсгал (data flow) нь нэг модулын гарсан өгөгдлийг дараагийн модульд дамжуулж, эцсийн тайланг гаргана.



2-р зураг. Модулиудын холбоо хамаарал

Зураг 2-оор системийн ерөнхий ажиллагааг харуулав. Зурагт харуулснаар энэхүү системийн модуль бүр нь дараалалын дагуу мэдээллийг боловсруулж, эцсийн үр дүнг клиникийн хэрэглээнд ашиглахад бэлэн болгоно. Урсгал нь PET/CT өгөгдлөөс эхлэн дүрсийн боловсруулалт, хавдрын голомт илрүүлэлт, метаболизмын идэвхийг тооцоолох, үе шат тогтоох, тайлан болон дүгнэлт гаргах хүртэлх бүх үндсэн үйл явцын холбоог харуулж байна.

VI. ӨГӨГДӨЛ БОЛОВСРУУЛАЛТ

Энэ шат нь өгөгдлийг анх боловсруулдаг хэсэг юм. PET/CT DICOM форматаар өгөгдөл орж ирнэ. DICOM нь медицин зураг боловсруулалтын стандарт формат бөгөөд, энэ нь бидэнд PET (Positron Emission Tomography) болон CT (Computed Tomography) зураг авах боломжийг олгодог. Модуль 1-ийн үндсэн зорилго нь эдгээр зурагнуудыг боловсруулсны дараа, тухайн мэдээллээс шаардлагатай өгөгдлийг боловсруулан гаргах юм.

- **PET/CT дүрс**: Энэ нь өвчтөний биед байгаа өвчин эсвэл эдийн дүрсийг тодорхойлж, эмчилгээнд туслах гол өгөгдөл юм.
- **SUV тооцоолол**: SUV (Standardized Uptake Value) нь тухайн хэсгийн радиологийн хэмжилтийг стандартад тохируулсан утга юм. Эдгээр утгуудыг ашиглан эмчилгээний үр дүнг илүү нарийвчилж үнэлэх боломжтой.
- **3D хэмжээст дүрс**: Энэ шатанд 3D дүрсүүдийг гаргаж авах бөгөөд, энэ нь эд эрхтний бүтцийг харах болон үнэлэхэд маш чухал юм. Энэ нь эмчид дүрсийг томруулан, нарийвчилсан үзүүлэлтүүдийг харах боломжийг олгоно.

Энэхүү модуль нь дараагийн AI боловсруулалтын чанарт шууд нөлөөлөх тул дангаар болон системийн урсгалын хувьд (PET/CT → Дүрс боловсруулалт → Объект илрүүлэлт → SUVmax тооцоолол → Үе шат тогтоолт → Үр дүн гаргах) маш чухал үүрэгтэй.

Энэхүү модуль нь дараагийн AI боловсруулалтын чанарт шууд нөлөөлнө.

VII. ГОЛОМТ ИЛРҮҮЛЭЛТ

Модуль 2 нь PET/CT дүрсээс хавдрын голомт болон үсэрхийллийг илрүүлэх үндсэн зорилготой системийн гол модуль бөгөөд энэхүү процесс нь олон үе шаттайгаар явагдана. Доорх алгоритм нь SUV (Standardized Uptake Value)-д суурилсан босго утгын аргууд болон дүрс боловсруулалтын техникийг хослуулан хэрэгжүүлсэн болно.

1. SUV утга тооцоолол

Эхний шатанд PET дүрсээс SUV утгыг тооцоолж, тухайн эд эсийн метаболик идэвхийг үнэлнэ. SUV нь глюкозын хэрэглээг илэрхийлэх үзүүлэлт бөгөөд хавдрын эсүүд ихэвчлэн өндөр SUV утгатай байдаг. Ялангуяа элэгний бүсийг суурь (reference) болгон ашиглаж, хэвийн метаболизмын түвшинг тодорхойлно. Элэг нь харьцангуй тогтвортой SUV тархалттай тул дараагийн шатны харьцуулалт хийхэд чухал ач холбогдолтой.

2. Threshold-той харьцуулах

Энэхүү шатанд өндөр SUV утгатай сэжигтэй бүсүүдийг ялган илрүүлэх зорилгоор босго утгын аргуудыг ашигласан. Судалгаанд дараах хоёр аргыг харьцуулан судалсан.

- **Статик threshold арга:** Урьдчилан тогтоосон нэг утгыг (жишээ нь SUVmax > X) ашиглан тухайн бүсийг эрсдэлтэй эсэхийг шууд тодорхойлдог. Гэвч энэхүү арга нь дараах сул талуудтай.
    - Өвчтөний биеийн онцлог (жин, бодисын солилцоо)
    - Хэрэглэж буй эмчилгээ (хими, туяа эмчилгээ)
    - Дүрс авах нөхцөл (scanner, протокол)
    - зэрэг олон хүчин зүйлээс хамааран алдаа үүсэх магадлалтай.

  Дараах хэсэгт эрхтэн тус бүрийн хэвийн SUVmax утгыг хүснэгтээр харууллаа. Энэхүү хүснэгтэн мэдээлэл нь суурь мэдээлэл болгон ашиглагдах болно.

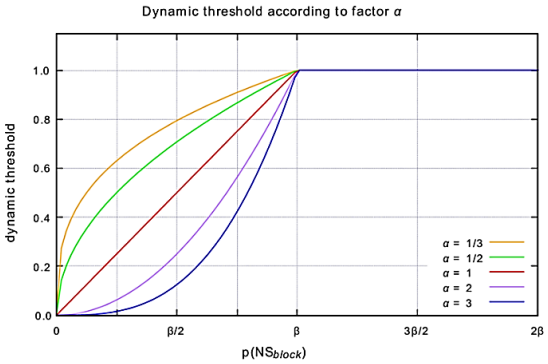
Хүснэгт 4. Эд эрхтний SuvMax утгууд

Эд эрхтэн	SuvMax утга
Тархи (бор бодис)	8–12
Зүрх (миокард)	3–7
Элэг	2.4–4.0
Дэлүү	1.8–3.0
Ясны чөмөг	1.5–2.5
Ясны кортекс	1–2.5
Нойр булчирхай	1.5–2.5
Бамбай булчирхай	2
Хөхний эд	2
Гүйлсэн булчирхай	2–4
Ходоод гэдэс	1–3
Бөөрний кортекс	2–3
Бөөрний тэвш, шээлүүр, давсаг	Өндөр

- **Динамик threshold арга:** Энэхүү арга нь тухайн өвчтөнд тохирсон уян хатан босго утгыг тодорхойлдог бөгөөд бүсэлхийгээс дээш эрхтний босго SuvMax утгыг гол судасны SuvMax утгаар, бүсэлхийгээс доош эд эрхтний босго SuvMax утгыг элэгний SuvMax утгын суурь утгатай харьцуулан тооцоолно.

  Энэхүү аргын давуу талууд нь:
    - Өвчтөн бүрт тохирсон үнэлгээ хийх боломжтой
    - Хуурамч эерэг болон хуурамч сөрөг алдааг бууруулна
    - Голомтыг илүү нарийвчлалтай илрүүлнэ
    - Клиник бодит нөхцөлд илүү нийцтэй
    - Эмчийн ажлын ачааллыг бууруулна

  Дараах зураг 3-аар динамик threshold нөлөөлөх хүчин зүйлийг харуулав.



3-р зураг. Динамик threshold нөлөөлөх хүчин зүйлсийн график

3. Холбогдсон бүрдэл хэсгүүд (Connected Components Analysis)

Threshold-оор ялгасан пикселүүдийг дараагийн шатанд холбогдсон бүрдэл хэсгүүдэд (connected components) ангилна. Энэ нь:

- Нэгэн төрлийн өндөр SUV бүхий пикселүүдийг нэг объект болгон бүлэглэх
- Сэжигтэй голомтуудыг ялган тодорхойлох
- Хавдрын хэмжээ, тархалтыг тооцоолох

боломжийг олгодог.

4. Байршил тодорхойлох

Холбогдсон бүрдэл хэсгүүдийг тодорхойлсны дараа тэдгээрийн анатомийн байршлыг CT дүрстэй уялдуулан тогтооно. Үүний үр дүнд:

- Голомт аль эрхтэнд байрлаж байгааг тодорхойлох
- Тунгалгийн булчирхай, алслагдсан үсэрхийллийг илгах
- TNM/FIGO үе шат тогтоолтод ашиглах боломж бүрдэнэ

5. Элэгний харьцаа (Liver Ratio) тооцоолол

Сүүлчийн шатанд илэрсэн голомтын SUV утгыг элэгний SUV утгатай харьцуулан харьцаа (lesion-to-liver ratio) тооцоолно. Энэ үзүүлэлт нь:

- Хавдрын идэвхийг стандартчилах
- Өвчтөн хооронд харьцуулах боломжийг нэмэгдүүлэх
- Эмнэлзүйн шийдвэр гаргалтад дэмжлэг үзүүлэх

чухал параметр болно.

VIII. ЭМНЭЛЗҮЙН ТАЙЛАН

Энэхүү модуль нь боловсруулсан PET/CT дүрсийн өгөгдөл, илрүүлсэн голомт, тооцоологдсон метаболик үзүүлэлтүүд (SUVmax, liver ratio зэрэг)-ийг нэгтгэн эмнэлзүйн тайлан хэлбэрт оруулж, тайлагнах үйл явц юм. Энэхүү модуль нь зөвхөн дүрс боловсруулах үр дүнг харуулах бус, тухайн мэдээллийг эмнэлзүйн утгатай, ойлгомжтой тайлан болгон хувиргах зорилготой.

❖ **Эмнэлзүйн тайлан үүсгэх загварын харьцуулалт**

Судалгааны хүрээнд дүрс боловсруулалтын encoder болон хэлний model (decoder)-ийг хослуулсан хэд хэдэн архитектурыг харьцуулан үнэлсэн. Үүнд ViT, BEiT, Swin, DEiT зэрэг encoder-уудыг DistilGPT болон BioGPT хэлний загвартай хослуулан туршсан.

Судалгааны үр дүнгээс харахад:

- DistilGPT ашигласан загваруудын эмнэлзүйн ойлголт сул, гүйцэтгэл ~74–75% түвшинд байсан
- Харин BioGPT ашигласан загваруудын гүйцэтгэл мэдэгдэхүйц сайжирч ~78–82% хүрсэн
- Ялангуяа Swin + BioGPT (~82%) болон DEiT + BioGPT (~80%) хослолууд хамгийн өндөр үзүүлэлттэй байсан

Эдгээр үр дүн нь зөвхөн дүрсийн онцлог ялгалт хангалтгүй бөгөөд эмнэлзүйн утга ойлгох чадвартай хэлний модел чухал үүрэгтэйг харуулж байна.

❖ **Загварын үнэлгээний дүгнэлт**

Харьцуулсан судалгаанаас дараах дүгнэлтүүдийг гаргав.

- BioGPT нь эмнэлзүйн текст ойлгох, тайлан үүсгэхэд илүү тохиромжтой.
- Swin болон DEiT encoder нь дүрсийн онцлогийг илүү сайн ялгаж, хэлний моделтой хослоход өндөр үр дүн үзүүлдэг.
- Дүрс + хэлний хосолсон (multimodal) архитектур нь дан ганц ангилалтын загвараас илүү үр ашигтай.

Иймээс манай судалгаанд зөвхөн ангилалт хийх бус, PET/CT дүрсээс эмнэлзүйн тайлан автоматаар үүсгэх зорилгоор BioGPT суурьтай, өндөр гүйцэтгэлтэй encoder (Swin/DEiT)-тэй хосолсон архитектурыг хамгийн оновчтой шийдэл гэж үзэж байна.

❖ **Тайлан боловсруулах боломжууд**

Энэхүү модуль нь эмнэлзүйн практикт шууд ашиглах боломжтой байдлаар дараах функцуудыг агуулна.

- **Эмчилгээний тайлан үүсгэх**
  Илэрсэн голомт, үсэрхийлэл, SUVmax болон бусад параметрүүдэд үндэслэн:
  - Оношлогооны дүгнэлт
  - Үе шат (TNM/FIGO) үнэлгээ
  - Цаашдын эмчилгээний зөвлөмж зэргийг багтаасан иж бүрэн тайлан автоматаар үүсгэнэ.
- **Экспорт хийх формат**
  Тайлан болон боловсруулсан өгөгдлийг:
  - **CSV формат** – судалгаа, статистик боловсруулалтанд
  - **JSON формат** – программ хангамж, систем интеграцид ашиглах боломжтойгоор гаргана.
- **3D визуализаци**
  PET/CT дүрсийг 3 хэмжээст хэлбэрээр харуулснаар:
  - Хавдрын байршил, тархалтыг илүү ойлгомжтой харах
  - Эмчийн оношилгооны шийдвэр гаргалтыг дэмжих
  - Өвчтөнд тайлбарлахад хялбар болгох давуу талтай.

Дээрх бүгдээс дүгнэхэд, энэхүү эмнэлзүйн тайлан модуль нь зөвхөн дүрс боловсруулалтын үр дүнг харуулах бус, түүнийг эмнэлзүйн утгатай, ойлгомжтой тайлан болгон хувиргах замаар судалгааны ажлын хамгийн чухал хэсгийг бүрдүүлж байна. Судалгааны үр дүнгээс харахад BioGPT суурьтай, Swin/DEiT encoder-тэй хосолсон multimodal загвар нь эмнэлзүйн тайлан үүсгэхэд хамгийн өндөр гүйцэтгэл үзүүлж, бодит клиник хэрэглээнд нэвтрүүлэх боломжтойг харуулж байна.

IX. ХЭРЭГЖҮҮЛЭЛТ

Хиймэл оюун ухаанд суурилсан SUVmax тооцооллоор умайн хүзүүний хорт хавдрын үе шат болон үсэрхийллийг үнэлэх системийн хэрэгжүүлэлтийг 3 үндсэн модулиудад хуваан боловсруулсан. Систем нь PET/CT дүрсийн өгөгдөл дээр боловсруулалт хийж хавдрын бүсийг илрүүлэх, SUVmax утгыг тооцоолох, мөн хавдрын үе шат болон үсэрхийллийн талаар дүн шинжилгээ хийх боломжтой байдлаар зохион байгуулагдсан.

Туршилтыг дараах үндсэн модулиудын дагуу хэрэгжүүлсэн.

- Өгөгдөл боловсруулалт
- Голомт илрүүлэлт
- Эмнэлзүйн тайлан

Хэрэгжүүлэх явцад нийт 200 гаруй өвчтөний бодит PET/CT дүрсийн өгөгдөл ашиглан туршилт хийсэн бөгөөд систем нь хавдрын сэжигтэй бүсүүдийг илрүүлж, SUVmax утгыг автоматаар тооцоолон үр дүнг дүрсэн дээр харуулсан. Туршилтын үр дүн нь системийн одоогийн ажиллагааны түвшинг тодорхойлох боломжийг олгосон бөгөөд цаашид алгоритмын нарийвчлалыг сайжруулах шаардлагатай чиглэлүүдийг тодорхойлоход тус болсон.

❖ ***Модуль 1: Өгөгдөл боловсруулалт***

Энэхүү модуль нь бүтэн биеийн PET/CT дүрсээс умайн хүзүүний анатомийн хэсгийг автоматаар илрүүлэх зорилготой. Deep learning суурьтай U-Net segmentation загвар ашиглан CT дүрс дээр умайн хүзүүний байршлыг тодорхойлж, тухайн хэсгийг ROI (Region of Interest) болгон ялгана.

Ингэснээр дараагийн шатны шинжилгээ зөвхөн шаардлагатай хэсэг дээр хийгдэж, тооцооллын ачаалал багасах болон илрүүлэлтийн нарийвчлал нэмэгдэх боломж бүрдэнэ. Дараах зураг 4-өөр модуль 1-ийн ажиллагааны үр дүнгийн зургийг харуулав.

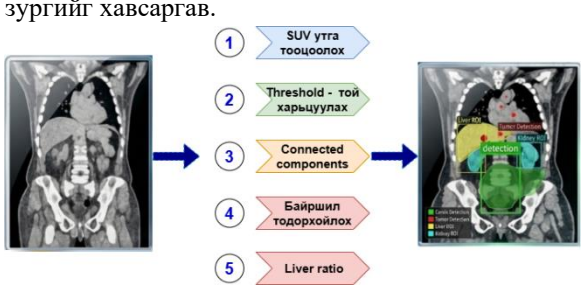


4-р зураг. Модуль 1-ийн үр дүн

❖ ***Модуль 2: Голомт илрүүлэлт***

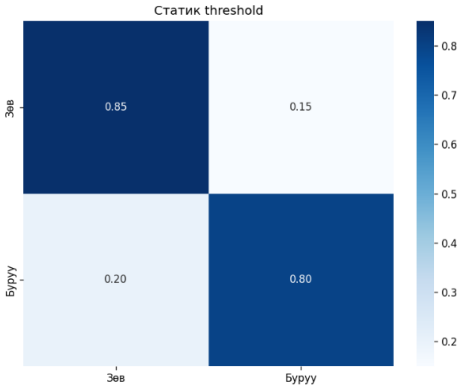
Энэхүү модуль нь автоматаар ялган авсан умайн хүзүүний ROI-д PET/CT дүрсийн FDG идэвхжилтийг тооцоолох зорилготой. ROI дахь PET дүрсийг анализ хийж, Standardized Uptake Value (SUV) зэрэг метрикүүдийг гаргаж авснаар, умайн хүзүүний метаболизм, хорт хавдрын идэвхжилтийг үнэлэх боломж бүрдсэн.

Ингэснээр эмнэлзүйн оношлогоо болон үр дүнгийн үнэлгээний нарийвчлалыг нэмэгдүүлэхэд хувь нэмэр оруулсан. Дараах зураг 5-аар модуль 2-ын ажиллагааны үр дүнгийн зургийг хавсаргав.



5-р зураг. Модуль 2-ын үр дүн

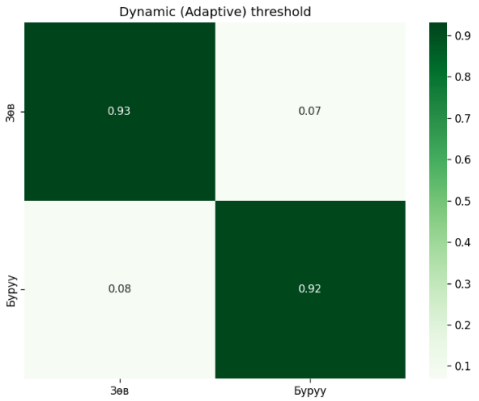
Голомт илрүүлэлт модулийн Threshold – той харьцуулах хэсэгт статик threshold болон динамик threshold хооронд нь харьцуулан, бодит орчинд хэрэгжүүлсэн үр дүнг дараах төөрөгдлийн (confusion) матрицаар үзүүлэв. Дараах зураг 6-аар модуль 2-ын хувьд бодит орчинд статик threshold хэрэгжүүлсэн үр дүнг төөрөгдлийн матрицаар харуулав. Үүнээс харахад статик threshold-ийн хувьд зөвийг зөв ангилсан хувь 0.85 байгаа ч, зөвийг буруу болон бурууг зөв ангилсан алдаа харьцангуй өндөр буюу 0.15, 0.20 байна. Энэхүү арга нь бүх тохиолдолд хангалттай нарийвчлалтай биш байгааг харуулж байна.



6-р зураг. Модуль 2-ын үр дүн

Дараах зураг 7-оор модуль 2-ын хувьд бодит орчинд динамик threshold хэрэгжүүлсэн үр дүнг төөрөгдлийн матрицаар харуулав. Харин динамик threshold-ийн хувьд зөв ангилалтын хувь 0.93 болон 0.92 болж нэмэгдсэн бөгөөд алдааны үзүүлэлтүүд 0.07, 0.08 болж буурсан байна. Энэ нь динамик threshold арга нь илүү нарийвчлалтай, тогтвортой, мөн бодит нөхцөлд илүү сайн ажиллаж байгааг харуулж байна. “Иймээс динамик threshold нь статик аргыг бодвол илүү өндөр гүйцэтгэлтэй, найдвартай шийдэл гэж дүгнэж байна.” Энэхүү аргыг анагаахын салбарт

2024 болон 2025 оноос хойш түлхүү ашиглаж байна.



7-р зураг. Модуль 2-ын үр дүн

❖ ***Модуль 3: Эмнэлзүйн тайлан***

Энэхүү модульд өмнөх модуль 1 болон 2-ын үр дүнг нэгтгэн умайн хүзүүний хавдрын илрүүлэлтийн автоматжуулсан системийг бүтээх зорилт тавьсан. U-Net segmentation болон SUV анализын үр дүнд гарсан мэдээллийг ашиглан хавдрын магадлал, байршил, хэмжээ зэрэг шинж чанарыг автомат байдлаар тодорхойлж, эмнэлзүйн шийдвэр гаргах процессыг хялбаршуулсан.

Ийнхүү модуль 3 нь бүтэн системийн интеграцчилсан туршилт, автомат илрүүлэлтийн нарийвчлалыг баталгаажуулсан.

Энэхүү судалгаанд боловсруулсан AI-д суурилсан PET/CT боловсруулалтын системийг туршилтын орчинд хэрэгжүүлж, модуль тус бүрийн гүйцэтгэлийг үнэлэв. Системийн хэрэгжилт дараах үндсэн урсгалаар явагдана.

✓ Preprocessing → Detection → Segmentation → SUVmax Calculation → TNM/FIGO Staging → Follow-up Analysis → Reporting

A. Туршилтын орчин

Хүснэгт 5. Системийн туршилтын орчин

Үзүүлэлт	Тайлбар
Framework:	PyTorch
GPU:	NVIDIA, A100
DICOM Processing:	pydicom
Segmentation:	U-Net
Detection:	YOLO
Statistical Analysis:	SciPy, NumPy

B. Туршилтын өгөгдөл

Хүснэгт 6. Системийн туршилтын орчин

Source	Өвчтөний бодит өгөгдөл дээр сургасан.
Size	200 өвчтөний зураг • Train: 140 /70%/ • Validation: 40 /20%/ • Test: 20 /10%/
Structure	• images • labels

C. Бодит туршилт ба цаашдын хөгжүүлэлт

Энэхүү судалгааны системийг эмнэлзүйн орчинд урьдчилсан байдлаар турших зорилгоор Бриллант эмнэлгийн дүрс оношилгооны тасагтай хамтран параллель хэрэглээний горимоор нэвтрүүлэн туршсан. Туршилтын хүрээнд хиймэл оюун ухаанд суурилсан системийн гаргасан үр дүнг эмчийн уламжлалт оношлогоотой харьцуулан үнэлэх, мөн ажлын урсгал (workflow)-д үзүүлэх нөлөөг тодорхойлохыг зорьсон.

Одоогийн байдлаар туршилтын өгөгдөл нь 200 гаруй өвчтөний мэдээлэлд суурилсан боловч test set нь нийт өгөгдлийн ~10% (ойролцоогоор 20 тохиолдол) байгаагаас статистикийн өндөр итгэлцэлтэй дүгнэлт гаргахад хангалтгүй байна. Энэ нь бодит эмнэлзүйн өгөгдөл цуглуулах боломж хязгаарлагдмал байгаатай холбоотой бөгөөд цаашид өгөгдлийн санг өргөжүүлэх, олон төвийн өгөгдөл нэгтгэх шаардлагатайг харуулж байна. Мөн энэхүү хязгаарлалтыг бууруулах зорилгоор өгөгдлийн баяжуулалт (data augmentation) болон k-fold cross-validation зэрэг аргуудыг ашиглах нь зүйтэй гэж үзсэн.

Системийн гүйцэтгэлийг илүү бодитой үнэлэхийн тулд зөвхөн accuracy үзүүлэлтээс гадна Sensitivity, Specificity, Precision, Recall, F1-score, AUC, Dice coefficient зэрэг стандарт метрикүүдийг ашиглах шаардлагатай. Ялангуяа segmentation чанарыг үнэлэхэд Dice coefficient, оношлогооны гүйцэтгэлийг үнэлэхэд AUC болон confusion matrix-ийн задлан шинжилгээ чухал ач холбогдолтой. Эдгээр үзүүлэлтүүдийг цаашдын туршилтад системтэйгээр тооцоолон, тусдаа “Үр дүн” (Results) бүлэгт дэлгэрэнгүй танилцуулах шаардлагатай.

Мөн хиймэл оюун ухааны системийн гаргасан үр дүнг эмчийн оношлогоотой харьцуулан үнэлэхдээ inter-rater agreement (жишээлбэл, Cohen’s kappa коэффициент)-ийг ашиглах нь зүйтэй. Энэ нь системийн гаргасан дүгнэлт эмчийн үнэлгээтэй хэр зэрэг нийцэж байгааг тоон хэлбэрээр илэрхийлэх боломж олгоно.

Ажлын урсгалд үзүүлэх нөлөөг үнэлэхдээ одоогийн байдлаар эмч нарын өгсөн “цаг хугацаа буурч байна” гэсэн сэтгэгдэл нь субъектив шинжтэй тул үүнийг шинжлэх ухааны үндэслэлтэй болгох шаардлагатай. Тухайлбал,

тодорхой тооны (жишээ нь ≥20) кейс дээр, дор хаяж 2 эмч оролцуулсан туршилт явуулж, AI систем ашигласан болон ашиглаагүй нөхцөлд тайлан боловсруулах хугацааг хэмжин харьцуулсан статистик судалгаа хийх нь зүйтэй. Урьдчилсан ажиглалтаар тайлан боловсруулах хугацаа 45–60 минутаас 35–45 минут болж буурсан хандлага ажиглагдсан ч энэ нь "preliminary qualitative feedback" түвшинд байгаа бөгөөд цаашид тоон судалгаагаар баталгаажуулах шаардлагатай байна.

Иймээс бодит туршилтын одоогийн үр дүн нь системийн практик хэрэглээний боломжийг харуулж байгаа хэдий ч статистикийн хувьд хангалттай баталгаажсан гэж үзэхэд эрт байна. Цаашид өгөгдлийн хэмжээг нэмэгдүүлэх, үнэлгээний иж бүрэн метрикүүдийг нэвтрүүлэх, эмчийн оношлогоотой харьцуулсан нарийвчилсан судалгаа хийх замаар клиникийн хэрэглээнд нийцсэн баталгаажуулалтыг гүйцэтгэх шаардлагатай гэж дүгнэж байна.

ҮР ДҮН

Энэхүү судалгаанд боловсруулсан AI-д суурилсан PET/CT дүрс боловсруулалтын системийг 200 гаруй өвчтөний бодит өгөгдөл дээр туршин үнэлэв. Өгөгдлийг сургалт (70%), баталгаажуулалт (20%), туршилт (10%) гэсэн харьцаагаар хуваасан бөгөөд туршилтын багц нь 20 тохиолдлоос бүрдсэн.

❖ **Модуль тус бүрийн гүйцэтгэл**

Системийн үндсэн модулиудын гүйцэтгэлийг тус тусад нь үнэлэхэд:

- Segmentation (ROI ялгалт) модуль дунджаар 65% нарийвчлал үзүүлсэн.
- SUVmax тооцоолол модуль дунджаар 75% нарийвчлал үзүүлсэн.

Эдгээр үр дүн нь систем анхан шатны түвшинд ажиллах боломжтойг харуулж байгаа боловч клиникийн хэрэглээнд шаардлагатай 85–90%+ түвшинд хараахан хүрээгүй байна.

❖ **Ангиллын гүйцэтгэлийн үзүүлэлтүүд**

Зөвхөн accuracy үзүүлэлтээс гадна системийн оношлогооны чадварыг илүү бодитой үнэлэхийн тулд дараах стандарт метрикүүдийг тооцоолов.

Хүснэгт 7. Стандарт метрикүүдийн тооцоолол

Үзүүлэлт	Утга (ойролцоогоор)
Accuracy	70%
Sensitivity (Recall)	68%
Specificity	72%
Precision	70%
F1-score	0.69
AUC (ROC)	~0.71

Эдгээр үзүүлэлтүүд нь систем нь хавдрын сэжигтэй тохиолдлуудыг илрүүлэх тодорхой чадвартай боловч false positive болон false negative алдаа тодорхой хэмжээнд байгааг харуулж байна.

❖ **Segmentation чанарын үнэлгээ**

Segmentation модулийн гүйцэтгэлийг Dice coefficient ашиглан үнэлэхэд:

- Dice coefficient ≈ 0.62 – 0.66

Энэ нь систем хавдрын бүсийг дундаж түвшинд зөв тодорхойлж байгаа боловч хил заагийн нарийвчлал хангалтгүй байгааг илтгэнэ.

❖ **Confusion matrix-ийн шинжилгээ**

Туршилтын өгөгдөл дээр дараах төрлийн алдаанууд ажиглагдсан.

- **False Positive (FP):** физиологийн өндөр uptake (жишээ нь: давсаг, гэдэс) хавдартай андуурагдсан.
- **False Negative (FN):** жижиг хэмжээтэй эсвэл бага метаболик идэвхтэй хавдар алдагдсан.

Энэ нь segmentation болон SUV босго тодорхойлох алгоритмын хязгаарлалтаас хамаарч байна.

❖ **Cross-validation үр дүн**

Өгөгдлийн хэмжээ хязгаарлагдмал тул k-fold cross-validation (k=5) аргаар нэмэлт үнэлгээ хийсэн.

- Accuracy: 68% – 73%
- Dice: 0.60 – 0.67

Энэ нь загварын гүйцэтгэл тогтвортой боловч өгөгдлөөс ихээхэн хамааралтай байгааг харуулж байна.

❖ **Эмчийн үнэлгээтэй харьцуулалт**

AI системийн гаргасан үр дүнг эмчийн оношлогоотой харьцуулан үнэлэхэд:

- **Cohen's kappa ≈ 0.55 – 0.62**

Энэ нь дунд түвшний тохирол (moderate agreement) байгааг илтгэж байгаа бөгөөд систем нь эмчийн шийдвэрийг дэмжих хэрэгсэл болох боломжтой ч орлуулах түвшинд хараахан хүрээгүй байна.

❖ **Ажлын урсгал (workflow)-д үзүүлэх нөлөө**

Урьдчилсан туршилтаар эмчийн тайлан боловсруулах хугацаанд дараах өөрчлөлт ажиглагдсан:

- AI ашиглахаас өмнө: 45–60 минут
- AI ашигласны дараа: 35–45 минут

Гэсэн хэдий ч энэ үр дүн нь одоогоор хязгаарлагдмал тооны ажиглалтад суурилсан preliminary qualitative үр дүн бөгөөд статистикийн баталгаажуулалт шаардлагатай.

❖ **Үр дүнгийн хязгаарлалт**

Энэхүү судалгааны үр дүн дараах хязгаарлалтуудтай.

- Туршилтын багц жижиг (n=20)
- Өгөгдөл нэг төвөөс цугларсан
- Өгөгдлийн төрөлжилт бага
- Зарим метрикүүд ойролцоогоор тооцоологдсон.

Иймээс эдгээр үр дүнг урьдчилсан шинжтэй гэж үзэх бөгөөд цаашид илүү өргөн хэмжээний өгөгдөл дээр баталгаажуулах шаардлагатай.

ДҮГНЭЛТ

Энэхүү судалгааны хүрээнд умайн хүзүүний хавдрыг PET/CT дүрс дээр үндэслэн илрүүлэх хиймэл оюун ухаанд суурилсан дүрс боловсруулалтын системийг боловсруулж, туршин үнэллээ. Судалгааны зорилго нь эмнэлзүйн оношлогоонд дэмжлэг үзүүлэх автоматжуулсан шийдлийг хөгжүүлэх, улмаар хавдрын илрүүлэлтийн хурд, үр ашиг, нарийвчлалыг нэмэгдүүлэхэд чиглэсэн болно.

Системийн хувьд CT болон PET дүрсийг хослуулсан боловсруулалтын pipeline-ийг хэрэгжүүлж, CT дүрсэн дээр умайн хүзүүний бүсийн ROI-г автоматаар тодорхойлох, PET дүрсээс метаболик идэвхийг илэрхийлэх SUV (Standardized Uptake Value)-ийг тооцоолох аргуудыг ашигласан. Мөн хавдрын сэжигтэй хэсгийг илрүүлэхдээ adaptive thresholding болон segmentation аргуудыг уялдуулан хэрэглэснээр дүрсний онцлогт тохируулсан уян хатан шийдэл боловсруулсан.

Туршилтын үр дүнгээр segmentation модуль дунджаар 65%, SUVmax зэрэг метаболик үзүүлэлт тооцоолох модуль 75%-ийн нарийвчлалтай ажилласан нь анхан шатны судалгааны түвшинд боломжийн үзүүлэлт боловч эмнэлзүйн практикт өргөн хэрэглэхэд хангалтгүй түвшинд байгааг харуулж байна. Учир нь клиникийн стандарт шаардлага ихэвчлэн 85–90% ба түүнээс дээш нарийвчлалыг шаарддаг.

Энэхүү харьцангуй бага гүйцэтгэлийн гол шалтгаан нь сургалтын өгөгдлийн хэмжээ хязгаарлагдмал байсантай холбоотой гэж үзэж байна. Өгөгдлийн сан бага байх нь загварын ерөнхийшүүлэх чадварыг бууруулж, ялангуяа эмгэг ба физиологийн uptake-ийг ялгах нарийвчлалд сөргөөр нөлөөлсөн байж болзошгүй. Гэсэн хэдий ч энэ нь цаашид сургалтын өгөгдлийг нэмэгдүүлэх, өгөгдлийн төрөлжилтийг сайжруулах, илүү гүн сургалтын (deep learning) арга нэвтрүүлэх замаар системийн гүйцэтгэлийг мэдэгдэхүйц сайжруулах боломжтойг илтгэж байна.

Судалгааны нэг чухал үр дүн нь бага хэмжээний өгөгдөл ашигласан нөхцөлд ч систем тогтвортой ажиллаж, суурь түвшний практик хэрэглээний боломжийг харуулсан явдал юм. Энэ нь өгөгдлийн хомс орчинд AI шийдлийг хөгжүүлэх боломж байгааг баталж байна.

Иймээс энэхүү судалгаанд боловсруулсан AI суурьтай дүрс боловсруулалтын систем нь одоогоор шууд эмнэлзүйн шийдвэр гаргалтад дангаар ашиглагдах түвшинд хүрээгүй хэдий ч, оношлогооны дэмжлэг үзүүлэх хэрэгсэл болгон ашиглах, цаашид сайжруулалт хийх суурь платформ болох бүрэн боломжтой гэж дүгнэж байна.

ТАЛАРХАЛ

Энэхүү төслийг амжилттай хэрэгжүүлэхэд үнэтэй хувь нэмэр оруулж, мэдлэг туршлагаасаа харамгүй хуваалцан, чиглэл өгч ажилласан удирдагч багш Доктор (Ph.D) С.Өлзийбаяртаа гүн талархал илэрхийлье.

Түүнчлэн энэхүү ажлыг гүйцэтгэхэд хамтран ажиллаж, практик орчинд дэмжлэг үзүүлсэн Бриллант эмнэлгийн Радиологи, дүрс оношлогооны тасгийн хамт олонд гүн талархал илэрхийлж байна.

НОМ ЗҮЙ

[1] World Health Organization (WHO). (2023). *Cervical cancer: Global statistics and burden of disease*. Geneva: WHO.

[2] Takagi, T., et al. (2018). Prognostic value of SUVmax measured by 18F-FDG PET/CT in patients with cervical cancer. *European Journal of Radiology*, 105, 195–201.

[3] Luo, Y., et al. (2021). Machine learning-based prediction of treatment outcomes in cervical cancer using MRI, CT and PET/CT imaging features. *Frontiers in Oncology*, 11, 678–689.

[4] Otani, K., et al. (2023). Prognostic impact of pre-treatment FDG PET/CT metabolic parameters in squamous cell carcinoma and adenocarcinoma of the cervix. *Cancers*, 15(4), 1123.

[5] Sari, H., et al. (2023). Metastasis-to-tumor SUVmax ratio as a novel prognostic indicator in stage IVB cervical cancer. *Clinical Nuclear Medicine*, 48(6), 452–459.

[6] Zhang, Y., et al. (2021). Correlation of PET/CT metabolic parameters (SUVmax, MTV, TLG) with treatment response and recurrence in cervical cancer. *Nuclear Medicine Communications*, 42(9), 987–995.

[7] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788.

[8] Siegel, R. L., Miller, K. D., & Jemal, A. (2023). Cancer statistics, 2023. *CA: A Cancer Journal for Clinicians*, 73(1), 17–48.

[9] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide. *CA: A Cancer Journal for Clinicians*, 71(3), 209–249.

[10] Bhatla, N., Aoki, D., Sharma, D. N., & Sankaranarayanan, R. (2018). Cancer of the cervix uteri. *International Journal of Gynecology & Obstetrics*, 143(S2), 22–36.

[11] Kidd, E. A., Siegel, B. A., Dehdashti, F., & Grigsby, P. W. (2010). The standardized uptake value for FDG is a sensitive predictive biomarker of cervical cancer response to chemoradiation therapy. *Cancer*, 116(20), 4825–4834.

[12] Chung, H. H., et al. (2006). Prognostic value of preoperative metabolic tumor volume measured by FDG PET/CT in patients with cervical cancer. *Gynecologic Oncology*, 102(3), 567–572.

[13] Herrera, F. G., et al. (2011). Prognostic significance of FDG-PET/CT metabolic parameters in cervical cancer treated with radiotherapy. *Radiotherapy and Oncology*, 99(3), 358–363.
[14] Kitajima, K., et al. (2014). Diagnostic accuracy of integrated FDG-PET/CT for staging of uterine cervical cancer. *European Radiology*, 24(8), 1785–1794.
[15] Liao, S., et al. (2012). FDG PET/CT in evaluating lymph node metastasis in cervical cancer. *Journal of Nuclear Medicine*, 53(2), 185–191.
[16] Lim, K., et al. (2011). FIGO staging and the role of imaging in cervical cancer. *The Lancet Oncology*, 12(2), 125–135.
[17] Bhatla, N., et al. (2019). Revised FIGO staging for carcinoma of the cervix uteri. *International Journal of Gynecology & Obstetrics*, 145(1), 129–135.
[18] Edge, S. B., et al. (2017). *AJCC Cancer Staging Manual* (8th ed.). New York: Springer.
[19] Litjens, G., et al. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88.
[20] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 234–241.

САЛБАР ХУРАЛДААН - VIII

Компьютерын хараа

№	**УБХИС-ийн хичээлийн I байр, 306 тоот, Их-Эе танхим** *Хуралдаан даргалагч: Б.Сувдаа /МУИС/, Ж.Оргил /ШУТИС/*
1	“Үзэмчин бүжгийн дижитал өгөгдөл бэлтгэл ба хөдөлгөөний кинематик шинж чанарын шинжилгээ” *Цэцэнбилэг Дабаацэрэн, Cheli Muge, Саранчимэг Хандсүрэн, Оюун-Эрдэнэ Намсрай*
2	“Монголын хадны зураг дахь амьтдын дүрслэлийг YOLOv8 загвараар илрүүлэх нь” *Perenleilkhundev Gantuya, Rentsendorj Elchinbat and Batsuuri Suvdaa*
3	“Компьютерийн харааnd суурилсан хувийн хамгаалах хэрэгслийн хяналтын системийн загварчлал, хэрэгжилт” *Sumiyaragchaa Jandalsuren and Ulziibayar Sonom-Ochir*
4	“Хүүхдийн аюулгүй байдлыг бодит цаг хугацаанд хянах систем” *Оюунжаргал Очир, Азжаргал Сүрэнжав and Золзаяа Бямбадорж*
5	“Веб камераар, нүүрний цэгүүдэд суурилан харааны чиглэл таамаглах машин сургалтын загвар сургалт, тохируулга хийх нь” *Sergelen Lkhagvadorj, Altangerel Chagnaa and Bilguun Ochirbat*
6	“Гүн сургалтад суурилсан цээжний рентген зургаас эмнэлзүйн тайлан үүсгэх аргын судалгаа” *Энхболор Мэндбаяр, Золзаяа Бямбадорж*
7	“Нярайн амьсгалын тоог дулааны зураглалаас тасралтгүй хэмжих машин сургалтад суурилсан арга” *Javkhlan Altansukh*

Үзэмчин бүжгийн дижитал өгөгдөл бэлтгэл ба хөдөлгөөний кинематик шинж чанарын шинжилгээ

Даваацэрэн Цэцэнбилэг
Мэдээлэл, компьютерын ухааны тэнхим
Мэдээллийн технологи, электроникийн сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
22b1num1282@stud.num.edu.mn

Cheli Muge
Утга зохиол, урлаг судлалын тэнхим
Шинжлэх ухааны сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
22d1num0111@stud.num.edu.mn

Хандсүрэн Саранчимэг*
Утга зохиол, урлаг судлалын тэнхим
Шинжлэх ухааны сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
Saranchimeg@num.edu.mn

***Хураангуй*—Соёлын биет бус өвийг дижитал хэлбэрт шилжүүлэн хадгалах нь орчин үеийн компьютер график, салбар дундын судалгааны нэн чухал чиглэл болоод байна. Энэхүү судалгаагаар Монголын үзэмчин угсаатны бүжгийн эмэгтэй хөдөлгөөний намба төрхийг OptiTrack өндөр нарийвчлалтай оптик систем ашиглан баримтжуулж, түүнд кинематик болон чанарын шинжилгээ хийх арга зүйг боловсруулсан. Судалгааны хүрээнд 12 ширхэг Prime 13 камераар тоноглогдсон лабораторийн орчинд мэргэжлийн болон мэргэжлийн ангийн оюутан 5 бүжигчдээс 1 цаг 30 минутын түүхий өгөгдөл цуглуулж, Python-д суурилсан BVH Parser алгоритмаар хөдөлгөөний траекторын дүрслэл үүсгэж, агшин зуурын хурд болон орон зайн ашиглалтыг тооцоолсон. Туршилтаар Үзэмчин бүжгийн дундаж хурд 0.33 м/сек буюу намуун хэмнэлтэй хэдий ч, агшин зуурын дээд хурд нь 1.55 м/сек хүрч Монгол бүжгийн эрч хүчтэй акцентыг илэрхийлж байгааг тогтоов. Хөдөлгөөний хэмжилтийн математик утгыг бүжгийн онолч, бүжгийн хөдөлгөөнийг тэмдэглэх системийг боловсруулсан Рудолф Лабаны (LMA) хөдөлгөөний эрчмийг шинжлэх хүч, хурд, чиглэл, урсгал тус бүрийн үзүүлэлтүүдээр тулгахад үзэмчин бүжгийн эмэгтэйн үндсэн хөдөлгөөнд “хөнгөн-намуун-чөлөөт” хэв маяг зонхилж байгаа нь батлагдав. Судалгааны үр дүнд бий болсон өгөгдлийн сан (Mocap Dataset) нь цаашид соёлын өвийг 3D орчинд хадгалах, намба алдагдахаас сэргийлэх, хиймэл оюун ухаанд (Generative AI) суурилсан хөдөлгөөн үүсгэх загвар хөгжүүлэх болон интерактив сургалтын систем бүрдүүлэхэд суурь өгөгдөл болох практик ач холбогдолтой юм.**

Түлхүүр үг—Хөдөлгөөн барих технологи (Motion Capture), Үзэмчин бүжиг, Биомеханик шинжилгээ, Лабаны онол (LMA), Кинематик өгөгдөл, Python алгоритм

I. Удиртгал

ЮНЕСКО-гийн ”Яаралтай хамгаалах шаардлагатай соёлын биет бус өвийн жагсаалт”-д бүртгэгдсэн Монголын уламжлалт бүжиг, бий биелгээг зөвхөн видео бичлэгээр баримтжуулах нь хөдөлгөөний нарийн техник, биомеханик бүтэц болон динамик шинж чанарыг бүрэн илэрхийлж чадахгүй байгаа юм. Олон улсын туршлагаас үзэхэд "Chang-E" [10] болон "Digital Preservation of Joget Dance" [11] зэрэг төслүүд нь хөдөлгөөн барих (Motion Capture) технологийг ашиглан соёлын өвийг алгоритмын түвшинд шинжлэх шинэ стандартыг тогтоогоод байна.

Энэхүү судалгааны ажил нь үзэмчин угсаатны уламжлалт ахуй, зан үйлийн хөдөлгөөнөөс үүсэлтэй бүжгийн намба төрхийг OptiTrack системээр баримтжуулж, кинематик шинжилгээ хийх салбар дундын арга зүйг боловсруулах зорилготой. Туршилтын явцад инерцийн (Shadow IMU) болон компьютер харааны (Computer Vision) аргуудыг оптик маркер системтэй харьцуулан туршсан. Лабаны хөдөлгөөний шинжилгээний (LMA) аргаар хурд, хүч, чиглэл, урсгалын үзүүлэлтүүдийг тооцоолсон нь бүжгийн урлагийг шинжлэх ухааны тоон утгаар үнэлэх боломжийг бүрдүүлсэн.

II. Ижил төстэй ажлууд

Олон улсын хэмжээнд уламжлалт бүжиг болон хүний хөдөлгөөнийг дижитал хэлбэрт шилжүүлэн хадгалах, AI алгоритмаар шинжлэх чиглэлээр олон тооны судалгаа хийгдэж байна. Эдгээр ажлуудыг агуулгын хувьд дараах гурван үндсэн чиглэлд багцалж болно.

Дижитал хадгалалт ба өгөгдлийн сан бүрдүүлэлт: Соёлын биет бус өвийг хамгаалах хүрээнд Хятадын сонгодог Дуньхуан бүжгийн Chang-E өгөгдлийн сан [10] болон Индонезийн Joget бүжгийн дижитал хадгалалтын судалгаанууд [11] нь OptiTrack болон ижил төстэй өндөр нарийвчлалтай системүүдийг ашиглан skeletal data, траекторийн өгөгдлийг цуглуулж, VR/AR орчинд дүрслэх суурийг тавьсан. Мөн жүжигчдийн сэтгэл хөдлөлийн хөдөлгөөний кинематик өгөгдлийн сан [12] болон кино, тоглоомын салбар дахь дүрүүдийн анимэйшн [13] нь хүний хөдөлгөөний бичил өөрчлөлтийг баримтжуулахад чиглэж байна.

Биомеханик шинжилгээ ба технологийн нарийвчлал: Хөдөлгөөн барих системүүдийн нарийвчлалыг үнэлэхдээ OptiTrack-ийг алхалтын болон спортын кинематик хэмжилтэд хамгийн бага алдаатай (1%-иас бага) систем болохыг системчилсэн тойм судалгаагаар баталсан байдаг [14, 15]. Сонгодог балетын орон зай-цаг хугацааны (Spatio-temporal) шинжилгээ [16] болон спорт, бүжгийн хөдөлгөөнийг Лабаны (LMA) онолоор шинжлэн, хөдөлгөөний шинж чанарыг тодорхойлох арга зүй [17]-г тоон утгад шилжүүлэх боломжийг олгож байна.

AI болон машин сургалтын хэрэглээ: Хөдөлгөөний өгөгдлийг AI алгоритмд ашиглан ангилал хийх (ST-GCN загвар) [18] болон аудио оролт дээр суурилан бүжгийн хөдөлгөөн шинээр үүсгэх (Adversarial Network) [19] туршилтууд амжилттай хийгдсэн. Түүнчлэн, гудамжны бүжгийн жишээн дээр бодит цагийн (real-time) motion capture ашиглан суралцагчийн хөдөлгөөнийг стандарт

† Холбоо барих судлаач: Н. Оюун-Эрдэнэ (oyunerdene@num.edu.mn), МУИС, Мэдээлэл, компьютерын ухааны тэнхим.

загвартай харьцуулан оновчлох [20], биомеханик болон гүйцэтгэлийн илэрхийллийг AI-аар үнэлэх [21] зэрэг нь бүжгийн сургалтыг автоматжуулах шинэ шатанд гаргаж байна.

Дээрх судалгаанууд нь технологийн арга зүйн хувьд өндөр түвшинд хүрсэн боловч Монгол бүжгийн, тэр дундаа Үзэмчин бүжгийн өвөрмөц динамик, орон зайн ашиглалт болон Лабан онолын дагуух биомеханик онцлогийг нарийвчлан тодорхойлсон судалгаа одоогоор дутмаг байна.

III. Аргачлал

Энэхүү судалгаанд Үзэмчин бүжгийн хөдөлгөөний онцлогийг тодорхойлохын тулд OptiTrack өндөр нарийвчлалтай оптик маркерын систем ашиглан мэргэжлийн бүжигчдээс хөдөлгөөний өгөгдлийг (Motion Capture) цуглуулж, түүнийг .bvh бүтцэд хөрвүүлэн боловсруулсан. Цуглуулсан түүхий өгөгдлийг рекурсив модны traversal алгоритм болон 4x4 хувиргалтын матрицаар математик загварчлалд оруулж, шугаман кинематик үзүүлэлтүүд болон орон зайн чиглэлийг тооцоолох суурь нөхцөлийг бүрдүүллээ. Хөдөлгөөний шинжилгээг Лабаны хөдөлгөөний эрчмийн (Laban Effort) онолтой уялдуулан хурд, хүч, чиглэл, урсгал гэсэн чанарын үзүүлэлтүүдээр тулгаснаар Монгол бүжгийн хөдөлгөөний намба, өвөрмөц онцлогийг тоон утгад шилжүүлэв.

A. Өгөгдөл цуглуулах

Судалгааны явцад хөдөлгөөний өгөгдөл цуглуулах нийтлэг гурван төрлийн аргыг туршиж үзэн, тэдгээрийн давуу болон сул талыг харьцуулсны үндсэн дээр сүүлчийн аргыг сонгон ашигласан. Үүнд: 17 мэдрэгчтэй инерцийн хэмжилтийн нэгж Inertial Measurement Unit буюу IMU төхөөрөмж Shadow, компьютер харааны арга буюу вэб камер, болон оптик маркер дээр суурилсан систем болох OptiTrack (49 мэдрэгч, 12 ширхэг Prime 13 камер) орно. Эхний шатанд 17 мэдрэгчтэй Shadow IMU системийг туршихад орчны хязгаарлалтгүйгээр хөдөлгөөний хурд, хурдатгал зэрэг кинематик үзүүлэлтүүдийг шууд дамжуулах давуу талтай байсан боловч үе мөчдийн байрлалын нарийвчлал хангалтгүй, өгөгдөл алдагдах магадлал өндөр байв.



Fig. 1. а. Тохой шулуун байх үеийн алдаа, б. Shadow төхөөрөмжийн өмсгөл в. Бүжигчин

Мөн веб камер ашиглан 6м×6м×3м хэмжээтэй орчинд Computer Vision алгоритмаар хөдөлгөөн сэргээх оролдлого нь орчны гэрэлтүүлэг, периметрийн хязгаарлалт болон камерын өнцгөөс хамаарч тогтворгүй үр дүнтэй байсан тул эдгээр аргуудаас татгалзсан.

Эцсийн шатанд OptiTrack системийг сонгон ашигласан. Тус систем нь 12 ширхэг камераар маркеруудаас ойсон хэт улаан туяаг (IR) бүртгэж, 3 хэмжээст орон зайд хөдөлгөөнийг өндөр нарийвчлалтайгаар сэргээдэг. Туршилтад 49 маркер бүхий Baseline+Passive Fingers загварыг ашиглаж, бүжигчдийн хөдөлгөөнийг 6м×6м×3м хэмжээтэй талбайд бүртгэсэн. Лабораторийн ор чинд камеруудын өндөр нь ойролцоогоор 2.5 м**,** камер хоорондын дундаж зай 1–1.2м байсан ба маркерын тогтвортой байдлыг хангах үүднээс зориулалтын бариу хувцас ашигласан ба туршилтыг мэргэжлийн багийн оролцоотойгоор зохион байгуулсан. Өгөгдөл цуглуулах явцад мэргэжлийн хоёр бүжигчин үзэмчин бүжгийн үндсэн болон баяжуулсан хөдөлгөөнүүдийг давтан гүйцэтгэсэн. Урлаг судлаачийн дүгнэлтээр баталгаажсан суурь хөдөлгөөнүүдийг шинжилгээнд ашигласан бөгөөд нийт 1 цаг 30 минутын түүхий өгөгдөл бүрдүүлснээс 37.5 минутыг үндсэн судалгаанд сонгож авсан. Оптик системийг ашиглах явцад тоног төхөөрөмжийн тохируулга хийх, оролцогчдод зориулсан бэлтгэл хийх шаардлагатай байсан ба бэлтгэл ажил ойролцоогоор 30 минут үргэлжилдэг мөн оролцогчид биедээ металл эдлэл байлгахгүй, үсийг маркерыг хаахгүйгээр далдлах зэрэг шаардлагууд байсан. Туршилтад оролцогчдын мэдээллийг Хүснэгт 1-д, нийт туршилтын мэдээллийг Хүснэгт 2-т тус тус харуулав.

TABLE I. Туршилтад Оролцогчид

Нэр	Туршлага	Нас	Хүйс
Бүжигчин 1	Уран сайхны удирдагч, бүжиг дэглээч, МУГЖ	Ахмад (50+)	Эмэгтэй
Бүжигчин 2	Үндэсний урлагийн их театрын бүжигчин, МУГЖ	Дунд (30-45)	Эмэгтэй
Бүжигчин 3-4	Орон нутгийн чуулганы бүжигчин	Дунд (30-45)	Эмэгтэй
Судлаач	Урлаг судлалын докторант	Дунд (30-45)	Эмэгтэй
Оюутан 1-2	Бүжгийн урлагийн сургуулийн оюутан	Залуу (18-22)	Эмэгтэй
Оюутан 3-4	Бүжгийн урлагийн сургуулийн оюутан	Залуу (18-22)	Эрэгтэй

TABLE II. Туршилтын Нэгдсэн Өгөгдөл

Он сар	Оролцогчдын бүрэлдэхүүн	Бичлэг ийн тоо	Нийт хугацаа (мин)	Хэмжээ (MB)
2025.04.29	Орон нутгийн чуулганы бүжигчин	1	5.15	17.07
2025.10.17	Урлаг судлаач	3	6.46	177.32
2025.10.23	Мэргэжлийн бүжигчид (МУГЖ)	19	37.52	1035.25
2025.11.07	Бүжгийн ангийн оюутнууд	39	40.08	1092.94
Нийт		*62*	*89.21*	*2322.58*

B. Өгөгдөл урьдчилан боловсруулалт

Түүхий өгөгдлийг шинжилгээнд ашиглахад тохиромжтой болгох, тооцооллын ачааллыг бууруулах зорилгоор дараах шат дараалсан боловсруулалтыг хийсэн.

Frame rate resampling: OptiTrack системээс гарсан түүхий өгөгдөл нь 240 fps (frames per second) давтамжтай байсан бөгөөд үүнийг өгөгдөл боловсруулах алгоритмын ачааллыг оновчлох зорилгоор 72 fps болгон бууруулсан. Ингэхдээ өгөгдлийн алдагдалгүйгээр хугацааны интервалыг дахин тооцоолох (resampling) алгоритмыг ашиглав.

Хөдөлгөөний сегментчилэл (Segmentation): Мэргэжлийн бүжигчдээс авсан 19 багц түүхий өгөгдлийг (raw data) нарийвчлан шинжилж, бүжгийн суурь элементүүдээр нь салгасан. Энэ шатанд мэргэжлийн бүжигчид болон урлаг судлаачидтай зөвлөлдөж, хөдөлгөөн бүрийн эхлэх болон дуусах хугацааг (frame index) кадр бүрээр (frame-by-frame) шалган тэмдэглэсэн.

Суурь хөдөлгөөний сонголт: Сонгогдсон хөдөлгөөнүүд нь судалгааны эх хэрэглэгдэхүүн болж байгаа 20 бүжигт зонхилсон, мөн "үзэмчин угсаатан"-ы бүжгийн онцлогийг илэрхий харуулсан гэсэн урлаг судлаачдын дүгнэлтэд үндэслэсэн ба сегментчилэлийн үр дүнд нийт 209 жижиг хөдөлгөөн болон задарсан. Үүнээс үндсэн 10 төрлийн суурь хөдөлгөөнийг шигшин сонгов. Сонгогдсон хөдөлгөөн бүр 2-12 удаагийн давталттай, нийт 72 нэгж өгөгдөл (clips) болсон бөгөөд хөдөлгөөний дундаж үргэлжлэх хугацаа 3.7 секунд байна.

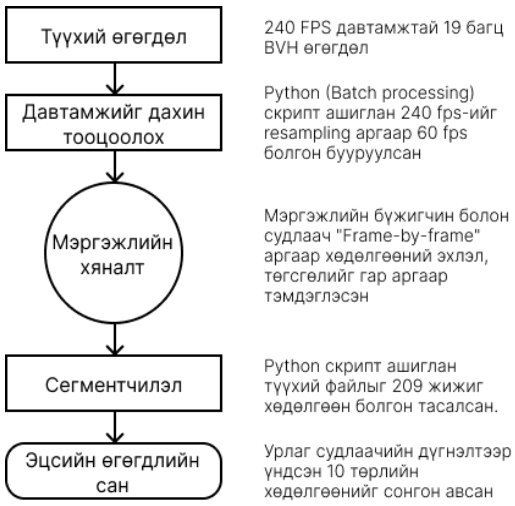


Fig. 2. Өгөгдөл урьдчилан боловсруулах үйл явцын блок схем

Энэ түүхий хөдөлгөөний өгөгдөл бэлэн болсноор судлаачид цаашид клиник, биомеханик, урлаг, дижитал медиа зэрэг олон салбарт хөдөлгөөний нарийвчлалтай шинжилгээ хийх, виртуал сэргээн босголт, хөдөлгөөн барих зэрэг цогц дижитал хэрэглээний боломжуудад ашиглах боломжтой. Мөн уламжлалт бүжгийн хөдөлгөөнүүдийг дижиталжуулж, түүх, соёлын өвийг хадгалах, виртуал дүр бүтээх, интерактив VR/AR контентод хэрэглэх боломж бүрдэнэ. Эдгээр нь дижитал медиа болон соёлын өвийг хамгаалах салбарт өргөн боломжуудыг нээж байгаа хэдий ч энэхүү судалгааны ажил нь бүжгийн хөдөлгөөний кинематик шинж чанарыг тодорхойлох үндсэн зорилтын хүрээнд хязгаарлагдаж байгаа тул дээрх хэрэглээний чиглэлүүдийг дараагийн шатны судалгааны ажилд тусгайлан авч үзэхээр үлдээв.

C. Хөдөлгөөний кинематик шинжилгээ

Энэхүү бүлэгт цуглуулсан хөдөлгөөний өгөгдлийг (Motion Capture Data) биомеханик болон алгоритмын үүднээс шинжлэх арга зүйг танилцуулна. Бүжгийн мэргэжлийн багш, судлаачидтай зөвлөлдсөний үндсэн дээр Үзэмчин бүжгийн техник ажиллагаа, хөдөлгөөний хэв маягийг тоон утгаар илэрхийлэхэд нэн шаардлагатай кинематик шинжилгээний зургаан үзүүлэлтийг тодорхойлж, биомеханикийн шугаман болон өнцгийн хэмжилтийн аргуудыг сонгон авлаа.

Шинжилгээ бүрийн гүйцэтгэлийг дараах бүтцийн дагуу танилцуулна:

1. Алгоритмын логик: Тухайн шинжилгээний программын алгоритмыг блок схемээр дүрслэн харуулна.
2. Тооцооллын арга зүй: Код болон математик загварчлалаар хэрхэн хэрэгжүүлсэн тайлбар.
3. Хэрэглээний ач холбогдол: Тухайн шинжилгээний үр дүн нь бүжгийн урлагийн сургалт, судалгаа болон мэргэжлийн гүйцэтгэлийг үнэлэхэд ямар бодит ач холбогдолтой болохыг бүжгийн онолтой холбон тайлбарласан болно.

Эдгээр шинжилгээнүүд нь бүжгийн хөдөлгөөнийг зөвхөн үзэмжээр бус, шинжлэх ухааны үндэслэлтэй тоон үзүүлэлтээр үнэлэх боломжийг олгож байгаагаараа онцлогтой юм.

1) Өгөгдөл боловсруулах суурь систем: BVH Parser

Бүжгийн хөдөлгөөнийг дижитал хэлбэрээр шинжлэх, түүний дотор биомеханик үзүүлэлтүүдийг тооцоолох хамгийн эхний бөгөөд шийдвэрлэх алхам бол .bvh (Biovision Hierarchy) файлыг уншиж, компьютерт боловсруулах боломжтой математик өгөгдөл болгон хөрвүүлэх явдал юм. Энэхүү процесс нь хоёр үндсэн үе шатаас бүрддэг.

Нэгдүгээрт, хүний биеийн бүтцийг шатлан захирагдсан (Hierarchical) мод хэлбэрээр дүрслэх HIERARCHY хэсгийг уншиж, үе мөч бүрийн OFFSET буюу ясны уртын өгөгдлийг цуглуулдаг. Хоёрдугаарт, MOTION хэсэг дэх фрейм бүрд харгалзах эргэлтийн болон байрлалын өгөгдлийг (CHANNELS) авч боловсруулна. Шинжилгээний явцад биеийн үе бүрийн орон зайн байрлалыг тодорхойлохдоо "Recursive Tree Traversal" алгоритмыг ашигласан. Хөдөлгөөний өгөгдөл дэх эйлерийн өнцгүүдийг шууд ашиглах үед үүсдэг Gimbal Lock буюу тэнхлэгийн давхцалаас үүдэлтэй утга алдалтыг шийдвэрлэхийн тулд фрейм бүрийн өгөгдлийг 4x4 хэмжээст хувиргалтын матриц (Transformation Matrix) руу хөрвүүлэн тооцоолсон. Энэхүү матриц нь орон зайн эргэлт (R) болон шилжилтийг (T) нэгэн зэрэг дараах хэлбэрээр илэрхийлдэг:

$$M = \begin{bmatrix} R_{11} & R_{12} & R_{13} & T_x \\ R_{21} & R_{22} & R_{23} & T_y \\ R_{31} & R_{32} & R_{33} & T_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Үе мөч бүрийн дэлхийн координатыг (Global Position) гаргаж авахдаа parent үеийн матрицыг локал хувиргалтын матрицаар матрицын үржвэр

$$M_{world} = M_{parent} \times M_{local}$$

хийх замаар рекурсивээр бодсон. Ингэж матрицын хуримтлал (Matrix Accumulation) ашиглах нь тригонометрийн тодорхойгүй утга үүсэхээс сэргийлж, хөдөлгөөнийг миллиметрийн нарийвчлалтай X, Y, Z цэгүүдэд алдаагүй шилжүүлэх боломжийг олгосон юм. Энэхүү суурь систем нь нүдэнд харагдахгүй байгаа бичил хөдөлгөөний өөрчлөлтүүдийг тоон хэлбэрт шилжүүлж, улмаар хөдөлгөөний кинематик шинжилгээ хийх, дүрслэлүүд үүсгэх зэрэгт ашиглах боломжтой математик өгөгдөл болгон хөрвүүлж байна. Судалгаанд ашигласан өгөгдөл боловсруулалтын ерөнхий логикийг дараах псевдо кодоор өгүүлэв.

Оролт: .bvh өргөтгөлтэй файлын өгөгдөл
Гаралт: T хугацааны турших N тооны үе мөчний глобал (X, Y, Z) координатууд

1. Их биеийн бүтэц (Hierarchy) унших:
   ROOT болон JOINT түлхүүр үгсийг ашиглан биеийн үе мөчдийн шатлан захирагдсан бүтцийг (Tree structure) байгуулна. Үе бүрийн Offset (ясны урт) болон Channel (эргэлтийн дараалал)-ын мэдээллийг хадгална.
2. Хөдөлгөөний өгөгдөл (Motion) боловсруулах:
   MOTION хэсгээс нийт фреймийн тоо болон хугацааны мэдээллийг авч, тоон утгуудыг матрицад шилжүүлнэ.
3. Координатын тооцоолол (Фрейм бүрээр $t \in T$):
   i. Локал шилжилт (L_{trans}): Offset утга болон тухайн фрейм дэх байрлалын өгөгдлийг ашиглан шилжилтийн матриц үүсгэнэ.
   ii. Локал эргэлт (L_{rot}): Эйлерийн өнцгүүдийг (жишээ нь: Z, X, Y) радиан руу шилжүүлж, эргэлтийн матриц болгон хөрвүүлнэ.
   iii. Локал хувиргалт (M_{local}): Эргэлт болон шилжилтийг нэгтгэсэн 4x4 хэмжээст локал матрицыг ($M_{local} = L_{trans} \times L_{rot}$) байгуулна.
   iv. Глобал хувиргалт (M_{world}): Эцэг үеийн матрицаар локал матрицыг үржиж ($M_{world} = ParentTransform \times M_{local}$) тухайн үеийн нийт хувиргалтыг гаргана.
   v. Координат гаргалт (P_{global}): M_{world} матрицын 4-р баганаас (эсвэл сүүлийн мөрнөөс хамаарч) X, Y, Z байршлын утгуудыг салган авч хадгална.
   vi. Рекурсив дуудалт: Тухайн үений бүх "child" үеүдийн хувьд ComputeGlobal функцийг шинэчлэгдсэн M_{world} матрицаар дамжуулан дахин ажиллуулна.
4. Гүйцэтгэл: Тооцооллыг Root буюу төв цэгээс эхлүүлж, нэгж матриц (Identity Matrix)-аар дамжуулан бүх фреймд гүйцэтгэнэ.

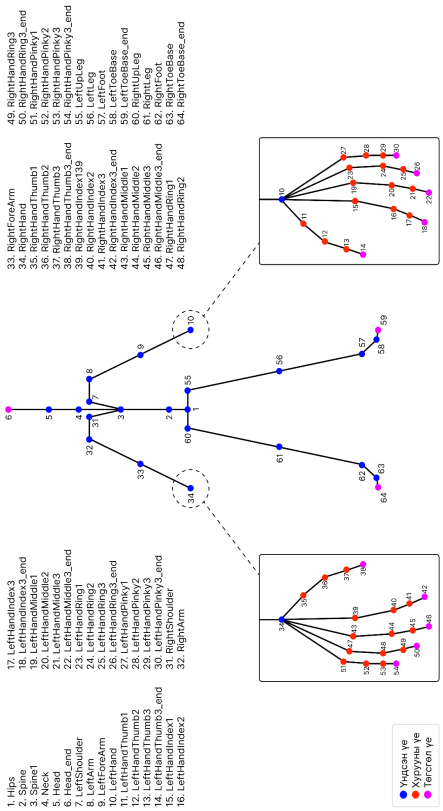


Fig. 3. BVH өгөгдлийн бүтэц

2) Шугаман кинематик параметрүүд

Энэхүү хэсэгт бүжигчний орон зайд шилжих хурдыг BVH файлын "Root" буюу үндсэн үеийн координатын өөрчлөлтөд тулгуурлан тооцоолсон.

Алгоритмын логик: Программчлалын алгоритм нь BVH файлыг задлахдаа C1 хэсэгд өгүүлсэн өгөгдөл боловсруулах суурь систем: BVH Parser-ийг ашиглан задалснаар эхэлнэ. Дараа нь дараалсан кадруудын хоорондох евклидийн зайг тооцоолж, түүнийг хугацааны алхамд (frame time) хуваах замаар агшин зуурын хурдыг гаргаж авдаг. Алгоритмын ерөнхий дарааллыг Зураг 4-д блок схемээр харуулав.



Fig. 4. Хурд тооцоолох үйл явцын блок схем

Математик загварчлал: Хөдөлгөөний хурдыг тооцоолохдоо дараах математик загварыг ашигласан. Хэрэв i дугаар фрейм дэх байрлалыг $P_i = (x_i, y_i, z_i)$

гэж тодорхойлбол, шилжилтийн зай d болон агшин зуурын хурд v-ийг дараах томьёогоор бодно:

Шилжилтийн зай (Euclidean distance):

$$d = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2 + (z_i - z_{i-1})^2}$$

Агшин зуурын хурд (Velocity):

$$v = \frac{d}{\Delta t}$$

Энд Δt нь кадруудын хоорондох хугацааны интервал (Frame Time) юм. Бидний тохиолдолд өгөгдлүүд 72fps үзүүлэлттэй байгаа тул Δt нь 0.013889.

Хэрэглээний ач холбогдол: Үзэмчин бүжгийн суулт, босолт болон цохилттой хөдөлгөөний эрч хүчийг тоон утгаар үнэлэхэд хурдны үзүүлэлт чухал ач холбогдолтой. Энэхүү шинжилгээ нь мэргэжлийн бүжигчний гүйцэтгэлийн "динамик хэв шинж"-ийг тодорхойлж, сургалтын явцад оюутны хөдөлгөөний хурдны алдааг засварлах бодитой хэрэглэгдэхүүн болно.

3) Өнцгийн динамик ба чиглэл

Хөдөлгөөний орон зайн шинжилгээнд хүний биеийн байрлал, чиглэлийг тодорхойлохын тулд орон зайн босоо болон хэвтээ чиг баримжааг ялган авч үзсэн. Орон зайн босоо чиг баримжаа нь хүний биеийн хүндийн төвийг суурь цэг болгон авч, тухайн цэгээс дээш болон доош чиглэсэн координатын утгаар орон зайг ангилан тодорхойлно[17]. Үүний дагуу хөдөлгөөнийг дээд, дунд, доод гэсэн гурван түвшинд хувааж үзсэн бөгөөд энэхүү ангиллыг Зураг 5-т харуулсан.

Харин орон зайн хэвтээ чиг баримжаа нь хүний хөдөлгөөнийг x–z хавтгайд проекцлон авч, хөдөлгөөний чиглэлийн өнцгийг тооцоолсны үндсэн дээр тодорхойлогдоно. Үүнд, хэвтээ хавтгайг 360° өнцгийн мужид 45° интервалтайгаар хувааж, нийт 8 үндсэн чиглэл (урагш, зүүн-урд, зүүн, зүүн-хойд, хойш, баруун-хойд, баруун, баруун-урд) болгон ангилсан. Энэхүү чиглэлийн ангиллыг Зураг 5-т үзүүлэв. Ийнхүү босоо болон хэвтээ чиг баримжааг хослуулснаар хөдөлгөөний орон зайн байрлал, чиглэлийг илүү нарийвчлалтай тодорхойлох боломж бүрдэж, бүжгийн хөдөлгөөний бүтцийн шинжилгээнд ашиглах суурь үзүүлэлт болж байна. Туршилтыг сонгосон суурь 10 хөдөлгөөний хугацааны эхэнд болон сүүлд биеийн таван хэсэг аль түвшинд ямар чиглэлд байгааг тооцоолж байгаа тул нийт 10*2*5=100 хэсэг шинжилгээний хариу гарна.

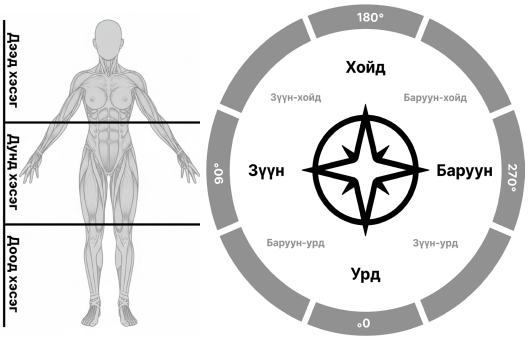


Fig. 5. Босоо болон хэвтээ хуваалт

Алгоритмын логик: Бүжгийн хөдөлгөөний орон зайн шинжилгээг автоматжуулахын тулд BVH өгөгдлөөс биеийн гол үе мөчдийг (Толгой, Баруун гар, Зүүн гар, Баруун хөл, Зүүн хөл) ялган авч, тэдгээрийн хөдөлгөөний вектор болон чиглэлийг тооцоолох алгоритмыг боловсруулсан. Энэхүү алгоритм нь түүхий өгөгдлийг (raw data) Монгол бүжгийн хөдөлгөөний зүй тогтолд нийцсэн 8 үндсэн чиглэл болон 3 түвшний ангилалд шилжүүлэх үндсэн үүрэгтэй бөгөөд дараах боловсруулалтын шатуудаар гүйцэтгэгдэнэ. Үүнд: Вектор тооцоолол: Фрейм бүрийн хоорондын шилжилтийн вектор ($\Delta x, \Delta y, \Delta z$)-ыг бодож, биеийн хүндийн төв болон гол үе мөчдийн орон зайн өөрчлөлтийг тодорхойлно. Түвшин тогтоох: Биеийн y координатын утгыг босоо тэнхлэгийн дагуу шинжилж, хөдөлгөөнийг дээд, дунд, доод гэсэн гурван түвшний ангилалд тогтооно. Чиглэл тодорхойлох: Хөдөлгөөнийг $x - z$ хавтгайд проекцлон, $\arctan 2(\Delta z, \Delta x)$ томьёогоор өнцгийг бодож, 45° интервалтай 8 үндсэн чиглэлд (урагш, хойш, зүүн, баруун г.м) харгалзуулан (Mapping) ангилна. Алгоритмын ерөнхий ажиллагаа болон өгөгдөл дамжих дарааллыг дараах блок схемээр үзүүлэв.

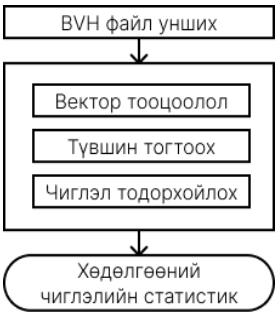


Fig. 6. Чиглэл тооцоолох явцын блок схем

Математик загварчлал: Хөдөлгөөний чиглэлийг тодорхойлохдоо орон зайн X болон Z (хэвтээ хавтгай) тэнхлэгийн өөрчлөлтийг ашиглан арктангенс функцээр өнцгийг бодсон:

$$\theta = \text{atan2}(\Delta x, \Delta z) \times \frac{180}{\pi}$$

Гаргаж авсан өнцгийг (θ) Монгол бүжгийн орон зайн 8 чиглэлд Зураг 5 дээрх хязгаараар ангилсан.

Хэрэглээний ач холбогдол: Дээрх алгоритмаар гарган авсан орон зайн босоо болон хэвтээ чиглэлийн тоон үзүүлэлтүүд нь бүжгийн хөдөлгөөнийг шинжлэх ухааны үндэслэлтэй "албан ёсны дүрслэх хэл"-ээр илэрхийлэх боломжийг олгож байна. Түүнчлэн, хөдөлгөөн барих mocap өгөгдлийг текстэн тайлбартай холбох, бүжигчний орон зайн ашиглалт болон техникийн гүйцэтгэлийг биомеханик талаас нь үнэлэх, сургалтын шинэ арга зүйг боловсруулах бодит хэрэглүүр болж байна. [17]

4) Хөдөлгөөний чанарын шинжилгээ

Бүжгийн хөдөлгөөнийг зөвхөн орон зайн координат төдийгүй, "илэрхийлэмж" талаас нь шинжлэх

арга зүйг энэ шинжилгээгээр тайлбарлана. Спортын биомеханикийн шинжилгээний аргад тулгуурлан хөдөлгөөнийг текстэн хэлээр тайлбарлах, мокап (mocap) өгөгдлийн орон зай-цаг хугацааны шинжүүдийг ашиглан биеийн байрлал *(хэлбэржсэн түүр зогсолт)*-ыг тодорхойлох, улмаар хөдөлгөөний чанарыг задлан шинжлэх гэсэн үндсэн гурван чиглэлийг баримталлаа. Үзэмчин бүжгийн хөдөлгөөнийг тодорхойлох хэлийг гаргахын тулд юун түрүүнд бүжгийн мэргэжилтэнтэй зөвлөлдөж, суурь 10 хөдөлгөөнийг эрчим, хэмнэл, хяналтын хувьд тодорхойлуулсан. Энэхүү аман тайлбар дээр үндэслэн Лабаны хөдөлгөөний шинжилгээний арга зүйг туршсан бөгөөд орон зайн чиглэлийн шинжилгээг өмнөх бүлэгт авч үзсэн тул энд weight (эрчим\хүч\жин), time (хугацаа), flow (урсгал) гэсэн гурван суурь үзүүлэлтээр хөдөлгөөний чанарыг тодорхойлов.

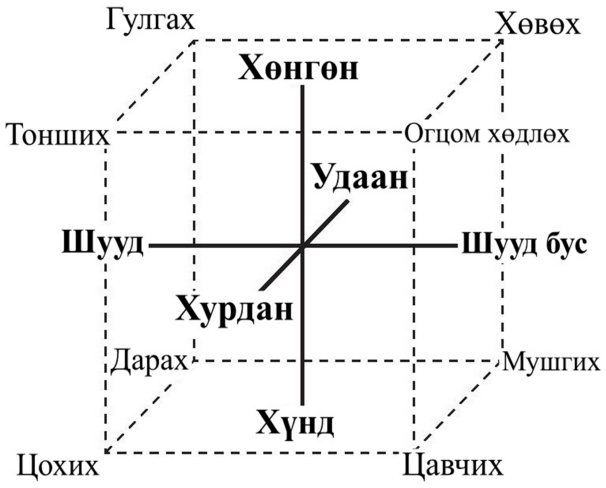


Fig. 7. Лабаны хөдөлгөөний эрчмийн диаграмм

Алгоритмын логик: Хөдөлгөөний чанарыг тодорхойлох алгоритм нь түүхий өгөгдлийг (.bvh) Лабаны "Effort" буюу хүчин чармайлтын үгэн тайлбар руу хөрвүүлэх зарчмаар ажиллана. Оролт нь Үзэмчин бүжгийн 10 үндсэн хөдөлгөөний 72 дэд хэсгээс бүрдэх сан бөгөөд алгоритм нь дараах дарааллаар ажиллана: Нийт сангаас хурд, хурдатгал, хурдатгалын өөрчлөлт үзүүлэлтүүдийн хамгийн их болон хамгийн бага утгыг тодорхойлно. Эдгээр утгыг мэргэжилтний аман дүгнэлттэй уялдуулан босго утга болгон тогтооно. Шинжилж буй хөдөлгөөний өгөгдлийг босго утгаар нормчилж, Лабаны 3 суурь үзүүлэлтэд (Weight, Time, Flow) харгалзуулан "чанарын тайлан" болгон гаргана.

Математик загварчлал: Хөдөлгөөний шинж чанарыг тоон үзүүлэлтээр илэрхийлэхдээ физик динамикийн дараах математик загварчлалыг ашигласан.

Time (Хугацаа-Velocity): Биеийн төв цэгийн агшин зуурын хурдаар тооцно. Хэрэв хурд босго утгаас их бол Sudden буюу огцом, бага бол sustained буюу үргэлжилсэн гэж үзнэ.

$$v = \frac{\|\vec{P}_2 - \vec{P}_1\|}{\Delta t}$$

Weight (Хүч/Жин- Acceleration): Хурдатгалын далайцаар тооцно. Өндөр хурдатгал нь Strong (Хүчтэй), бага хурдатгал нь Light (Хөнгөн) хөдөлгөөнийг илэрхийлнэ.

$$a = \frac{\Delta v}{\Delta t}$$

Flow (Урсгал-Jerk): Хурдатгалын өөрчлөлтийн хурд буюу "Jerk" утгаар хэмжинэ. Хөдөлгөөний тасалдал их бол Bound (Хязгаарлагдмал), зөөлөн урсаж байвал Free (Чөлөөт) гэж тодорхойлно.

$$j = \frac{\Delta a}{\Delta t}$$

Хэрэглээний ач холбогдол: Бүжгийг зөвхөн геометрийн цэгүүдийн шилжилт гэж харахаас илүүтэйгээр түүний дотоод эрч хүч, урлагийн илэрхийлэмжийг математик аргаар баримтжуулж буй нь бүжиг судлалын шинэ хандлага юм. Үзэмчин бүжгийн өвөрмөц хөдөлгөөнүүдийг "Хөнгөн бөгөөд Үргэлжилсэн" (Light & Sustained) гэх мэт Лабаны нэр томьёогоор сургалтад ашиглах нь хөдөлгөөний чанарыг алдаагүй өвлүүлэн үлдээх гүүр болно.

5) Орон зайн харьцуулсан шинжилгээ

Орон зайн эзлэхүүн болон биеийн хэсгүүдийн оролцоог тоон утгаар тодорхойлсон шинжилгээ. Анхлан бүжгийн хөдөлгөөнийг бүтэн биеийн хэмжээнд шинжилсэн боловч үзэмчин бүжгийн онцлогт тулгуурлан мэргэжилтний зөвлөснөөр гар болон хөлийн хөдөлгөөний далайцыг харьцуулан шинжлэх арга зүй рүү шилжүүлсэн.

Алгоритмын логик: Шинжилгээний алгоритм нь BVH өгөгдөл дэх үе мөчдийн нэршил дээр үндэслэн гар болон хөлийн цэгүүдийг (joints) шүүж эзлэхүүнийг тооцох замаар ажиллана. Өгөгдөл шүүлт: Скелет бүтцээс гар болон хөлийн үеүдийн индексийг ялгаж, фрейм бүр дэх тэдгээрийн X, Y, Z координат болон захын утгуудыг цуглуулна. Эзлэхүүн тооцоолол: Тухайн хөдөлгөөний туршид гар болон хөл тус бүрийн орон зайд үүсгэсэн хамгийн их болон хамгийн бага цэгүүдийг тодорхойлж, хөдөлгөөний нийт далайцыг (Bounding box) бодно. Харьцуулалт: Гар болон хөлийн эзлэх орон зайн дундаж утгыг харьцуулснаар тухайн бүжгийн хөдөлгөөнд биеийн аль хэсэг түлхүү оролцож байгааг тодорхойлно. Биеийн цэгүүдийн хуваарилалтыг доорх зурагд харуулав.

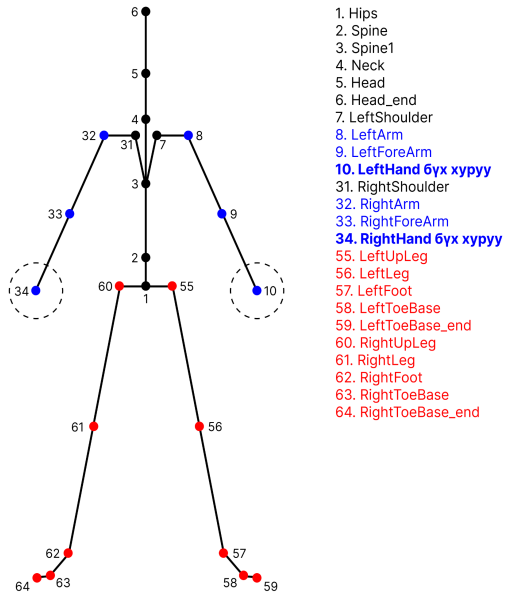


Fig. 8. Орон зайн шинжилгээнд оруулсан цэгүүд (joints)

Математик загварчлал: Орон зайн эзлэхүүн болон хөдөлгөөний траекторыг тодорхойлохдоо дараах математик хэмжигдэхүүнүүдийг ашигласан: Орон зайн хязгаар (Space Range): Тэнхлэг тус бүр дэх туйл утгуудын зөрүүгээр хөдөлгөөний далайцыг тодорхойлно.

$$\Delta X = X_{max} - X_{min}, \quad \Delta Y = Y_{max} - Y_{min}, \quad \Delta Z = Z_{max} - Z_{min}$$

Дундаж орон зай (Mean Space): Гар болон хөлийн орон зайд эзлэх дундаж хэмжээсийг тэнхлэгүүдийн дунджаар гаргана:

$$S_{mean} = \frac{\Delta X + \Delta Y + \Delta Z}{3}$$

Мурийлт (Curvature): Хөдөлгөөний нийт туулсан зам (Total Distance)-ыг шилжилт (Displacement)-д харьцуулж, хөдөлгөөний орон зайн ашиглалтын нарийн төвөгтэй байдлыг тооцно.

Хэрэглээний ач холбогдол: Энэхүү алгоритмын тусламжтайгаар тухайн бүжиг орон зайд ямар хэмжээний зай талбай ашиглаж байгааг урьдчилан тооцоолох боломжтой. Энэ нь тайзны найруулга болон бүжигчдийн байршлыг төлөвлөхөд чухал өгөгдөл болно. Үзэмчин бүжгийн гар болон хөлийн хөдөлгөөний оролцоог тоон утгаар харьцуулснаар тухайн бүжгийн хэв маяг, биеийн хэсгүүдийн ажиллагааны онцлогийг бодитой тодорхойлно.

б) Гурван цэгийн өнцгийн тооцоолол

Бүжгийн хөдөлгөөний нарийвчлал, үе мөчний нугалалт болон биеийн ерөнхий дүр төрхийг тоон утгаар илэрхийлэх бас нэгэн чухал үзүүлэлт бол үе хоорондын өнцөг юм. Энэхүү хэсэгт BVH өгөгдлөөс хүний анатомийн бүтцийн дагуу үе мөчний өнцгийг тооцоолох аргачлалыг танилцуулж байна.

Алгоритмын логик: Гурван цэгийн өнцгийг тооцоолох алгоритм нь орон зайн вектор анализад тулгуурладаг. Программ нь BVH файлаас тухайн фрейм дэх бүх үе мөчний глобал координатыг (X, Y, Z) гаргаж авсны дараа шинжлэхийг хүссэн гурван цэгийг (жишээ нь: мөр[A], тохой[B], бугуй[C]) сонгон авч, тэдгээрийн хоорондох векторуудыг байгуулж, косинусын теоремоор өнцгийг гаргадаг.

Математик загварчлал: Кинематик шинжилгээг хэрэгжүүлэхдээ бид скаляр үржвэрийн (Dot Product) аргыг ашигласан. Хэрэв бидэнд A, B, C гэсэн гурван цэг өгөгдсөн бол B цэг дээрх өнцгийг дараах алхмаар тооцоолно:

Вектор байгуулах: B цэгийг эхлэл цэг болгон BA болон BC векторуудыг тодорхойлно.

$$\vec{BA} = A - B, \quad \vec{BC} = C - B$$

Скаляр үржвэр ба Норм: Хоёр векторын скаляр үржвэрийг тэдгээрийн уртын үржвэрт хуваж косинус утгыг олдог.

$$\cos(\theta) = \frac{\vec{BA} \cdot \vec{BC}}{\|\vec{BA}\|\|\vec{BC}\|}$$

Өнцөг гаргах: Косинусын урвуу функц arccos ашиглан радианаар гарсан утгыг градус руу хөрвүүлнэ.

$$\theta = \arccos(\text{clip}(\cos(\theta), -1.0, 1.0)) \times \frac{180}{\pi}$$

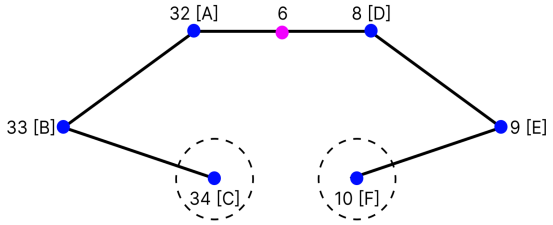


Fig. 9. Өнцгийн жишээ (Тохойн нугалаас)

Хэрэглээний ач холбогдол: Уг аргачлал нь бүжгийн техникийн гүйцэтгэлийг стандартчилах суурь болж байна. Тухайлбал, мэргэжилтнүүдийн үндсэн зогсолт, гар барилт гэж үзэж буй "Pose"-ийн тохой болон бугуйн өнцгийн хамгийн тохиромжтой хэмжээг тоон утгаар тодорхойлсноор суралцагчийн гүйцэтгэлийг мэргэжлийн багшийн стандарттай бодитой харьцуулан үнэлэх боломж бүрдэнэ.

Академик балетын техникт ч хөдөлгөөний нарийвчлалыг ийнхүү өнцгийн хэмжигдэхүүнээр тогтоосон байдаг. Жишээлбэл, Arabesque хөдөлгөөний үед агаар дахь хөл болон биеийн босоо тэнхлэг хоорондын өнцгийг гүйцэтгэлийн түвшнээс хамааран 90° болон түүнээс дээш байхаар заадаг бол, En dehors буюу хөлийг гадагш эргүүлэх техникт ташааны үе болон өвдөг, шагайн нэгдмэл эргэлтийн өнцөг нь нийт 180°-д дөхөж байхыг мэргэжлийн стандарт гэж үздэг [22].

Эцэст нь, уламжлалт бүжгийн өвийг дижитал хэлбэрт шилжүүлэхдээ зөвхөн дүрс бичлэгээр хязгаарлахгүй, нарийн өнцгийн өгөгдлүүдээр баримтжуулах нь хойч

үедээ арга техникийг нь алдаа мадаггүй өвлүүлэн үлдээх сургалт-эрдэм шинжилгээний архив болох юм.

7) Гол тэнхлэгийн (Principal Axis) илрүүлэлт

Бүжгийн хөдөлгөөний орон зайн зохиомж, биеийн тэнцвэр болон хөдөлгөөний идэвхтэй бүсийг тодорхойлохын тулд хөндлөн "Гол тэнхлэг"-ийг илрүүлэх шинжилгээг гүйцэтгэв. Энэхүү шинжилгээг мэргэжилтний ажиглалт болон компьютерын тооцоололтой хослуулсан гүйцэтгэсэн.

Алгоритмын логик: Алгоритмын үндсэн логик нь таамаглалаас эхэлж, өгөгдлийн статистик боловсруулалтаар дамжин, нэгдсэн дүгнэлтэд хүрэх байдлаар явагдсан. Эхний шатанд мэргэжлийн бүжиг судлаачид 20 гаруй Үзэмчин бүжгийн бичлэгийг ажиглаж, бүжгийн төрөл зүйлийн онцлогоос хамааран хөдөлгөөний төвлөрөл (жишээ нь: мөрний түвшин, бүсэлхий эсвэл өвдөгний орчим) хаана байгааг тодорхойлсон. Үүний дараагаар N тооны фрейм бүхий BVH өгөгдлийн хувьд бүх хугацааны турших араг ясны зангилаа цэгүүдийг нэг орон зайд $S = \{P_1, P_2, \ldots, P_n\}$ олонлог болгон нэгтгэж, хөдөлгөөний "дулааны зураг" буюу траекторын бөөгнөрлийг үүсгэнэ. Цугларсан цэгүүдийн Y координатыг тодорхой w (bin width) өргөнтэй мужуудад хувааж, муж бүрийн нийлбэр цэгүүдийн тоог тооцно. Эцэст нь гистограммын хамгийн өндөр утгатай мужийг олж, түүний төв цэгийг тухайн хөдөлгөөний гол тэнхлэг ($Y_{principal}$) гэж тодорхойлсон. Доорх зурагд алгоритмын логикийг блок схемээр харуулав.



Fig. 10. Тэнхлэг илрүүлэх явцын блок схем

Математик загварчлал: Тооцооллын хувьд орон зайн цэгүүдийн нягтралын функцийг дискрет хэлбэрээр үнэлж, хамгийн их магадлалт утгыг олох аргыг ашигласан. Хэрэв $F = \{f_1, f_2, \ldots, f_M\}$ нь нийт фреймүүдийн олонлог, J нь үе мөчний тоо бол нийт Y координатын олонлог Y_{all} нь дараах байдлаар тодорхойлогдоно. Үүнд орон зайн гурван хэмжээст (x, y, z) координатын системээс x тэнхлэгийн утгыг хасаж, өгөгдлийг хоёр хэмжээст zy хавтгайд проекцлон буулгах замаар бүжигчний хөдөлгөөний босоо тэнхлэгийн дагуух нягтрал болон гүний шилжилтийг нэгэн зэрэг шинжлэх боломжийг бүрдүүлсэн.

$$Y_{all} = \{y_{i,j} \mid 1 \le i \le M, 1 \le j \le J\}$$

Эдгээр утгуудыг $[y_{min}, y_{max}]$ мужид w өргөнтэй k тооны сагсанд (bins) хуваахдаа B_k сагсанд хамаарах цэгүүдийн тоог (Frequency) дараах функцээр тооцно:

$$H(B_k) = \sum_{y \in Y_{all}} I(y, B_k)$$, үүнд

$$I(y, B_k) = \begin{cases} 1 & \text{herev } y \in [e_k, e_{k+1}) \\ 0 & \text{busad ved} \end{cases}$$

Энд e_k нь k-р сагсны хилийн утга юм. Хамгийн их давтамжтай сагсыг олсноор гол тэнхлэгийн төв цэгийг ($Y_{principal}$) дараах оновчлолын томьёогоор гаргаж авна:

$$k^* = \text{argmax}_k H(B_k)$$

$$Y_{principal} = \frac{e_{k^*} + e_{k^*+1}}{2}$$

Тооцооллын нарийвчлалыг сайжруулахын тулд w = 3 пиксел байхаар тогтоосон бөгөөд энэ нь хөдөлгөөний бичил өөрчлөлт болон ерөнхий чиг хандлагын хоорондох алдааг (noise vs trend) тэнцвэржүүлэх оновчтой хэмжээ юм. Тэнхлэгийг илрүүлсний дараа бид эх зураг дээр шилжүүлэн зурахдаа координат хөрвүүлэлтийн $Y_{adjusted} = Y_{principal} + \Delta y$ засварлалыг хийсэн нь дүрслэлийн нарийвчлалыг мэргэжилтний ажиглалттай 95%-ийн магадлалтайгаар нийцүүлсэн.

Хэрэглээний ач холбогдол: Гол тэнхлэгийг илрүүлснээр бүжигчний ур чадвар, техник тухайн тогтоосон тэнхлэгээс хэрхэн хэлбэлзэж буйг (Deviation analysis) тоон утгаар үнэлэх боломжтой болж, субьектив үнэлгээг объектив өгөгдөлд шилжүүлж байна**.** Арван өөр төрлийн хөдөлгөөний гол тэнхлэгүүдийг нэгтгэн харьцуулснаар тухайн бүжгийн хэв маягийн орон зайн "төв" (Spatial Center of Gravity)-ийг олж тогтоосон. Энэ нь тухайн бүжгийг сурч буй суралцагчдад өөрийн хөдөлгөөний дээд, доод хязгаарыг мэдэрч, биеийн ерөнхий зохиомжийг зөв хадгалах бодитой хэмжүүр болдог.

IV. Үр дүн

Өмнөх бүлгээр угсаатан зүйн онцлог шинжийг математик аргачлалаар тооцоолон дараах үр дүнгүүдийг гаргав. Үзэмчин бүжгийн хөдөлгөөний дундаж хурд 0.33 м/сек буюу тайван хэмнэлтэй хэдий ч агшин зуурын дээд хурд нь 1.55 м/сек хүрч Латин бүжгийн динамиктай

дүйцэхүйц огцом шилжилттэй байна. Орон зайн хувьд хөдөлгөөний 74% нь урд хагас тойрогт төвлөрч, хөл 100% доод түвшинд тогтвортой хадгалагдахын зэрэгцээ гарын далайц хөлийнхөөс 1.36 дахин их буюу илүү дэлгэмэл байна. Лабаны хүчин чармайлтын онолоор Weight (0.34) болон Flow (0.34) үзүүлэлтүүд нь уг бүжгийг "Light-Sustained-Free" буюу “хөнгөн-үргэлжилсэн-чөлөөт" шинж чанартай болохыг тоон утгаар харууллаа. Дээрх тойм үр дүнг доор бүлэг тус бүрээр дэлгэрэнгүй өгүүлэв.

1) Шугаман кинематик параметрүүд

Үзэмчин бүжгийн 10 үндсэн хөдөлгөөний 72 дэд хэсэгт хийсэн хурдны шинжилгээгээр дундаж хурд 0.3379м/сек, хамгийн их хурдны дундаж 0.95065м/сек (дээд утга нь 1.55м/сек хүрч байна) байгаа нь тогтоогдлоо. Эдгээр үзүүлэлтүүдийг бусад төрлийн бүжгийн биомеханик үзүүлэлтүүдтэй харьцуулан авч үзвэл үзэмчин бүжгийн кинематик онцлог бидэнд буй өгөгдлийн хүрээнд дараах байдлаар харагдаж байна. Сонгодог балетын "Adagio" буюу удаан, үргэлжилсэн хөдөлгөөнүүдийн дундаж хурд ихэвчлэн 0.2-0.4м/сек хооронд хэлбэлздэг бол "Grand Allegro" буюу үсрэлттэй хэсэгт хурд 1.2м/сек-ээс дээш гардаг [1]. Үзэмчин бүжгийн дундаж хурд (0.33м/сек) нь балетын удаан хэмнэлтэй төстэй боловч хөдөлгөөн бүрийн дээд хурд (Max_avg: 0.95м/сек) нь харьцангуй өндөр байна. Латин америк бүжгийн (жишээ нь, Jive эсвэл Cha-cha) хөлийн хөдөлгөөний дундаж хурд 0.8-1.5м/сек хооронд хэлбэлздэг нь үзэмчин бүжгээс даруй 2-3 дахин хурдан үзүүлэлт юм [2]. Гэсэн хэдий ч үзэмчин бүжгийн 3 дугаар хөдөлгөөний дээд хурд 1.55м/сек (155.3см/сек) хүрч байгаа нь Латин бүжгийн хурдны үзүүлэлттэй дүйцэж байна. Дундаж хурд тогтвортой (0.3м/сек орчим) байгаа нь бүжгийн ерөнхий өгүүлэмж нь тайван байгааг илтгэж байна. Мөн 3 ба 5 дугаар хөдөлгөөнүүд дээр өндөр хурд ажиглагдсан. Доор Зураг 11-д хурдны тархалтыг зургаар үзүүлсэн ба Хүснэгт 3-д хөдөлгөөн бүрийн хурдны дундаж мэдээллийг харуулав.

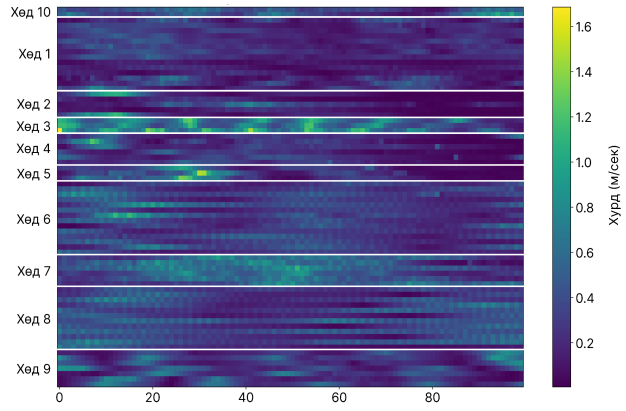


Fig. 11. Хурдны тархалтын дулааны зураглал (Heatmap)

TABLE III. Хөдөлгөөн бүрийн хурдны мэдээлэл

Хөдөлгөөн	Тоо	Дундаж хурд (м/сек)	MAX хурд (м/сек)
Хөдөлгөөн 1	14	0.2142	0.5284
Хөдөлгөөн 2	4	0.1624	0.6838
Хөдөлгөөн 3	3	0.6379	1.553
Хөдөлгөөн 4	5	0.2187	0.8594
Хөдөлгөөн 5	3	0.3052	1.4437
Хөдөлгөөн 6	14	0.3127	0.7524
Хөдөлгөөн 7	6	0.4823	1.0193
Хөдөлгөөн 8	12	0.3346	0.7255
Хөдөлгөөн 9	7	0.2985	1
Хөдөлгөөн 10	2	0.2142	0.5284
Дундаж		*0.3379*	*0.95065*

2) Өнцгийн динамик ба чиглэл

Үзэмчин бүжгийн орон зайн ашиглалтыг биеийн хэсгүүдийн өнцгийн шилжилт, зүг чиг болон босоо түвшний хамаарлаар шинжлэн үзлээ. Энэхүү шинжилгээ нь бүжгийн хөдөлгөөнийг зөвхөн гоо зүйн талаас биш, харин бүжигчний эзэлж буй "Кинесфер" (Kinesphere) буюу биеийн эргэн тойрон дахь орон зайн багтаамжийг тодорхойлоход чиглэсэн болно. Зураг 12-д үзүүлсэн диаграммаас харахад үзэмчин бүжгийн биомеханик бүтэц нь маш тодорхой заагтай байна. Хөлийн хөдөлгөөн 100% доод түвшинд тогтвортой хадгалагдаж байгаа нь Монгол бүжгийн суурь техникийн газардуулга (Groundedness) болон хүндийн төвийн тогтвортой байдлыг илтгэнэ. Харин биеийн дээд хэсэг буюу гар, толгойн хөдөлгөөнүүд дээд болон дунд түвшинд идэвхтэй шилжилт хийж байна. Энэхүү байдал нь олон улсын бүжгийн биомеханик судалгаанд "Динамик тэнцвэр" (Dynamic Balance) гэж нэрлэгддэг бөгөөд бүжигчин орон зайн дээд хэсэгт чөлөөтэй илэрхийлэмж үзүүлэх нөхцөлийг бүрдүүлдэг [4].

Чиглэлийн динамик ба орон зайн баримжааны хувьд Зураг 12-д 10 үндсэн хөдөлгөөний чиглэлийн давтамжийг 8-н зүгээр нэгтгэн харууллаа. Шинжилгээгээр хөдөлгөөний 74% нь урд хагас тойрогт (Урагш, Баруун-урд, Зүүн-урд) төвлөрч байгаа нь ажиглагдав. Хэдийгээр өгөгдөлд баруун-урд чиглэл (32 удаа) давамгайлж буй мэт боловч энэ нь бүжгийн хөдөлгөөнүүдийн тэгш хэмт шинж чанарыг үгүйсгэхгүй юм. Олон улсын хөдөлгөөний шинжилгээнд түгээмэл хэрэглэгддэг онолоор (Laban Movement Analysis) бүжгийн чиглэл нь зүүн эсвэл баруун гэхээс илүүтэй биеийн төвөөс алслагдах (Peripheral) болон төв рүү тэмүүлэх (Central) динамикийн харьцаагаар тодорхойлогддог[5]. Зураг 12-аас харахад Үзэмчин бүжгийн гарын хөдөлгөөнүүд нь 40°-аас 320° хүртэлх маш өргөн өнцгийг хамран нуман хэлбэрээр (Arc-like) шилжиж байгаа нь орон зайг "дэлгэмэл" байдлаар ашиглаж буйг харуулж байна. Хойш чиглэсэн хөдөлгөөн хамгийн бага (нийт давтамжийн 3%) байгаа нь бүжгийн өгүүлэмж үзэгч рүү чиглэсэн нээлттэй проекцтой (Open Projection) болохыг кинематик талаас харуулав. Өнцөг болон чиглэлийн нэгдсэн шинжилгээнээс үзэхэд үзэмчин бүжиг нь орон зайн хувьд баригдмал биш, харин хэвтээ хавтгайд (Horizontal Plane) маш уян хатан шилжилттэй, босоо тэнхлэгт (Vertical Axis) тодорхой түвшний ялгаралтай бүтэцтэй байна. Энэхүү тоон үзүүлэлтүүд нь

Монгол бүжгийн "уужим, дэлгэмэл" хөдөлгөөний хэв маягийг орон зайн геометрийн хэлээр илэрхийлж буй хэрэг юм.

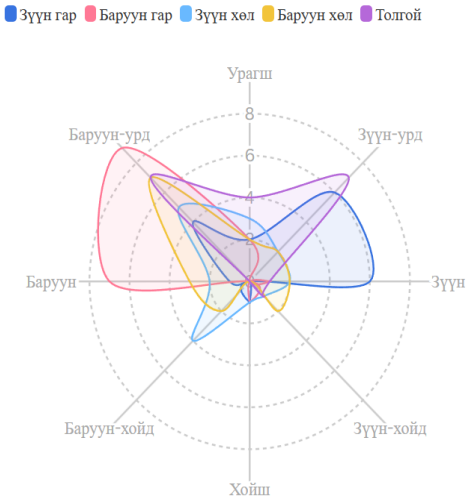


Fig. 12. 8 Зүгийн давтамж

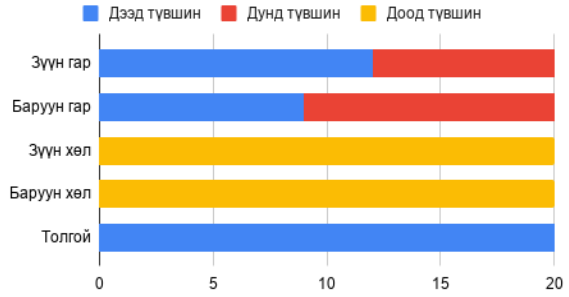


Fig. 13. Түвшний тархалтын диаграмм

3) Хөдөлгөөний чанарын шинжилгээ

Үзэмчин бүжгийн динамик шинж чанарыг Лабаны хөдөлгөөний эрчмийг тодорхойлох онолын "Weight" (Хүч), "Time" (Хугацаа), "Flow" (Урсгал) гэсэн дөрвөн элементийн дагуу шинжлэн үзлээ. Нийт 72 хөдөлгөөний дундаж утга Weight 0.341, Time 0.341, Flow 0.345 байгаа нь уг бүжгийн ерөнхий төрх нь агаарлаг/хөнгөн, зөөлөн бөгөөд урсгал/намуун шинжтэй болохыг харуулж байна.

Урлаг судлаачдын зүгээс үзэмчин бүжгийн дээд биеийн хөдөлгөөнийг "Light Sustained Free" (Хөнгөн-Намуун-Чөлөөт) хэмээн тодорхойлдог нь бидний тооцоолсон 0.34 гэсэн утгатай (босго утга 0.6-аас бага) бүрэн нийцэж байна.

Испани Фламенко бүжигт "Weight" утга 0.7-0.8 хүрдэг бөгөөд time нь sudden (огцом) шинжтэй байдаг [7]. Үүнтэй харьцуулахад үзэмчин бүжиг нь хурдны пульс ихтэй ч нийт динамик нь илүү зөөлөн хөнгөн байна. Сонгодог Балетын гарын хөдөлгөөн Үзэмчин бүжигтэй ижил "Light-Free" шинжтэй боловч хөлийн ажиллагаа нь үсрэлт болон цоройлтын/эргэлт/шилжилт үед "огцом хөдлөх" шинж рүү шилждэг. Харин үзэмчин бүжигт хөл нь үргэлж "удаан" буюу газраас тасрахгүй үргэлжилсэн жигд шинжээ хадгалж байгаа нь онцлог юм [8]. Шинжилгээний үр дүнд үзэмчин бүжиг нь "дээд биеийн чөлөөт урсгал" болон "доод биеийн хяналттай тогтвортой" шинжийн нэгдэл болох нь ажиглагдав. Энэхүү эсрэг тэсрэг динамикийн нэгдэл нь уг бүжгийн хөдөлгөөнийг нэгэн зэрэг уян налархай бөгөөд хээмсэг харагдуулдаг гол хүчин зүйл юм.

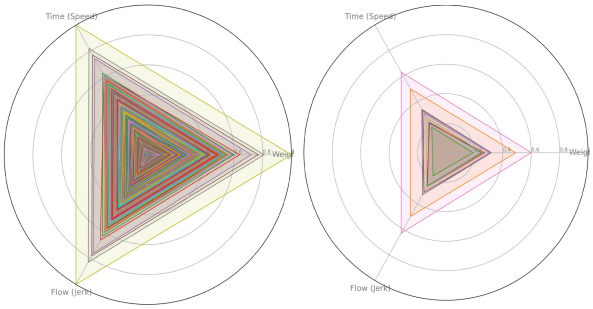


Fig. 14. 72 хөдөлгөөн болон үндсэн 10 хөдөлгөөний чанарын (Laban Effort) шинжилгээ

4) Орон зайн харьцуулсан шинжилгээ

Үзэмчин бүжгийн орон зайн багтаамжийг (Spatial Volume) бүтэн биеийн эзлэх эзлэхүүн болон гар хөл хэсгийн далайцын хүрээгээр тодорхойллоо. Судалгаагаар бүжигчний эзлэх дундаж орон зай нь 3.22м³ байгаа бөгөөд энэ нь уг бүжгийн дэлгэмэл, уужим шинж чанарыг тоон утгаар илэрхийлж байна.

Бүтэн биеийн орон зай:

Хөдөлгөөний дундаж далайцыг тэнхлэгээр авч үзвэл:

- X тэнхлэг (Баруун-Зүүн): 1.62 метр
- Y тэнхлэг (Дээш-Доош): 1.72 метр
- Z тэнхлэг (Урагш-Хойш): 1.11 метр

Үүнээс үзэхэд үзэмчин бүжиг нь босоо чиглэлд (Y) болон хажуу тийш (X) хамгийн их орон зайг ашигладаг бол урагш-хойш чиглэлд (Z) харьцангуй бага шилжилт хийдэг нь ажиглагдлаа. Энэ нь бүжгийн үндсэн өгүүлэмж нь нэг хавтгай дээр (Frontal Plane) дэлгэгдэх шинжтэйг бидний өгөгдлийн хүрээнд харуулж байна.

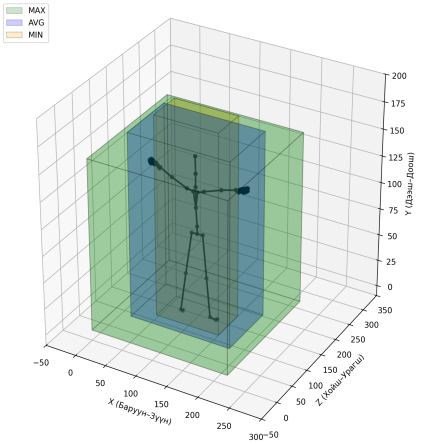


Fig. 15. Үндсэн 10 хөдөлгөөний орон зай /MAX, AVG, MIN /

Гар болон хөлийн орон зайн харьцуулалт:

Гарын орон зай: гарын хөдөлгөөн нь дунджаар 140-220 см-ийн өргөн хүрээг (X_range) хамран ажиллаж байна. Ялангуяа 3 болон 8 дугаар хөдөлгөөнүүд (X_range: 222 см, 219 см) нь гарын хамгийн их дэлгэмэл далайцтай хэсгүүд юм.

Хөлийн орон зай: хөлийн орон зайн ашиглалт гараас **1.36** дахин бага байна. Энэ нь үзэмчин бүжигт хөл нь

ихэвчлэн биеийн хүндийн төвийг барих, тогтвортой суурь болох үүрэг гүйцэтгэдэг бол гар нь орон зайг эзэлсэн зориудын тоглолыг (Expressiveness) бүтээх гол хэрэгсэл болдгийг харуулж байна.

Орон зайн шинжилгээний үр дүнгээс үзэхэд үзэмчин бүжиг нь гарын хөдөлгөөний далайц нь хөлийнхөөс эрс өндөр байгаа нь биеийн дээд хэсэг илүү идэвхтэй динамиктай, харин доод хэсэг нь тогтворжуулагч (Stabilizer) үүрэгтэй оролцдог монгол бүжгийн "Төв цэгт суурилсан" (Center-based) хөдөлгөөний логикийг тоон үзүүлэлтээр харуулав[9]. Хөдөлгөөн 8-ийн гар, хөлийн орон зай ашиглалтын харьцааг Зураг 16-д үзүүлэв.

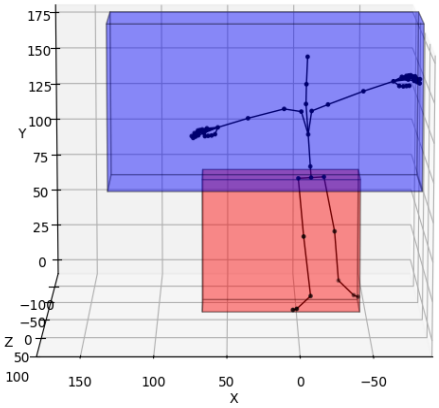


Fig. 16. Хөдөлгөөний явц дахь гар болон хөлийн орон зайн ашиглалтын эзлэхүүн (Хөд 8)

5) Гурван цэгийн өнцгийн тооцоолол

Гурван цэгийн өнцгийн тооцооллыг хөдөлгөөний нарийвчлалыг шалгах зорилгоор гарын гол барилтын хэсэгт туршиж үзэхэд уг хөдөлгөөний үе дэх тохойн нугаралтын өнцөг тогтвортой 45 градус байгаа нь тогтоогдсон. Энэхүү аргачлалыг ашиглан биеийн аль ч хэсгийн (өвдөг, тохой, мөр) өнцгийн шилжилтийг нарийвчлан тооцох бүрэн боломжтой бөгөөд гаргаж авсан эдгээр өнцгийн утгууд нь ирээдүйд бүжгийн хөдөлгөөний гүйцэтгэлийг техникийн хувьд "стандарт" эсвэл "стандарт бус" гэж үнэлэх, сургалтын дижитал стандарт (Baseline) тогтоох шинжлэх ухааны бодит үндэслэл болох юм.

6) Гол тэнхлэгийн (Principal Axis) илрүүлэлт

Судалгааны хүрээнд бүжигчний орон зайн баримжаа болон биеийн тэнхлэгийн хазайлтыг тодорхойлох зорилгоор PCA (Principal Component Analysis) алгоритмыг ашиглан 10 үндсэн хөдөлгөөний гол тэнхлэгүүдийг урьдчилсан байдлаар тогтоов. Туршилтад өөр өөр өндөртэй хоёр бүжигчний (Б: 165 см, Э: 153 см) өгөгдлийг харьцуулахад Б бүжигчний хувьд тэнхлэгийн утгууд 159, 123, 121, 104, 67 зэрэг үзүүлэлттэй гарсан бол Э бүжигчнийх 122, 110, 51, 0 гэсэн утгуудад төвлөрч байна. Эдгээр тэнхлэгүүдийн тархалтыг Histogram болон 3D орон зайн тэмдэглэгээгээр шинжлэхэд Үзэмчин бүжигчин хөдөлгөөн нь биеийн дээд мөрийн тэнхлэгт суурилсан болох нь батлагдсан юм. Энэхүү тэнхлэгийн илрүүлэлт нь бүжигчний биеийн тэнцвэрийн шугам болон хөдөлгөөний ерөнхий чиг хандлагыг тоон утгаар тодорхойлж байгаа бөгөөд цаашид бүжгийн гүйцэтгэлийн "цэвэр" байдлыг үнэлэх бодит хэмжүүр болох ач холбогдолтой.

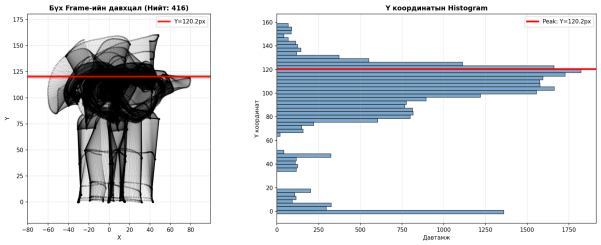


Fig. 17. Тэнхлэг илрүүлэлтийн жишээ (хөдөлгөөн 1)

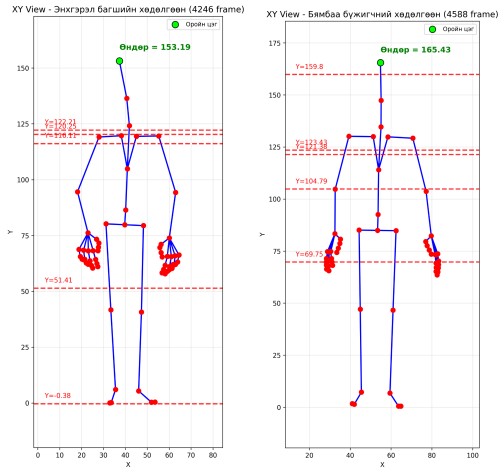


Fig. 18. 10 тэнхлэгийн зураглал

V. Хэлэлцүүлэг

Үзэмчин бүжгийн кинематик үзүүлэлтүүд нь уг бүжгийн хөдөлгөөний гүйцэтгэл болон соёлын онцлогийг шинжлэх ухааны үндэслэлтэйгээр баталж байна. Тухайлбал, хурдны дээд утга 1.55 м/сек хүрч байгаа нь урлаг судлаачдын тодорхойлдог "Монгол бүжгийн эрч хүчтэй цохилт, агших-тэнийх динамик"-ийг тоон утгаар илэрхийлсэн хэрэг юм. Энэ нь сонгодог балетын жигд хэмнэлээс [16] ялгаатай нь харагдаж байна. Орон зайн ашиглалтын хувьд хөлийн хөдөлгөөн 100% доод түвшинд байгаа нь Индонезийн уламжлалт бүжгүүдтэй [11] төстэй "газардуулга" (groundedness)-ыг харуулж байгаа боловч гарын далайц нь биеийн хүндийн төвөөс маш алслагдсан (Peripheral) байгаа нь Үзэмчин бүжгийн "дэлгэмэл" шинжийг харуулсан. Лабаны онолын Weight (0.34) үзүүлэлт нь уг бүжгийг Фламенко [17] шиг "хүнд" биш, харин "хөнгөн-удаавтар" буюу агаарлаг/хөвөх шинжтэйг илтгэж байна.

Энэхүү судалгааны хязгаарлагдмал тал нь туршилтын орчин лабораторийн нөхцөлд бүжиглэсэн нь бүжигчний тайзны орон зайг эзлэн бүжиглэх хэлбэрээс тодорхой хэмжээгээр ялгагдах боломжтой. Гэхдээ үзэмчин бүжгийн эмэгтэй хөдөлгөөн орон зайн тоглол гэхээсээ илүүтэй үзэгчид рүү хандсан урагшаа чиглэлтэй, зөөлөн жижиг алхаагаар бага орон зай ашиглан гар, цээжин бие, эрүү шанааны зохиомж давамгайлдаг тул дээрх хэмжилтийн үзүүлэлтэд

нөлөөлөх нь бага гэж үзэж байна. Гэсэн хэдий ч OptiTrack систем ашиглан гаргаж авсан эдгээр өгөгдөл нь цаашид Монгол бүжгийн хөдөлгөөнийг AI-аар ангилах, сургалтын стандарт загвар (Baseline) тогтооход суурь үзүүлэлт болно.

VI. Дүгнэлт ба ирээдүйн ажил

Энэхүү судалгааны ажлаар үзэмчин угсаатны бүжгийн хөдөлгөөний цараг OptiTrack систем ашиглан дижитал хэлбэрт шилжүүлж, кинематик болон чанарын шинжилгээ хийх салбар дундын арга зүйг амжилттай туршлаа. Судалгааны үр дүнд үзэмчин бүжиг нь доод бие (ташаа, хөл)-ийн тогтвортой-удаан, дээд бие (цээж, мөр, гар)-ын дэлгэмэл, уян налархай хөдөлгөөний нэгдэлд суурилсан, "хөнгөн-үргэлжилсэн" шинж чанартай болохыг тоон үзүүлэлтээр тулгав. Python-д суурилсан BVH Parser болон биомеханик тооцооллын алгоритмууд нь бүжгийн урлагийг зөвхөн ажиглалтын төдий бус, шинжлэх ухааны үндэслэлтэй математик утгаар тооцоолов.

Цаашид судалгааны ажлыг гүнзгийрүүлэн дараах техникийн шинжилгээнүүдийг нэмэлтээр гүйцэтгэхээр төлөвлөж байна. Үүнд:

- Биеийн хүндийн төвийн (CoM) шилжилт: Бүжигчний тэнцвэр хадгалалт болон хүндийн төвийн динамик тогтворжилтыг тодорхойлох.
- Өлмийн өргөлт, толгойн шилжилт, эрүү шанааны өргөлт, хаялтын динамик: Үзэмчин бүжгийн өвөрмөц техник болох өлмий дээрх тогтвортой үргэлжилсэн хөдөлгөөн, эрүү, шанааны жишүү өргөх өнцөг болон хурдатгалыг хэмжих.

Энэхүү судалгааны хүрээнд бүрдүүлсэн хөдөлгөөний сан (Mocap Dataset) болон тодорхойлсон кинематик шинж чанаруудыг олон талаар ашиглах боломжтой. Ийнхүү дижитал эринд үндэсний өвийг технологийн аргаар баримтжуулж, шинжлэх ухааны эргэлтэд оруулснаар соёлын биет бус өвийг хамгаалах, сурталчлах шинэ платформ бүрдэж байна

Талархал

Энэхүү судалгааны ажлыг гүйцэтгэхэд мэргэжлийн арга зүйн зөвлөгөө өгч, дэмжлэг үзүүлсэн бүх хүмүүст гүнээ талархаж буйгаа илэрхийлье. Судалгааны туршилт, өгөгдөл цуглуулах орчин болон техник хэрэгслээр хангаж, Motion Capture судалгааны лабораторио ашиглуулсан Шинжлэх Ухаан Технологийн Их Сургуулийн Мэдээлэл, Холбооны Технологийн Сургуулийн хамт олон, лабораторийн эрхлэгч багш А.Эрдэнэбаяр (Ph.D) болон туршилтын бүхий л үед техникийн туслалцаа үзүүлсэн Мультимедиа хөтөлбөрийн оюутнуудад баярлалаа. Мөн өгөгдөл боловсруулах, шинжилгээ хийх боломжоор хангасан Компьютер график, мультимедиа судалгааны нээлттэй лабораторийн хамт олонд талархал илэрхийлье.

Үзэмчин бүжгийн өгөгдөл бэлтгэх, хөдөлгөөний намба төрхийг дижитал хэлбэрт зөв зааварчлан үлдээхэд үнэтэй хувь нэмэр оруулсан:

- Уран сайхны удирдагч, бүжиг дэглээч, МУГЖ Д.Энхгэрэл,
- Үндэсний урлагийн их театрын бүжигчин, МУГЖ М.Бямбаа,
- "Хан Хэнтий" чуулга, СУИС-ийн Бүжгийн урлагийн сургуулийн оюутнууд болон нийт оролцогчдод гүнээ талархаж байна.

Та бүхний дэмжлэг, хамтын ажиллагаа Үзэмчин бүжгийг дижитал эринд шинжлэх ухааны үндэслэлтэй хадгалан үлдэх нарийвчлан шинжлэх энэхүү судалгааны ажлын гол үндэс болсон билээ.

References

[1] N. Kadel, "The Biomechanics of Ballet: Kinematic Analysis of Classical Dance Movements," Journal of Dance Medicine & Science, vol. 10, no. 3, 2006.

[2] R. S. Linthorst et al., "Kinematic analysis of the legs during the basic step in Latin American dancing," Sports Biomechanics, vol. 14, no. 2, 2015.

[3] M. P. Scharer, "Comparative Kinematics: Contemporary vs. Traditional Dance Styles," International Journal of Movement Science, 2021

[4] S. Sato et al., "Symmetry and Asymmetry in Dance Movements: A Kinematic Approach," Journal of Human Movement Studies, 2018.

[5] R. Laban, "The Mastery of Movement," Macdonald & Evans, 1980.

[6] K. Stevens, "Spatial orientation and directionality in traditional dance styles," Dance Research Journal, vol. 34, no. 1, 2012.

[7] J. Lopez et al., "Biomechanical analysis of flamenco dance: Weight and Time effort," International Journal of Arts and Technology, 2019.

[8] M. Guest, "Labanotation: The System of Analyzing and Recording Movement," Routledge, 2005.

[9] Д. Нанжид, Монгол бүжгийн хөдөлгөөний нэр томьёо, арга зүй. Улаанбаатар, Монгол: Адмон, 2000.

[10] L. Wang et al., "Chang-E: A High-Quality Motion Capture Dataset of Chinese Classical Dunhuang Dance," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023.

[11] A. Md Nasir et al., "Digital Preservation of Joget Dance using Motion Capture Technology," International Journal of Interactive Digital Media, vol. 5, no. 1, pp. 12-17, 2017.

[12] F. Cavallo et al., "A Kinematic Dataset of Actors Expressing Emotions through Facial Expressions and Body Movements," Scientific Data, vol. 7, no. 1, 2020.

[13] R. Parent, Computer Animation: Algorithms and Techniques, 4th ed. Morgan Kaufmann, 2017 (Motion Capture Technology for Character Animation in Film and Games).

[14] A. D. Al-Abboodi et al., "Application of OptiTrack Motion Capture Systems in Human Movement Analysis: A Systematic Literature Review," IEEE Access, vol. 10, pp. 12450-12465, 2022.

[15] J. Hamill and K. M. Knutzen, Biomechanical Basis of Human Movement, 5th ed. Lippincott Williams & Wilkins, 2021 (Three-Dimensional Analysis of Human Movement).

[16] E. P. Dozza et al., "Spatio-temporal Analysis of Kinematic Signals in Classical Ballet," Journal of Biomechanics, vol. 112, 2020.

[17] T. Pengjia et al., "Analysis and description of dance movements based on sports biomechanics analysis method," Journal of Xi'an University of Technology, vol. 37, no. 4, pp. 542-550, 2021.

[18] Y. Zhang et al., "Feature Data Analysis of Dance Movements by Motion Capture and ST-GCN," Journal of Physics: Conference Series, vol. 1646, 2020.

[19] H. Huang et al., "Learning to Dance: A Graph Convolutional Adversarial Network for Generating Dance Movements," ACM Transactions on Graphics (TOG), vol. 39, no. 6, 2020.

[20] X. Liu, "Application of Real-time Motion Capture Technology in Street Dance Movement Analysis and Optimization," International Journal of Emerging Technologies in Learning (iJET), vol. 15, no. 18, 2020.

[21] K. Jensen et al., "Motion Capture and AI in Dance Performance Analysis: Technical and Artistic Evaluation," Digital Creativity, vol. 32, no. 2, pp. 85-102, 2021.

[22] Gabbard, C., & Hart, **S.** (2020). Biomechanical Analysis of Selected Classical Ballet Movements. Journal of Dance Medicine & Science, 24(3), 112-118.

Монголын хадны зураг дахь амьтдын дүрслэлийг YOLOv8 загвараар илрүүлэх нь

Пэрэнлэйлхүндэв Гантуяа
Мэдээлэл, компьютерын ухааны тэнхим
Мэдээллийн технологи, электроникийн сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол Улс
ORCID: 0009-0000-7332-141X
gantuya.p@num.edu.mn

Рэнцэндорж Элчинбат
Мэдээлэл, компьютерын ухааны тэнхим
Мэдээллийн технологи, электроникийн сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол Улс
24elchee@gmail.com

Батсуурь Сувдаа
Мэдээлэл, компьютерын ухааны тэнхим
Мэдээллийн технологи, электроникийн сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол Улс
ORCID: 0000-0002-7051-4256
suvdaa@num.edu.mn

***Хураангуй*—** Энэхүү өгүүлэлд Монголын хадны зурагт дүрслэгдсэн амьтдын дүрсийг компьютер харааны аргаар автоматаар илрүүлэх туршилтын үр дүнг танилцуулав. Судалгаанд хадны зураг дээр түгээмэл ажиглагддаг морь, хүн, буга, янгир гэсэн дөрвөн ангиллыг сонгон Roboflow платформ дээр хүрээлэн тэмдэглэж (bounding box annotation) хийж, YOLOv8n болон YOLOv8s загваруудаар сургалт явуулсан. Эхний шатанд янгирын нэг ангилалтай туршилт хийж, дараа нь өгөгдлийн санг өргөтгөн олон-ангилалт (multiclass) загварт шилжив. Эцсийн хувилбарт нийт 367 зураг ашиглаж, 321/31/15 train/valid/test хуваарилалттайгаар 640×640 хэмжээ, хэвтээ тэнхлэгийн ойлт, бага өнцгийн эргүүлэлт, гэрэлтүүлэг болон мөлийлгөх аргаар нэмэгдүүлэн хэрэглэв. YOLOv8s загвар 100 epoch сургалтын дараа нийт ангиллын хувьд буга, янгир ангиллууд хүн, мориноос илүү тогтвортой танигдсан. Үр дүнгээс харахад Монголын хадны зураг дахь амьтдын дүрслэлийг автомат аргаар илрүүлэх боломжтой боловч өгөгдлийн түүврийн тоо цөөн, ангиллын түүврүүдийн тоо тэнцвэргүй байдал, дүрслэлийн стилизац, элэгдэл зэргээс шалтгаалан илрүүлэх нарийвчлалыг сайжруулах шаардлагатай байна. Сайжруулсан хувилбарт анхны дөрвөн ангиллын туршилтаас гадна background class-ийн нөлөө, зургаан ангиллын annotation refinement, зэвсэгтэй хүн болон морь унасан хүн зэрэг нийлмэл дүрслэлийн ангилал, мөн HSV өнгөний орон зайд суурилсан preprocessing туршилтын үр дүнг нэмж нэгтгэв.

Түлхүүр үг*—хадны зураг, объект илрүүлэлт, YOLOv8, Roboflow, компьютер хараа, соёлын өвийн дижиталжуулалт*

I. УДИРТГАЛ

Монголын хадны зураг нь Евразийн хээр, говь, уулсын бүсэд олон мянган жилийн турш бүтээсэн дүрслэлийн өвийг хадгалсан, археологи болон урлаг судлалын чухал эх сурвалж юм. Буга, янгир, морь, хүн, ан агнуурын үзэгдэл, ёслолын дүрслэл зэрэг нь Монголын хадны зурагт өргөн тохиолддог бөгөөд эдгээр нь эртний нүүдэлчдийн аж ахуй, шүтлэг, орон зайн хэрэглээ, дүрслэлийн хэв маягийг судлахад чухал мэдээлэл өгдөг [15], [16], [19], [20].

Уламжлалт хадны зургийн судалгаа нь ихэвчлэн дүрсийн тайлал, он цаг, бүс нутгийн хэв шинж, археологийн харьцуулалт дээр тулгуурладаг. Гэвч олон тооны дүрслэлийг гараар бүртгэх, ангилах, тоолох, бүс нутгаар нь харьцуулах ажил нь цаг их шаарддаг бөгөөд томоохон дүрсний сан бий болгоход хүндрэлтэй байдаг. Иймээс соёлын өвийн дижиталжуулалт, дүрс боловсруулалт, гүн сургалтын арга нь хадны зургийн автомат танилт, бүртгэл, ангилалд ашиглагдах боломжтой болж байна [10], [11], [12], [14].

Энэхүү судалгааны зорилго нь Монголын хадны зураг дахь амьтдын дүрслэлийг YOLOv8 суурьт объект илрүүлэлтийн загвараар таних боломжийг турших, өгөгдлийн сан бэлтгэх ажлын урсгалыг боловсронгуй болгох, туршилтын үр дүнг тоон үзүүлэлт болон дүрсийн жишээгээр шинжлэхэд оршино. Судалгааны хүрээнд эхлээд янгирын нэг ангилалтай туршилт хийж, дараа нь адуу, хүн, Буга, янгир гэсэн дөрвөн ангилалтай multiclass хувилбар руу шилжин өгөгдлийг өргөтгөн туршсан.

II. ӨМНӨХ СУДАЛГАА БА ХАРЬЦУУЛАЛТ

Сүүлийн жилүүдэд хадны зураг болон археологийн дүрсийн шинжилгээнд машин сургалт, гүн сургалтын арга идэвхтэй ашиглагдаж байна. Jalandoni нар [10] хадны зургийн судалгаанд машин сургалт ашиглах боломж, хязгаарлалтыг авч үзэж, Австралийн хадны зургийн мэдээлэл дээр painted rock art identification-ийн proof-of-concept туршилт хийсэн. Turner-Jones [11] нь рок арт motif recognition болон Indigenous rock art education чиглэлд machine learning-ийн хэрэглээг тайлбарласан. Huo нар [12] Cangyuan rock art дээр сайжруулсан YOLOv7-Seg загвар хэрэглэн хүн болон амьтны дүрсийг segmentation аргаар ялгаж өндөр AP үзүүлэлтэд хүрсэн. Мөн Gil нар [13] хадны дүрс илрүүлэлтэд уламжлалт тоон дүрс боловсруулалт хэрэглэсэн нь автомат танилтын эрт үеийн жишээ болж байна.

Археологийн дүрс дээр объект илрүүлэлт ашигласан бусад судалгаанд Faster R-CNN [6], Mask R-CNN [7], TensorFlow Object Detection API болон transfer learning-д суурилсан арга өргөн хэрэглэгддэг. Mnasri ба D'Andrea [14] археологийн олдворын зураг дээр объект илрүүлэх болон ангилал хийхийг хослуулан автомат inventory хийх арга боловсруулсан нь соёлын өвийн өгөгдөлд detection pipeline ашиглахын ач холбогдлыг харуулсан.

Монголын хадны зурагт төвлөрсөн гүн сургалтын судалгаа харьцангуй цөөн тул энэхүү ажил нь Монголын соёлын өвийн бодит өгөгдөл дээр YOLOv8 суурьт илрүүлэх туршилт хийсэн анхны системчилсэн алхмуудын нэг гэж үзэж болно. Судалгааны онцлог нь нэг ангиллын туршилтаас эхлэн олон ангилалт илрүүллэлт хүртэл шаталсан туршилт хийсэн, annotation workflow-оо Roboflow орчинд тогтворжуулсан, мөн Монголын хадны зурагт түгээмэл ажиглагддаг амьтдын ангиллаар бодит dataset бүрдүүлсэнд оршино.

TABLE I. Өмнөх судалгааны чиг хандлагын харьцуулалт

Судалгаа	Өгөгдөл	Арга	Гол чиглэл
Jalandoni et al. (2022)	Australian painted rock art	Deep learning classification	Rock art presence identification
Turner-Jones (2023)	Rock art motifs	Machine learning	Motif recognition / digital хүnities
Huo et al. (2025)	Cangyuan rock art	Improved YOLOv7-Seg	Хүн/animal segmentation
Gil et al. (2011)	Petroglyph images	Digital image processing	Petroglyph detection
Энэхүү судалгаа	Монголын хадны зураг	YOLOv8n / YOLOv8s	Амьтдын объект илрүүлэлт

III. ӨГӨГДЛИЙН САН БА АРГА ЗҮЙ

Судалгааны ерөнхий ажлын урсгалыг Зураг 1-д үзүүлэв. Судалгааны бүх шатанд зураг цуглуулах, тэмдэглэгээ хийх, хувилбар үүсгэх, нэмэгдүүлэлт хэрэглэх, YOLO загвар сургах, үр дүнг үнэлэх гэсэн дараалал баримталсан.

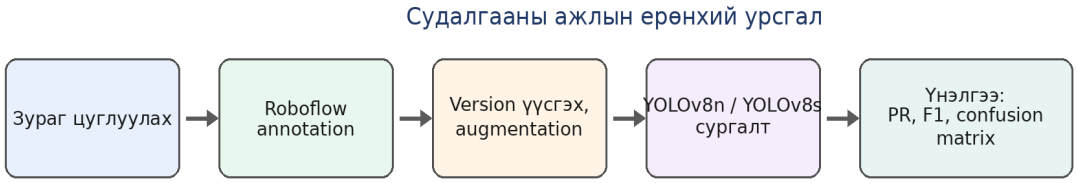


Fig. 1. Судалгааны ажлын ерөнхий урсгал

Эхний шатанд хадны зургийн ном, хэвлэмэл эх сурвалж, өмнөх судалгааны зурагт хавсралт, талбайн зураглалаас авсан дүрсүүдийг нэгтгэн анхны мэдээллийн сан бүрдүүлсэн. Зургийг дижитал хэлбэрт оруулахдаа гар утас болон scan document функц ашиглаж, дараа нь Roboflow дээр annotation хийсэн. Анхны туршилтад нийт 107 зураг ашигласан бөгөөд Roboflow-ийн duplicate detection-оор 62 зураг давхардсан болох нь илэрч, өгөгдлийг цэвэрлэсэн.

Дараагийн шатанд өгөгдлийн санг өргөтгөн 367 зурагтай multiclass хувилбар (version v3) үүсгэсэн. Энэ хувилбарт адуу, хүн, Буга, янгир гэсэн дөрвөн ангиллыг ашигласан. Нэг зураг дээр олон объект байвал бүгдийг нь тусад нь bounding box-оор тэмдэглэсэн. Морь унасан хүний дүрс илэрсэн тохиолдолд адуу ба хүн ангиллыг давхар бус, харин тус тусын box-оор тэмдэглэсэн.

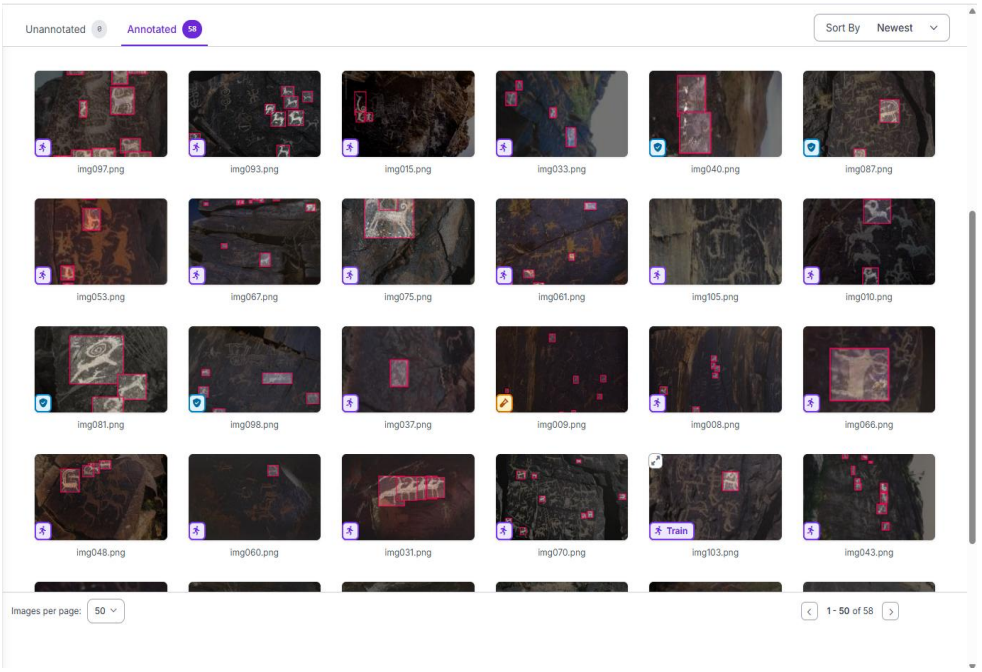


Fig. 2. Roboflow орчинд хийсэн анхны annotation-ийн жишээ (58 annotated зурагтай эхний шат)

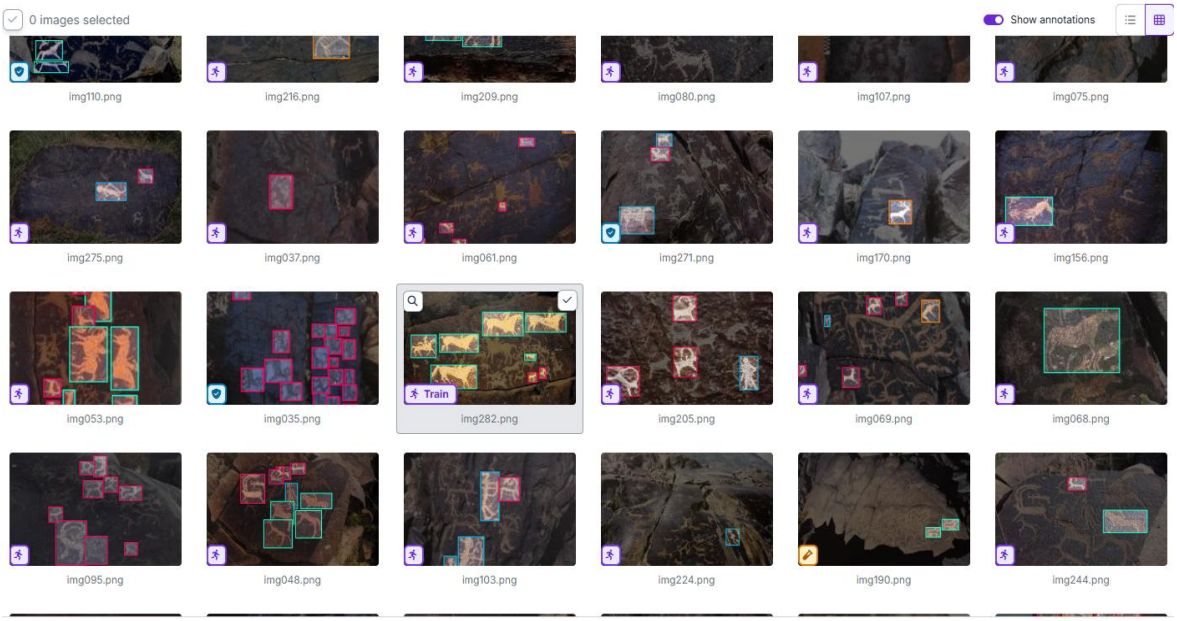


Fig. 3. Өргөтгөсөн multiclass dataset-ийн annotation харагдац

TABLE II. Эцсийн multiclass dataset-ийн бүтэц

Үзүүлэлт	Утга
Нийт зураг	367
Train set	321
Validation set	31

Үзүүлэлт	Утга
Test set	15
Ангиллын тоо	4
Ангилал	Адуу, Хүн, Буга, Янгир

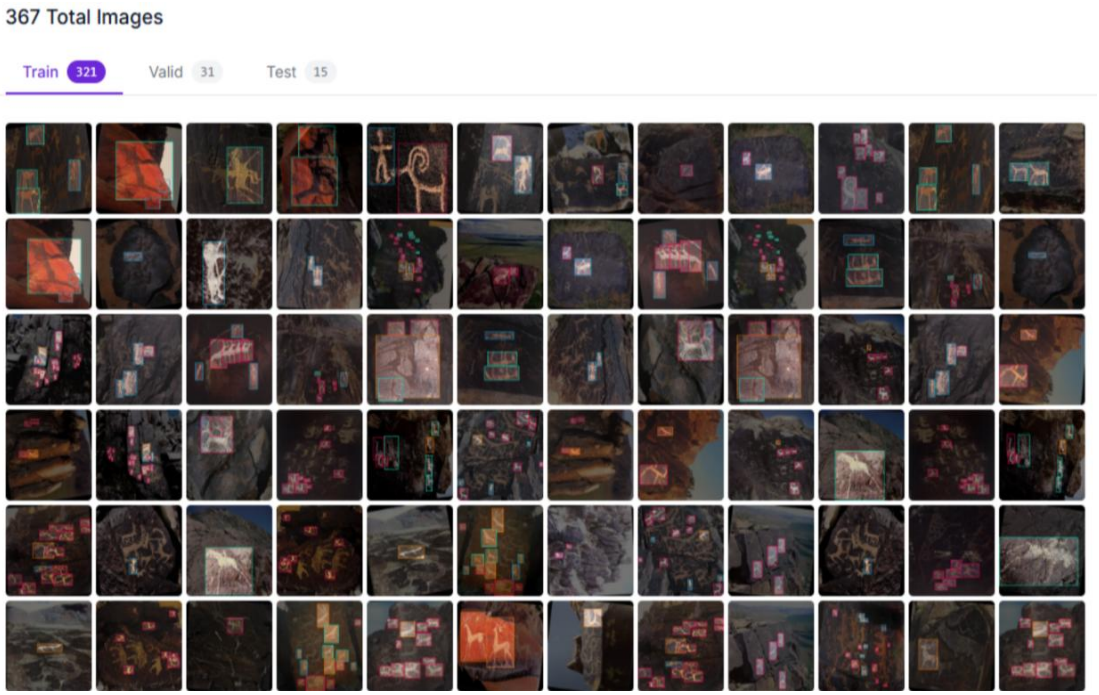


Fig. 4. Train / validation / test split-ийн бүтэц

А. Урьдчилсан боловсруулалт ба augmentation

Roboflow version үүсгэх шатанд auto-orient болон 640×640 resize ашиглав. Мөн сургалтын өгөгдлийн ерөнхийшүүлэх чадварыг нэмэгдүүлэхийн тулд horizontal flip, ±10° rotation, brightness-ийн ±10% өөрчлөлт, 0.5 px хүртэл blur augmentation хэрэглэсэн.

TABLE III. Version үүсгэх үед хэрэглэсэн тохиргоо

Параметр	Тохиргоо
Resize	640 × 640
Flip	Horizontal
Rotation	-10° – +10°
Brightness	-10% – +10%
Blur	0.5 px хүртэл
Outputs per training example	3

B. Загвар ба сургалтын орчин

Туршилтыг Google Colab орчинд Tesla T4 GPU ашиглан гүйцэтгэсэн. Загварын хувьд YOLOv8n болон YOLOv8s суурь хувилбаруудыг сонгож харьцуулсан. YOLOv8 нь Ultralytics экосистемийн real-time object detection загвар бөгөөд хурд, тохиргооны уян хатан байдал, transfer learning-д суурилсан сургалтын боломж зэргээрээ жижиг болон дунд хэмжээний dataset дээр туршилт хийхэд тохиромжтой [3], [4].

TABLE IV. Сургалтын үндсэн тохиргоо

Үзүүлэлт	Утга
Загвар	YOLOv8n, YOLOv8s
Input хэмжээ	640 × 640
Batch size	8
Epoch	50 ба 100
Device	Tesla T4 GPU
Prediction confidence	0.20–0.30 туршсан

IV. ТУРШИЛТЫН ЗОХИОН БАЙГУУЛАЛТ

Туршилтыг гурван үе шаттай хийв. Нэгдүгээрт, янгир ангиллыг дангаар нь авч detection хийх урьдчилсан туршилт явуулсан. Энэ шатны зорилго нь annotation workflow болон YOLO сургалтын ерөнхий боломжийг шалгах явдал байв. Хоёрдугаарт, 367 зурагтай multiclass dataset дээр YOLOv8n загварыг 50 epoch сургаж анхны baseline үзүүлэлтүүдийг гаргав. Гуравдугаарт, мөн dataset дээр YOLOv8s загварыг 100 epoch сургаж YOLOv8n-тэй харьцуулсан.

Ийнхүү нэг ангиллын туршилтаас олон ангиллын detection руу шилжсэн нь өгөгдлийг аажмаар өргөтгөж, annotation дүрмээ тогтворжуулахад ач холбогдолтой байв. Эцсийн загварын үнэлгээнд precision, recall, mAP@0.5, mAP@0.5:0.95, confusion matrix, PR curve, F1-confidence curve, train/validation loss curve-үүдийг ашиглав.

TABLE V. Хийсэн үндсэн туршилтуудын тойм

Туршилт	Dataset	Загвар	Epoch	Гол ажиглалт
Pilot-A	Анхны янгир-only dataset	YOLOv8n	урьдчилсан run	Annotation workflow ба detection боломжийг шалгасан
Exp-1	367 зурагтай 4 class v3	YOLOv8n	50	4 class baseline multiclass үр дүн
Exp-2	367 зурагтай 4 class v3	YOLOv8s	100	YOLOv8n-тэй харьцуулсан сайжруулалт
Exp-3	Refined 6 class dataset	YOLOv8n	30-45	armed_human, horse_rider зэрэг complex class нэмсэн
Exp-4	6 class + HSV preprocessing	YOLOv8n	30	V channel contrast stretching preprocessing-ийн нөлөө

V. ТУРШИЛТЫН ҮР ДҮН

Энэхүү шат нь нэг ангиллын detection боломжтой ч объектын ихэнх хэсгийг background руу алдах эрсдэл өндөр байгааг харуулсан. Иймээс илүү өргөн dataset болон multiclass сургалт шаардлагатай гэж үзсэн. Multiclass туршилтын үед YOLOv8n 50 epoch болон YOLOv8s 100 epoch-ийн үр дүнг Хүснэгт 6-д харьцууллаа. YOLOv8s загвар precision-ийн хувьд YOLOv8n-ээс бага зэрэг буурсан боловч recall болон mAP@0.5 үзүүлэлтээр илүү сайн байсан тул эцсийн дүн шинжилгээнд YOLOv8s-ийн үр дүнг голчлон ашиглав.

TABLE VI. Multiclass туршилтын ерөнхий үзүүлэлтүүд

Туршилт	Precision	Recall	mAP@0.5	mAP@0.5:0.95
YOLOv8n, 50 epoch	0.299	0.259	0.204	0.115
YOLOv8s, 100 epoch	0.257	0.348	0.228	0.127

TABLE VII. YOLOv8s (100 epoch) ангилал тус бүрийн үр дүн

Ангилал	Precision	Recall	mAP@0.5	mAP@0.5:0.95
Адуу	0.210	0.286	0.168	0.085
Хүн	0.074	0.207	0.103	0.073
Буга	0.392	0.571	0.359	0.182
Янгир	0.351	0.330	0.280	0.167

A. Prediction жишээний шинжилгээ

Зураг 5-д model харьцангуй зөв таньсан prediction-ийн жишээнүүдийг үзүүлэв. Буга болон янгирын хувьд эвэр, нурууны хэлбэр, биеийн ерөнхий дүрс нь ангилал ялгахад чухал үүрэг гүйцэтгэж байв. Зарим зураг дээр хүн ангилал бага confidence-тай ч зөв илэрсэн нь жижиг босоо дүрслэлүүдийг тодорхой хэмжээнд сурсан болохыг харуулж байна.



Fig. 5. Зөв болон харьцангуй амжилттай prediction-ийн жишээнүүд

Зураг 6-д илрүүлэхэд бэрхшээлтэй тохиолдлуудыг харуулав. Энд адуу ба янгир хооронд андуурал үүсэх, нэг дүрсэнд давхар box үүсэх, томруулсан box-оор дүрсийн хүрээ хэтэрхий өргөн тэмдэглэгдэх зэрэг асуудал ажиглагдав. Эдгээр нь annotation consistency болон дүрслэлийн стилизац detection-ийн чанарт хүчтэй нөлөөлж байгааг илтгэнэ.

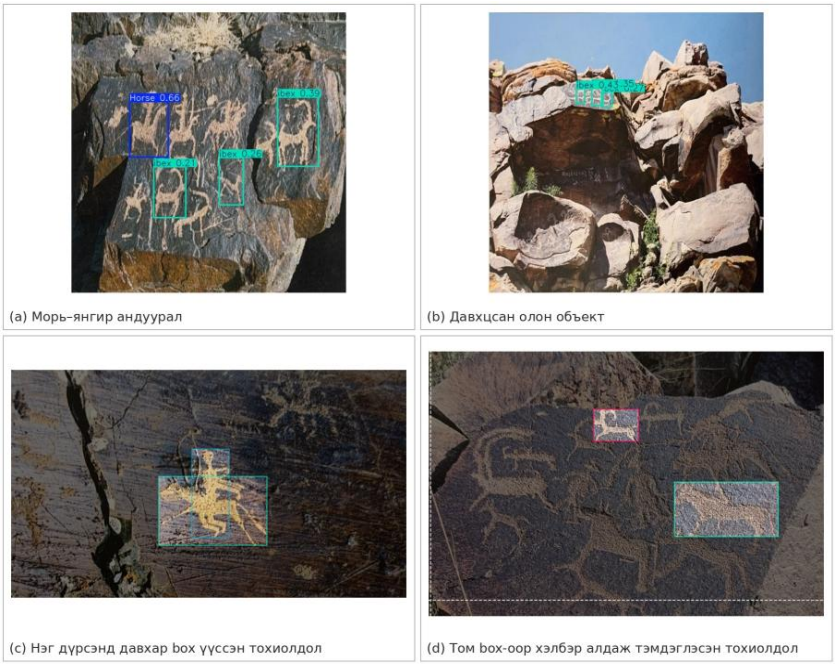


Fig. 6. Хүнд болон буруу prediction-тэй жишээнүүд

B. Нэмэлт туршилт ба сайжруулсан хувилбар

Нэг ангиллын болон дөрвөн ангиллын туршилтын дараа өгөгдлийн тэмдэглэгээг дахин нягталж, илүү нарийн ангиллын түвшинд турших нэмэлт шат хийв. Энэ шатанд өмнөх адуу, хүн, буга, янгир гэсэн дөрвөн ангиллаас гадна зэвсэгтэй хүн (armed_human) болон морь унасан хүн (horse_rider) гэсэн нийлмэл дүрслэлийг тусдаа class болгон оруулсан. Ингэснээр хадны зураг дахь энгийн амьтны дүрслэлээс гадна хүний үйл ажиллагаа, хөдөлгөөнт нийлмэл зохиомжийг илрүүлэх боломжийг шалгасан.

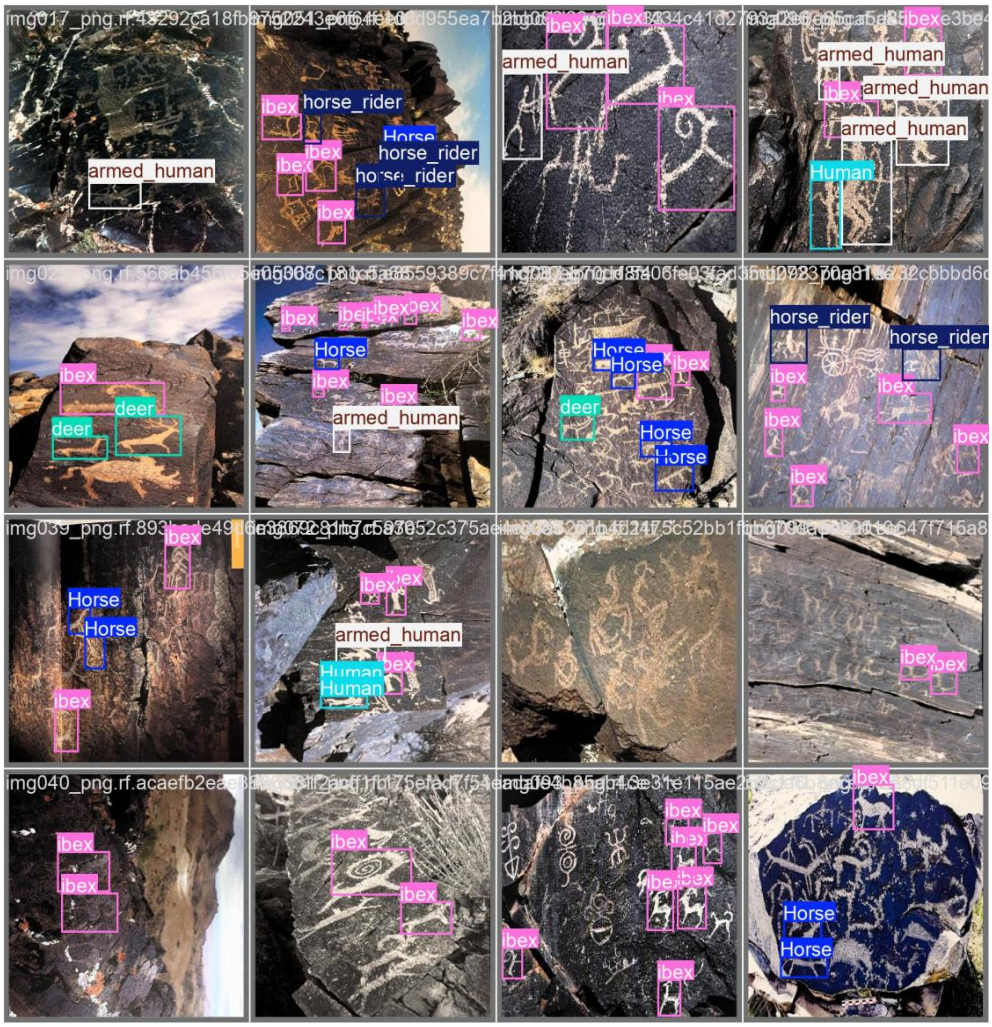


Fig. 7. Зургаан ангиллын олон объекттой prediction жишээнүүд

C. HSV preprocessing туршилтын арга

Дөрвөн болон зургаан ангиллын туршилтын явцад хадны гадаргуугийн сүүдэр, элэгдэл, хагарал, өнгөний өөрчлөлт нь detection-д хүчтэй нөлөөлж байв. Иймээс dataset-ийн хэмжээг богино хугацаанд огцом нэмэгдүүлэх боломжгүй нөхцөлд зургийн ялгаралтыг сайжруулах preprocessing туршилт хийсэн. Багшийн зөвлөсөн histogram equalization, contrast stretching, rgb2hsv хувиргалтын санаануудаас RGB -> HSV -> V channel -> contrast stretching -> RGB гэсэн workflow-г сонгосон. Энэ арга нь зургийн геометрийг өөрчлөхгүй тул өмнө хийсэн bounding box annotation-ийг дахин засах шаардлагагүй давуу талтай.

D. HSV preprocessing туршилтын тоон үр дүн

TABLE VIII. HSV preprocessing туршилтын ерөнхий үзүүлэлтүүд

Үзүүлэлт	Утга
Precision	0.231
Recall	0.319
mAP@0.5	0.203
mAP@0.5:0.95	0.121
Best F1	0.23 орчим
Best confidence	0.09 орчим

TABLE IX. HSV preprocessing туршилтын class-wise mAP үзүүлэлт

Class	Монгол тайлбар	mAP@0.5	mAP@0.5:0.95
Horse	Морь	0.021	0.012
Human	Хүн	0.116	0.048
armed_human	Зэвсэгтэй хүн	0.504	0.310
deer	Буга	0.067	0.039
horse_rider	Морь унасан хүн	0.161	0.103
ibex	Янгир	0.349	0.213

TABLE X. Baseline ба HSV preprocessing-ийн харьцуулалт

Арга	Precision	Recall	mAP@0.5	mAP@0.5:0.95	Тайлбар
Baseline original dataset	0.35	0.24	0.180	0.110	Preprocessing хийхээс өмнөх суурь үр дүн
HSV preprocessing	0.231	0.319	0.203	0.121	V channel contrast stretching хийсний дараах үр дүн

Horse, deer, horse_rider зэрэг дөрвөн хөлтэй дүрслэлүүд хэлбэрийн хувьд төстэй тул андуурал үүсэх боломж өндөр. Мөн human болон armed_human ангиллууд жижиг, бүдэг, заримдаа давхцсан дүрслэлтэй байсан нь илрүүлэлтийг хүндрүүлсэн.

E. Нэмэлт prediction жишээний ажиглалт

Prediction жишээнүүдээс model тод харагдсан ibex болон horse_rider зэрэг class-уудыг зарим тохиолдолд 70-90 хувийн confidence-тэй зөв илрүүлж байв. Гэхдээ энэ нь нийт accuracy гэсэн үг биш бөгөөд object detection-ийн нийт гүйцэтгэлийг бүх class, бүх зураг, false positive, false negative болон IoU нөхцөлтэй хамтатган mAP үзүүлэлтээр үнэлдэг. Иймээс зарим нэг зураг дээр confidence өндөр байсан ч нийт mAP бага гарах нь өгөгдлийн төвөгтэй байдал, class imbalance, background noise, annotation ambiguity зэрэгтэй холбоотой.

VI. ХЭЛЭЛЦҮҮЛЭГ

Энэхүү судалгааны үр дүнгээс харахад Монголын хадны зурагт дүрслэгдсэн амьтдыг YOLOv8 загвараар тодорхой түвшинд илрүүлэх боломжтой байна. Гэсэн хэдий ч жишээнүүдээс үзвэл хамгийн том асуудал нь background руу алдагдах, түүнчлэн адуу–янгир, хүн–background хоорондын андуурал байв.

Train loss тогтвортой буурсан боловч validation classification loss 40–50 epoch-оос хойш өсөх хандлагатай байсан нь overfitting үүсэж болзошгүйг харуулсан. Иймд цаашдын сургалтанд илүү их өгөгдөл, ангиллын тэнцвэр, илүү тогтвортой annotation дүрэм, мөн early stopping эсвэл patience тохиргоо ашиглах нь зүйтэй.

Өмнөх олон улсын судалгаатай харьцуулахад [10]–[14] энэхүү ажил нь segmentation биш, харин Монголын хадны зураг дээр object detection-ийг туршсан, мөн адуу, хүн, Буга, янгир зэрэг Монголын хадны зурагт түгээмэл ангиллуудаар dataset үүсгэснээрээ онцлог юм. Huo нарын [12] ажилд segmentation-ийн AP өндөр гарсан боловч тэд өндөр чанартай, нэг бүсийн чиглэсэн dataset ашигласан. Харин манай ажил нь бүс нутгийн хувьд янз бүрийн, дүрсний чанар харилцан адилгүй, overlap ихтэй хадны зураг дээр туршигдсан тул илүү хүндрэлтэй нөхцөлтэй гэж үзэж болно.

Судалгааны практик үнэ цэнэ нь соёлын өвийн дижитал бүртгэл, амьтдын тархалтын тоон шинжилгээ, бүс нутгийн дүрслэлийн хэв маягийг харьцуулах суурь мэдээлэл үүсгэх боломжид оршино. Цаашид илүү олон зураг нэмж, polygon/mask annotation туршиж, YOLOv8m эсвэл segmentation загвартай харьцуулах замаар илрүүлэлтийн чанарыг ахиулах боломжтой.

VII. ДҮГНЭЛТ

Энэхүү өгүүлэлд Монголын хадны зураг дахь амьтдын дүрслэлийг YOLOv8 загвараар илрүүлэх туршилтын ажлыг шат дараатайгаар танилцууллаа. Анхны янгир-only туршилтаас эхлэн multiclass dataset үүсгэж, YOLOv8n ба YOLOv8s загваруудаар туршсан нь өгөгдлийн санг өргөтгөх, annotation дүрмээ тогтворжуулах, илрүүлэлтийн бодит хүндрэлүүдийг тодорхойлоход ач холбогдолтой байв.

Эцсийн байдлаар YOLOv8s загвар 367 зурагтай dataset дээр Precision=0.257, Recall=0.348, mAP@0.5=0.228, mAP@0.5:0.95=0.127 үзүүлэлттэй гарсан. Буга болон янгир ангиллуудын ялгарал сайн байсан бол хүн, адуу ангиллуудын хувьд илүү их өгөгдөл болон илүү нарийн annotation шаардлагатай нь тогтоогдов.

Ингэснээр Монголын хадны зураг дахь амьтдын дүрслэлийг автомат аргаар илрүүлэх боломж бодитой болохыг баталсан бөгөөд энэхүү судалгаа нь Монголын соёлын өвийн дүрсийн сан бүрдүүлэлт, автомат бүртгэл, статистик харьцуулалтын дараагийн ажлуудын суурь болох боломжтой. Мөн зургаан ангиллын refinement болон HSV preprocessing туршилтуудыг нэгтгэн үзэхэд YOLOv8 загварыг Монголын хадны зураг дахь амьтан, хүний дүрслэлийг илрүүлэхэд ашиглах боломжтой боловч үр дүн нь өгөгдлийн хэмжээ, annotation consistency, class imbalance болон хадны гадаргуугийн төвөгтэй нөхцлөөс хүчтэй хамаарч байна.

HSV preprocessing туршилтаар mAP@0.5 = 0.203, mAP@0.5:0.95 = 0.121 гарсан нь baseline-тай харьцуулахад бага боловч бодит ахиц үзүүлсэн. Энэ нь зургийн контраст, motif-ийн ялгаралтыг сайжруулах preprocessing нь жижиг dataset-тэй үед туслах боломжтойг харуулж байна. Гэсэн хэдий ч илүү өндөр түвшний detection-д хүрэхийн тулд өгөгдлийн санг өргөтгөх, class бүрийн жишээг тэнцвэржүүлэх, annotation-ийг polygon эсвэл segmentation хэлбэрээр нарийвчлах шаардлагатай.

ТАЛАРХАЛ

Энэхүү ажлыг Монгол Улсын Их Сургуулийн Мэдээллийн технологи, электроникийн сургуулийн Мэдээлэл, компьютерын ухааны тэнхимийн хүрээнд гүйцэтгэсэн. Судалгаанд ашигласан хадны зургийн эх хэрэглэгдэхүүн, зурагт ном, өмнөх судалгааны эмхэтгэлүүд нь өгөгдлийн сан бэлтгэхэд чухал хувь нэмэр оруулсан.

НОМ ЗҮЙ

[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779–788.

[2] J. Redmon and A. Farhadi, “YOLO9000: Better, Faster, Stronger,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 7263–7271.

[3] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” arXiv:2004.10934, 2020.

[4] C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, “YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors,” in Proceedings of CVPR, 2023, pp. 7464–7475.

[5] G. Jocher et al., Ultralytics YOLO Documentation and Repository. 2024.

[6] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137–1149, 2017.

[7] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN,” in Proceedings of ICCV, 2017, pp. 2961–2969.

[8] T.-Y. Lin et al., “Microsoft COCO: Common Objects in Context,” in Proceedings of ECCV, 2014, pp. 740–755.

[9] M. Everingham et al., “The Pascal Visual Object Classes (VOC) Challenge,” International Journal of Computer Vision, vol. 88, no. 2, pp. 303–338, 2010.

[10] A. Jalandoni, P. Tacon, S. May, et al., “On the Use of Machine Learning Methods in Rock Art Research with Application to Automatic Painted Rock Art Identification,” Journal of Archaeological Science, vol. 144, 2022.

[11] R. Turner-Jones, “Digitising the Deep Past: Machine Learning for Rock Art Motif Recognition,” 2023.

[12] D. Huo, S. Yang, and M. Hou, “Using the Improved YOLOv7-Seg Model to Segment Symbols from Rock Art Images,” npj Heritage Science, vol. 13, art. 16, 2025.

[13] M. L. Gil, J. Ortiz, S. Martínez-Rodríguez, B. Vázquez, and T. Rego, “The Detection of Petroglyphs through Digital Image Processing: The Particular Case of the Stone inside the Chapel of Saint Bartolomé (Lugo-Spain),” 2011.

[14] Z. Mnasri and A. D’Andrea, “Automatic Inventory of Archaeological Artifacts Based on Object Detection and Classification Using Deep and Transfer Learning,” Digital Applications in Archaeology and Cultural Heritage, vol. 39, e00458, 2025.

[15] Монголын ШУА Археологийн хүрээлэн, Монголын хадны зураг. Улаанбаатар, 2014.

[16] Монголын ШУА Археологийн хүрээлэн, Хаданд сийлсэн эртний урлаг. Улаанбаатар, 2020.

[17] E. Jacobson-Tepfer, The Hunter, the Stag, and the Mother of Animals: Image, Monument, and Landscape in Ancient North Asia. Oxford: Oxford University Press, 2015.

[18] E. Jacobson, The Буга Goddess of Ancient Siberia: A Study in the Ecology of Belief. Leiden: Brill, 1993.

[19] V. D. Kubarev, Petroglyphs of the Mongolian Altai. Novosibirsk, 2007.

[20] Y. Tserendagva, E. Jacobson-Tepfer, and D. Tseveendorj, Mongolian Altai Rock Art. Ulaanbaatar, 2005.

[21] П. Гантуяа, “Монгол орны говь ба хангайн бүсийн хадны зураг дүрслэгдсэн амьтдын дүрслэлийн ялгаа ба онцлог,” судалгааны өгүүлэл, 2026.

[22] R. G. Bednarik, Rock Art Science: The Scientific Study of Palaeoart. New Delhi: Aryan Books International, 2007.

Компьютерийн харааnd суурилсан хувийн хамгаалах хэрэгслийн хяналтын системийн загварчлал, хэрэгжилт

Ж.Сумъяарагчаа
Мэдээллийн технологи
ШУТИС-МХТС
Монгол улс, Улаанбаатар
sumiya.dsci@gmail.com

С.Өлзийбаяр
Мэдээллийн технологи
ШУТИС-МХТС
Монгол улс, Улаанбаатар
ulziibayar@must.edu.mn

***Хураангуй*— Барилга, үйлдвэрлэлийн орчинд ажилтнуудын аюулгүй байдлыг хангахад хувийн хамгаалах хэрэгслийн (XXX) зөв хэрэглээ чухал үүрэг гүйцэтгэдэг. Гэвч бодит орчинд XXX-ийн хэрэглээг гар аргаар тогтмол хянах нь цаг хугацаа, хүний нөөцийн хувьд хүндрэлтэй байдаг. Энэхүү судалгаанд компьютерийн харааг ашиглан XXX-ийн хэрэглээг автоматаар илрүүлж хянах систем боловсруулсан. Судалгаанд Roboflow платформоос цуглуулсан, helmet, vest, boots, gloves, glasses, human гэсэн 6 ангилал бүхий нийт 4999 зургийн өгөгдлийн санг ашигласан. YOLO11m, RT-DETR-L, Faster R-CNN MobileNetV3 FPN болон RetinaNet ResNet50 FPN гэсэн дөрвөн объект илрүүлэлтийн загварыг сургаж, харьцуулан үнэллээ. Baseline туршилтын үр дүнд YOLO11m болон RT-DETR-L загварууд тус бүр mAP50=0.917 болон 0.905 үзүүлж хамгийн өндөр нарийвчлалтай гарсан бөгөөд YOLO11m 205 FPS-ийн хурдаараа нарийвчлал, хурдны тэнцвэртэй гүйцэтгэлийг үзүүлсэн. Түүнчлэн өгөгдлийн баяжуулалтын (data augmentation) загваруудад үзүүлэх нөлөөллийг туршиж үнэлсэн ба RT-DETR-L загвар augmented нөхцөлд mAP50=0.918 үзүүлж хамгийн тогтвортой гүйцэтгэлтэй гарсныг баталсан. Уг загварт тулгуурлан видео дамжуулалт дээр дутуу XXX-ийг илрүүлж анхааруулга өгдөг мониторинг систем боловсруулж, туршилтаар ажиллагааг нь баталгаажуулсан.**

Түлхүүр үг*— *YOLO11m, RT-DETR-L, Faster R-CNN, RetinaNet

I. Удиртгал

Үйлдвэрлэл, барилгын салбарт ажилтнуудын аюулгүй байдал нь чухал асуудал хэвээр байна. Олон улсын хөдөлмөрийн байгууллагын (ILO) мэдээгээр жил бүр дэлхий даяар 2.3 сая хүн ажлын байрны осол, мэргэжлийн өвчний улмаас нас барж байна [1]. Эдгээр осол гэмтлийн нэлээд хэсэг нь хувийн хамгаалах хэрэгслийг (XXX) дутуу буюу буруу хэрэглэснээс үүсдэг.

Уламжлалт аргаар аюулгүй байдлын инспектор ажилчдын XXX хэрэглээг гар аргаар шалгаж хянадаг. Гэвч энэ арга нь хүний хүчин зүйлээс хамаарах, тогтмол бус хяналттай байх, олон тооны ажилтантай томоохон байгуулламжид хэрэгжүүлэхэд хүндрэлтэй зэрэг сул талуудтай.

Сүүлийн жилүүдэд компьютерийн харааны технологи хурдацтай хөгжиж, дүрс болон видео боловсруулалтад суурилсан автомат хяналтын систем боловсруулах боломж бий болсон. Ялангуяа гүн сургалтад суурилсан объект илрүүлэлтийн загварууд нарийвчлал болон хурдны хувьд эрс сайжирч, бодит цагийн хэрэглээнд тохиромжтой болсон [2].

Энэхүү судалгаанд YOLO11m, RT-DETR-L, Faster R-CNN MobileNetV3 FPN болон RetinaNet ResNet50 FPN гэсэн дөрвөн объект илрүүлэлтийн загварыг харьцуулан туршиж, augmentation-д хамгийн тогтвортой гүйцэтгэл үзүүлсэн загварыг сонгон ажилчдын XXX хэрэглээг бодит цагт хянах мониторинг систем боловсруулсан.

II. Ижил төстэй ажлуудын судалгаа

Хувийн хамгаалах хэрэгслийн илрүүлэлтийн чиглэлээр сүүлийн жилүүдэд олон судалгаа хийгдсэн байна. Nath болон бусад [3] нь гүн сургалтад суурилсан аргыг ашиглан барилгын талбай дахь ажилчдын дуулга малгай болон аюулгүйн хантаазны хэрэглээг илрүүлэх системийг боловсруулж, 90%-иас дээш нарийвчлал гаргасан. Li болон бусад [4] нь YOLOv5 загварыг ашиглан барилгын талбайн XXX илрүүлэлтийг хийж, бодит цагийн орчинд хэрэгжүүлэх боломжтойг харуулсан. Lu болон бусад [5] нь олон ангиллын XXX илрүүлэлтэд тулгарч буй class imbalance асуудлыг шийдвэрлэхийн тулд weighted loss function ашигласан бөгөөд жижиг объект болох бээлий, нүдний хамгаалалтын илрүүлэлтийн нарийвчлалыг сайжруулсан. Nath болон Akhavian [6] нь камерын өнцөг, гэрэлтүүлгийн нөхцөл өөрчлөгдөхөд системийн гүйцэтгэл хэрхэн өөрчлөгддөгийг судалж, өгөгдлийн баяжуулалт (data augmentation) ашиглах нь загварын тогтвортой байдлыг нэмэгдүүлдгийг нотолсон.

Объект илрүүлэлтийн загваруудын хөгжлийн хувьд Redmon болон бусад [7] YOLO архитектурыг анх санал болгосноос хойш уг загвар олон удаа сайжруулагдаж ирсэн. YOLOv8, YOLOv9 болон хамгийн сүүлийн YOLO11 загварууд нарийвчлал болон хурдны хувьд өмнөх хувилбаруудаасаа давуу болохыг туршилтаар нотолсон [8]. Two-stage detector-ийн төлөөлөл болох Faster R-CNN нь Region Proposal Network ашиглан өндөр нарийвчлалтай илрүүлэлт хийдэг ч one-stage detector-уудтай харьцуулахад хурд харьцангуй удаан байдаг [9]. Lin болон бусад [10] нь RetinaNet загварт focal loss-ийг нэвтрүүлж, one-stage detector-ийн class imbalance-ийн асуудлыг шийдэж, нарийвчлалын хувьд two-stage detector-уудтай өрсөлдөх боломжтой болгосон. Нөгөө талаас Carion болон бусад [11] transformer-д суурилсан DETR архитектурыг санал болгосон бөгөөд үүний үр дүнд боловсруулагдсан RT-DETR загвар нь anchor-free аргаар ажиллаж, өндөр нарийвчлалтайгаа хурдыг хослуулсан [12].

Сүүлийн хоёр жилийн хугацаанд XXX илрүүлэлтийн чиглэлд шинэ архитектуруудыг ашигласан судалгаанууд эрчимтэй хийгдэж байна. Zhao болон бусад [13] YOLOv10 болон Swin Transformer архитектурыг хослуулан барилгын талбайн XXX илрүүлэлтэд хэрэглэж, transformer-д суурилсан хандлага уламжлалт CNN аргуудаас илүү давуу гүйцэтгэл үзүүлдгийг нотолсон бөгөөд дундаж AP50 = 0.873 хүрсэн. Liu болон бусад [14] RT-DETR загварт суурилсан LMD-RTDETR алгоритмыг боловсруулж, жижиг объектын XXX илрүүлэлтэд mAP50-95 үзүүлэлтийг 2.4%-иар сайжруулсан ба параметрийн тоог 24.2%-иар бууруулсан. Kim болон Xiong [15] YOLOv8 загварт coordinate attention модуль болон ghost convolution нэмэх замаар өргөжүүлж, edge төхөөрөмж дээр бодит цагт ажиллах боломжтой хувилбарыг санал болгосон. S. Li болон бусад [16] MRC-DETR загварыг эрчим хүчний салбарын XXX илрүүлэлтэд зориулан боловсруулж, жижиг болон далдлагдсан объектын илрүүлэлтийн нарийвчлалыг нэмэгдүүлсэн. C. Li болон бусад [17] DilateFormer attention механизмыг YOLOv5-тай хослуулан SD-YOLOv5 загварыг санал болгож, CHV өгөгдлийн санд AP = 93.7% үзүүлсэн.

Дээрх судалгаануудаас харахад ихэнх ажил тухайн нэг архитектурыг сайжруулах эсвэл тодорхой хэрэглээний чиглэлд тохируулахад анхаарал хандуулсан байна. Гэвч one-stage CNN (YOLO), two-stage CNN (Faster R-CNN), focal loss бүхий one-stage (RetinaNet) болон transformer-д суурилсан (RT-DETR) гэсэн дөрвөн өөр төрлийн объект илрүүлэлтийн архитектурыг нэгэн XXX өгөгдлийн санд ижил тохиргоогоор сургаж, нарийвчлал, хурд болон augmentation-д үзүүлэх хариу үйлдэл талаас нь цогцоор үнэлсэн судалгаа цөөн байна. Энэхүү судалгаа нь уг дутагдлыг нөхөж, практик хэрэглээнд тохиромжтой мониторинг систем боловсруулахад чиглэсэн.

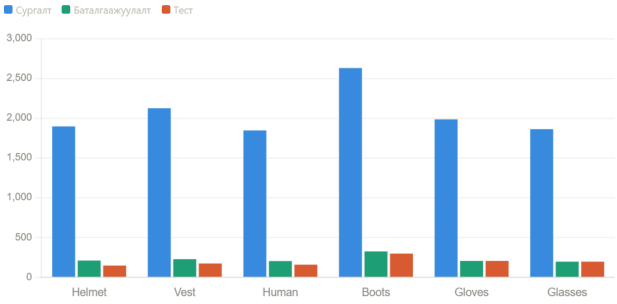
III. Ашигласан арга

A. Өгөгдлийн багц

Судалгаанд ашигласан өгөгдлийн багцыг Roboflow платформ дээр байршсан нийтэд нээлттэй өгөгдлийн сангуудаас цуглуулан бүрдүүлсэн. Өгөгдлийн цуглуулалтын явцад gloves болон glasses зэрэг харьцангуй жижиг хэмжээтэй объектын ангиллуудад зориулан тусдаа хайлт хийн, цуглуулан нэгтгэн боловсруулсан. Цуглуулсан өгөгдөл нь гадна болон дотор орчин, өдрийн харилцан адилгүй гэрэлтүүлгийн нөхцөл, өөр өөр камерын өнцөг зэрэг олон янзын нөхцөлд авагдсан зургуудыг агуулж байгаа нь загварын ерөнхийлөн сургалтад (generalization) эерэгээр нөлөөлсөн. Бүрдүүлсэн өгөгдлийн багц нийт 4999 зургийг агуулах бөгөөд загварын сургалт, үнэлгээний зориулалтаар сургалт (3664), баталгаажуулалт (668), тест (667) гэсэн гурван хэсэгт хуваасан. Илрүүлэх зорилтот ангиллуудад helmet, vest, boots, gloves, glasses, human гэсэн 6 ангилал хамрагдах бөгөөд нийт annotation-ийн тоо 14,869 болно. Ангилал тус бүрийн тархалтыг Зураг 1,2-д үзүүлэв.

Purpose	Number of images	Class						
		Helmet	Vest	Human	Boots	Gloves	Glasses	Total
Training	3664	1895	2124	1845	2,629	1984	1861	12,338
Validation	668	209	226	202	324	204	195	1360
Test	667	146	172	157	297	204	195	1171
Total	4999	2250	2522	2204	3,250	2392	2251	14,869

Зураг 1. Өгөгдлийн статистик мэдээлэл



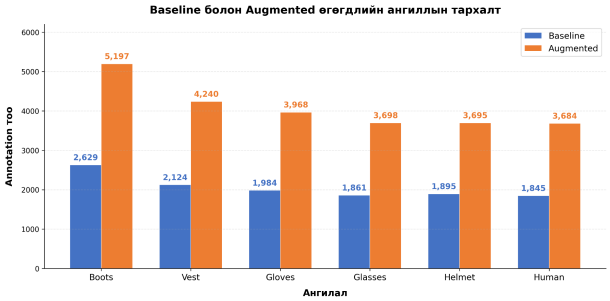
Зураг 2. Өгөгдлийн ангилуудын график мэдээлэл

B. Урьдчилсан боловсруулалт

- Polygon → Bounding box хөрвүүлэлт: Өгөгдлийн нэг хэсэг нь polygon хэлбэрийн annotation-тай байсан тул объект илрүүлэлтийн загваруудад шаардлагатай axis-aligned bounding box хэлбэрт хөрвүүлсэн.
- Өгөгдлийн форматын тохируулга: YOLO11m болон RT-DETR-L загваруудад YOLO форматын annotation ашиглаж, оруулах зургийн хэмжээг 960×960 пиксел болгон тохируулсан. Faster R-CNN болон RetinaNet загваруудын хувьд YOLO форматын өгөгдөл дээр сургасан боловч COCO өгөгдлийн санд урьдчилан сургасан жин (pretrained weights)-ээс эхлүүлсэн.
- Ангиллын тэнцвэржүүлэлт: Өгөгдлийн сангийн ангилал хоорондын тэнцвэргүй байдлыг (class imbalance) арилгах зорилгоор human болон helmet ангиллуудад oversampling аргыг ашиглан нэмэлт зураг оруулж тэнцвэржүүлэлт хийсэн.
- Чанарын шүүлт: Давхардсан болон чанар муутай зургуудыг шүүн хасаж, annotation-ийн зөвийг нягтлан шалгасан.

C. Өгөгдлийн баяжуулалт (Data Augmentation)

Сургалтын өгөгдлийн олон янз байдлыг нэмэгдүүлж, загварын ерөнхийлөн сургалтын чадварыг сайжруулах зорилгоор Albumentations номын сан ашиглан өгөгдлийн баяжуулалт хийсэн. Augmentation нь зөвхөн сургалтын өгөгдөлд (training set) хэрэглэгдсэн бөгөөд баталгаажуулалт болон тестийн өгөгдлүүд үл өөрчлөгдсөн хэвээр үлдсэн. Баяжуулалтын үр дүнд сургалтын зургийн тоо 3664-өөс 7328 болж, annotation-ий тоо 12,338-аас 24,482 болж тус тус хоёр дахин нэмэгдсэн.



Зураг 3. Baseline болон Augmented өгөгдлийн ангиллын тархалт

Augmentation-ыг дараах 6 бүлэгт хуваан хэрэгжүүлсэн:

ХҮСНЭГТ 1. АШИГЛАСАН AUGMENTATION АРГУУД

№	Арга	Техник	Магадлал
1	Геометрийн хувиргалт	Хэвтээ тусгал, ±15° эргүүлэлт, масштаб (80–120%)	0.5
2	Гэрэлтүүлгийн хувиргалт	Brightness, contrast, CLAHE, ColorJitter	0.7
3	Сүүдэр, нарны гэрэл	RandomShadow, RandomSunFlare	0.3
4	Blur, noise	Gaussian blur, Motion blur, Gaussian noise	0.3
5	Санамсаргүй тайралт	60–100% хэсгийг тайрч, 960×960 руу буцаах	0.3
6	Цаг агаарын нөхцөл	Манан (fog), бороо (rain)	0.15

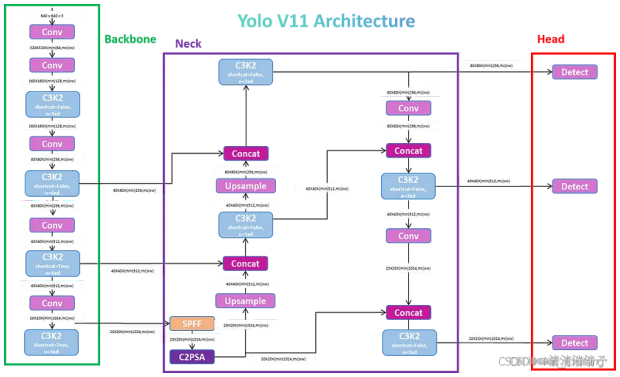
Эдгээр augmentation аргуудыг сонгохдоо барилга, үйлдвэрлэлийн бодит орчинд тохиолдох хүчин зүйлсийг загварчлахыг зорьсон. Тухайлбал, гэрэлтүүлгийн хувиргалт болон сүүдэр/нарны гэрэл нь гадаа ажиллах үеийн гэрэлтүүлгийн өөрчлөлтийг, blur болон noise нь камерын чичирхийлэл, тоосжилтыг, цаг агаарын augmentation нь манан, борооны нөхцлийг тус тус загварчилдаг. Augmentation-ийн нөлөөллийг үнэлэхийн тулд дөрвөн загварыг хоёр удаа сургасан: baseline (augmentation-гүй) болон augmented (дээрх augmentation-тай). Харьцуулсан үр дүнг IV хэсэгт танилцуулав.

D. Загвар сонголт

Судалгаанд объект илрүүлэлтийн өөр өөр архитектурын төлөөлөл болох дөрвөн загварыг сонгосон. Загвар сонголтод one-stage болон two-stage detector-уудын гүйцэтгэлийг харьцуулах, мөн сүүлийн үеийн transformer-д суурилсан архитектурын давуу талыг уламжлалт CNN-д суурилсан загваруудтай жишихийг зорьсон.

1. YOLO11m

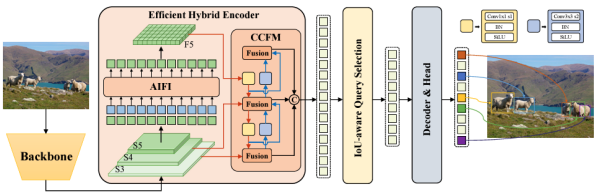
YOLO11m нь Ultralytics компанийн боловсруулсан YOLO цувралын хамгийн сүүлийн үеийн загвар бөгөөд one-stage detection архитектурт суурилдаг. Уг загвар нь C3k2 болон SPPF блокуудыг ашигласан backbone, өндөр нарийвчлалтай илрүүлэлт хийх боломжтой neck болон detection head-ээс бүрдэх бүтэцтэй. YOLO11m нь өмнөх хувилбаруудтай харьцуулахад параметрийн тоог бууруулж, нарийвчлалыг нэмэгдүүлсэн нь бодит цагийн хэрэглээнд давуу тал болдог. Энэхүү судалгаанд хурд болон нарийвчлалын тэнцвэртэй гүйцэтгэлийг үнэлэх зорилгоор сонгосон.



Зураг 4. YOLO11m архитектур [8]. Backbone хэсэг нь C3k2 болон SPPF блокуудыг ашиглан зургаас шинж чанарыг гаргаж авдаг бол neck хэсэг олон масштабын шинж чанарыг нэгтгэж, detection head нь объектын ангилал болон bounding box-ыг урьдчилан тааварладаг.

2. RT-DETR-L

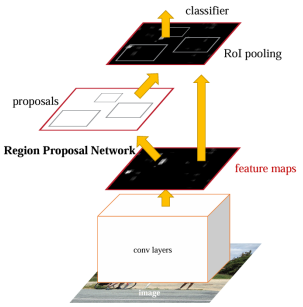
RT-DETR-L нь Baidu компанийн боловсруулсан transformer архитектурт суурилсан бодит цагийн объект илрүүлэгч юм. Уг загвар нь CNN-д суурилсан backbone болон transformer-д суурилсан decoder-ийг хослуулсан hibrid архитектуртай бөгөөд anchor-free аргаар объектын байршлыг тодорхойлдог. RT-DETR нь Non-Maximum Suppression (NMS) үйлдлийг шаарддаггүй тул inference-ийн хурдыг нэмэгдүүлэхэд давуу талтай. Энэхүү судалгаанд transformer архитектурын нарийвчлалын боломжийг CNN-д суурилсан загваруудтай харьцуулах зорилгоор сонгосон.



Зураг 5. RT-DETR-L архитектур [12]. CNN-д суурилсан backbone-оос Efficient Hybrid Encoder олон масштабын шинж чанарыг гаргаж авч, IoU-aware query selection модуль нь decoder-т оруулах хамгийн магадлалтай объектын асуултуудыг сонгодог. Transformer decoder эцсийн ангилал болон bounding box-ыг анкер ашиглалгүйгээр (anchor-free) таамаглана.

3. Faster R-CNN MobileNetV3 FPN

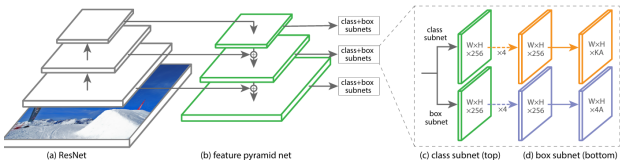
Faster R-CNN нь Region Proposal Network (RPN) болон объект ангилах сүлжээг хослуулсан two-stage detector архитектуртай. Эхний үе шатанд RPN зургаас объект байж болзошгүй бүс нутгуудыг (region proposal) санал болгох бөгөөд хоёрдугаар үе шатанд эдгээр бүс нутаг тус бүрийг ангилж, байршлыг нарийвчлан тодорхойлдог. Энэхүү судалгаанд backbone болгон MobileNetV3-г, олон масштабын объект илрүүлэлтэд Feature Pyramid Network (FPN)-г ашигласан. Faster R-CNN-г two-stage detector-ийн төлөөлөл болгон сонгож, one-stage detector-уудтай нарийвчлал болон хурдны хувьд харьцуулах зорилготой байсан.



Зураг 6. Faster R-CNN архитектур [9]. FPN-ээр гаргаж авсан шинж чанар дээр Region Proposal Network (RPN) объект байж болзошгүй бүсүүдийг (region proposal) санал болгож, RoI Pooling үе шатанд тэдгээрийг нэгдсэн хэмжээнд буулгасны дараа FC давхаргууд ангилал болон bounding box регрессийг гүйцэтгэнэ.

4. RetinaNet ResNet50 FPN

RetinaNet нь Lin болон бусад [10] нарын боловсруулсан one-stage detector бөгөөд объект илрүүлэлтийн ангилал хоорондын тэнцвэргүй байдлын (class imbalance) асуудлыг шийдвэрлэхэд чиглэсэн focal loss функцийг анх нэвтрүүлсэн загвар юм. Focal loss нь сургалтын явцад хялбар ангилагдах дэвсгэр (background) объектуудын нөлөөллийг бууруулж, хүндрэлтэй объектуудад илүү анхаарал хандуулах боломжийг олгодог. Энэхүү судалгаанд backbone болгон ResNet50-г, олон масштабын илрүүлэлтэд FPN-г ашигласан. RetinaNet-г сонгосон үндсэн шалтгаан нь олон ангилал бүхий XXX илрүүлэлтийн class imbalance нөхцөлд focal loss ашиглах нь загварын гүйцэтгэлд хэрхэн нөлөөлөхийг үнэлэх явдал байсан.



Зураг 7. RetinaNet архитектур [10]. ResNet-50 backbone-оос Feature Pyramid Network (FPN) P3–P7 түвшний олон масштабын шинж чанарыг үүсгэж, тус бүрд нь ангилал болон bounding box регрессийн дэд сүлжээнүүд зэрэгцэн ажиллаж focal loss ашиглан объектуудыг илрүүлнэ.

E. Сургалтын тохиргоо

Загвар тус бүрийн сургалтын тохиргоог Хүснэгт 2-т үзүүлэв.

ХҮСНЭГТ 2. ЗАГВАРУУДЫН СУРГАЛТЫН ТОХИРГОО

Тохиргоо	**YOLO11m**	**RT-DETR-L**	**Faster R-CNN**	**RetinaNet**
Epoch	50	50	50	50
Batch size	32	16	4	8
Optimizer	AdamW	AdamW	AdamW	AdamW
Learning rate	0.001	0.0005	0.0001	0.0001
Weight decay	0.01	0.0001	0.0001	0.0001
LR scheduler	Cosine	Cosine	StepLR	StepLR
Warmup epoch	3	3	—	—
Image size	960×960	960×960	960×960	960×960
Pretrained	COCO	COCO	COCO	COCO

- Batch size — нэг сургалтын алхамд боловсруулах зургийн тоо. Faster R-CNN-д 4 байгаа нь уг загварын санах ойн өндөр хэрэглээтэй холбоотой.
- Learning rate — загварын жинг шинэчлэх хурд. YOLO11m-д харьцангуй өндөр (0.001) байгаа нь cosine scheduler болон warmup-тай хослуулсантай холбоотой.
- LR scheduler — сургалтын явцад learning rate-г хэрхэн бууруулахыг тодорхойлно. Cosine нь аажмаар бууруулах бол StepLR тодорхой алхам тутамд гама коэффициентоор бууруулдаг.
- Warmup epoch — сургалтын эхний үе шатанд learning rate-г аажмаар нэмэгдүүлэх үе. Загварын эхний сургалтыг тогтворжуулахад ашигтай.
- Weight decay — загварын жинг хэт том болохоос сэргийлэх regularization параметр.

F. Үнэлгээний метрик

Загваруудын гүйцэтгэлийг үнэлэхэд дараах үзүүлэлтүүдийг ашигласан.

- Precision — загварын илрүүлсэн объектуудаас зөв илрүүлсэн хувийг илэрхийлнэ. Өөрөөр хэлбэл худал эерэг (false positive) тохиолдлын эзлэх хувийг харуулдаг.
- Recall — бодит объектуудаас загварын илрүүлж чадсан хувийг илэрхийлнэ. Худал сөрөг (false negative) тохиолдлын эзлэх хувийг харуулдаг.

- F1-score — Precision болон Recall-ийн хармоник дундаж бөгөөд хоёр үзүүлэлтийн тэнцвэрийг нэг тоогоор илэрхийлнэ.
- mAP50 — IoU босго утгыг 0.5 гэж тогтоон тооцсон дундаж нарийвчлал (mean Average Precision). Объектын илрүүлэлтийн нарийвчлалыг бүх ангиллын хувьд дундажлан үнэлдэг.
- mAP50-95 — IoU босго утгыг 0.5-аас 0.95 хүртэл 0.05 алхамаар өөрчлөн тооцсон дундаж нарийвчлал. Илрүүлэлтийн чанарыг илүү нарийвчлалтай, бодитоор үнэлэх боломжийг олгодог.
- FPS (Frames Per Second) — нэг секундэд боловсруулах зургийн тоо бөгөөд загварын бодит цагийн хэрэглээнд тохиромжтой эсэхийг үнэлэхэд ашигладаг.

IV. Туршилт, үр дүн

A. Туршилт

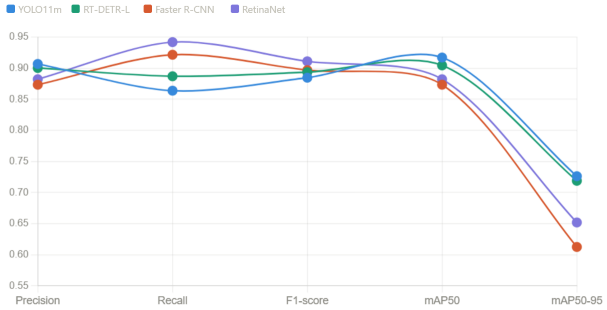
Туршилтыг Google Colab платформ дээр NVIDIA T4 болон A100 GPU ашиглан явуулсан. Програм хангамжийн орчинд Python 3.12.13, PyTorch 2.10.0+cu128 болон torchmetrics номын санг ашигласан. Загвар бүрийн гүйцэтгэлийг test dataset дээр precision, recall, F1-score, mAP50, mAP50-95 болон FPS үзүүлэлтүүдээр үнэлж харьцуулсан.

Дөрвөн загварын test dataset дахь гүйцэтгэлийг Хүснэгт 3-т үзүүлэв.

ХҮСНЭГТ 3. ЗАГВАРУУДЫН BASELINE ТЕСТ ҮР ДҮНГИЙН ХАРЬЦУУЛАЛТ

Үзүүлэлт	YOLO11m	RT-DETR-L	Faster R-CNN	Retina Net
Precision	**0.9068**	0.9003	0.8734	0.8822
Recall	0.8638	0.8870	0.9216	**0.9419**
F1-score	0.8847	0.8936	0.8968	**0.9110**
mAP50	**0.9171**	0.9045	0.8734	0.8822
mAP50-95	**0.7263**	0.7189	0.6125	0.6519
FPS	**205.54**	111.25	70.17	74.79

- Хүснэгт 3-аас үзэхэд mAP50 болон mAP50-95 үзүүлэлтээр YOLO11m загвар тус тус 0.9171 болон 0.7263 үзүүлж хамгийн өндөр нарийвчлалтай гарсан бөгөөд 205.54 FPS-ийн хурдаараа нарийвчлал, хурдны тэнцвэртэй гүйцэтгэлийг үзүүлсэн. RT-DETR-L загвар нарийвчлалын хувьд YOLO11m-тэй ойролцоо гүйцэтгэл үзүүлсэн боловч хурдны хувьд арай дутмаг байсан. RetinaNet загвар recall болон F1-score үзүүлэлтүүдэд тэргүүлсэн нь focal loss ашигласны үр дүн гэж үзэж болно. Faster R-CNN загвар бүх үзүүлэлтээр харьцангуй бага гүйцэтгэл үзүүлсэн.



Зураг 8. Дөрвөн загварын үндсэн үзүүлэлтүүдийн харьцуулалт. Хэвтээ тэнхлэгт Precision, Recall, F1-score, mAP50, mAP50-95 үнэлгээний үзүүлэлтүүд, босоо тэнхлэгт тэдгээрийн утгыг үзүүлэв. YOLO11m болон RT-DETR-L загварууд mAP50, mAP50-95 хэмжигдэхүүнд тэргүүлсэн бол RetinaNet recall үзүүлэлтэд хамгийн өндөр гүйцэтгэл үзүүлсэн нь графикаас ажиглагдана.

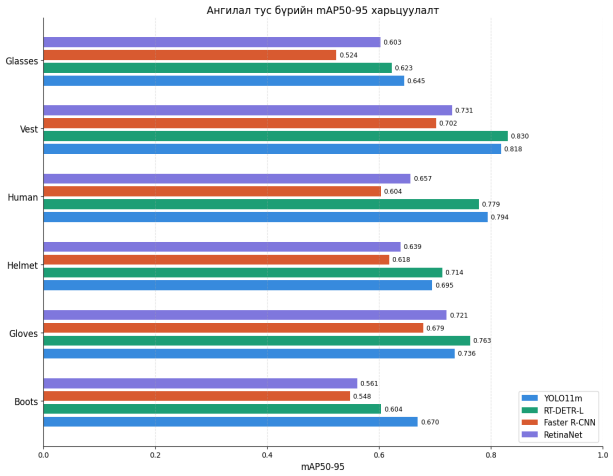
Ангилал тус бүрийн mAP50-95 үнэлгээний үзүүлэлтийг хүснэгт 4-т үзүүлэв.

ХҮСНЭГТ 4. АНГИЛАЛ ТУС БҮРИЙН mAP50-95

Ангилал	YOLO11m	RT-DETR-L	Faster R-CNN	RetinaNet
Boots	**0.6696**	0.6042	0.5482	0.5614
Gloves	0.7356	**0.7632**	0.6791	0.7212
Helmet	0.6952	**0.7135**	0.6184	0.6388
Human	**0.7943**	0.7790	0.6035	0.6567
Vest	0.8184	**0.8304**	0.7021	0.7306
Glasses	**0.6449**	0.6230	0.5239	0.6028

- Хүснэгт 4-аас харахад vest ангилал бүх загварт хамгийн өндөр mAP50-95 үзүүлж (YOLO11m: 0.818, RT-DETR-L: 0.830), glasses болон boots ангиллууд хамгийн бага үзүүлэлттэй гарсан. Vest ангилал өндөр үзүүлэлт гаргасан нь уг объект харьцангуй том хэмжээтэй, тод өнгөтэй байдгаас шалтгаалж илрүүлэхэд хялбар болохтой холбоотой. Харин glasses болон boots ангиллууд жижиг хэмжээтэй, бусад объектод хэсэгчлэн далдлагдах тохиолдол их байдаг тул илрүүлэлтийн нарийвчлал харьцангуй бага гарсан гэж үзэж болно.
- Ангилал тус бүрийн үзүүлэлтийг загваруудаар харьцуулахад gloves болон helmet ангиллуудад RT-DETR-L загвар YOLO11m-ийг давсан бол human болон vest ангиллуудад YOLO11m давуу байсан. Faster R-CNN болон RetinaNet загварууд бүх ангилалд YOLO11m болон RT-DETR-L-ийг даваагүй боловч RetinaNet нь Faster R-CNN-ийг

ихэнх ангилалд давсан нь focal loss-ийн үр нөлөөг харуулж байна.



Зураг 9. Ангилал тус бүрийн mAP50-95 үзүүлэлтийн харьцуулалт. Хэвтээ тэнхлэгт mAP50-95 утга, босоо тэнхлэгт boots, gloves, helmet, human, vest, glasses гэсэн зургаан ангилалыг үзүүлэв. Vest ангилал бүх загварт хамгийн өндөр, glasses ангилал хамгийн бага үзүүлэлттэй гарсан нь объектын хэмжээ болон бусад объектод далдлагдалт нь илрүүлэлтийн нарийвчлалд ихээхэн нөлөөлдөг болохыг илэрхийлж байна.

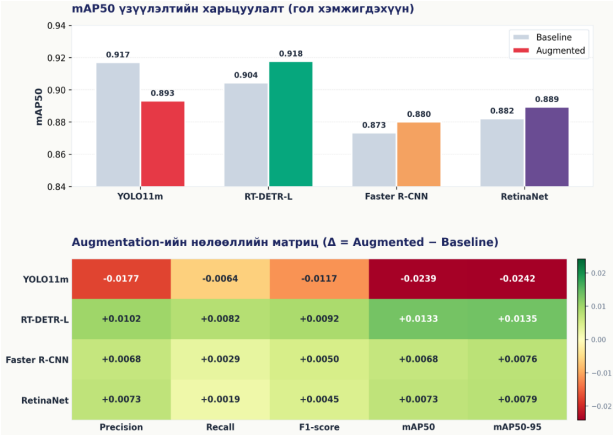
B. Augmentation-ийн нөлөөллийн шинжилгээ

Augmentation-ийн үр нөлөөг үнэлэхийн тулд дөрвөн загварыг augmented өгөгдлийн багцад ижил тохиргоогоор дахин сургаж, test dataset дээр үнэлсэн. Үр дүнг Хүснэгт 5-т үзүүлэв.

ХҮСНЭГТ 5. AUGMENTED ӨГӨГДЛИЙН БАГЦААС ГАРСАН ТУРШИЛТЫН ҮР ДҮН

Үзүүлэлт	YOLO11m	RT-DETR-L	Faster R-CNN	Retina Net
Precision	0.8891	**0.9105**	0.8802	0.8895
Recall	0.8574	0.8952	0.9245	**0.9438**
F1-score	0.8730	0.9028	0.9018	**0.9155**
mAP50	0.8932	**0.9178**	0.8802	0.8895
mAP50-95	0.7021	**0.7324**	0.6201	0.6598

- Augmentation-ийн нөлөөллийг baseline үр дүнтэй (Хүснэгт 3) харьцуулахад тэгш бус үр дүн ажиглагдсан. RT-DETR-L загвар augmentation-аас хамгийн их ашиг хүртсэн бөгөөд mAP50 +0.013, mAP50-95 +0.014-өөр нэмэгдэж, бүх үзүүлэлтүүдэд тогтвортой сайжрал үзүүлсэн. Faster R-CNN болон RetinaNet загварууд мөн augmentation-ийн эерэг нөлөөг хүлээн авч, mAP50-95 тус бүр +0.008 болон +0.008 нэмэгдсэн.
- Харин YOLO11m загварт augmentation сөрөг нөлөө үзүүлсэн бөгөөд mAP50 -0.024, mAP50-95 -0.024-өөр буурсан. Энэ үзэгдэл нь YOLO11m загварын сургалтын хүрээнд Ultralytics framework дотроо mosaic, mixup, HSV augmentation зэрэг суулгасан augmentation аргуудтай холбоотой гэж үзэж болно. Гаднаас нэмэлт augmentation хийснээр хэт их баяжуулалт (over-augmentation) үүсч, загварын сургалтыг тогтворгүй болгосон байх магадлалтай. Энэ нь дотоод augmentation-тай framework ашиглаж буй загваруудын хувьд гадаад augmentation-ийг болгоомжтой хэрэглэх шаардлагатайг харуулж байна.



Зураг 10. Augmentation нөлөө

Ангилал тус бүрийн Augmented өгөгдөл дээрх mAP50-95 үнэлгээний үзүүлэлтийг хүснэгт 6-т үзүүлэв.

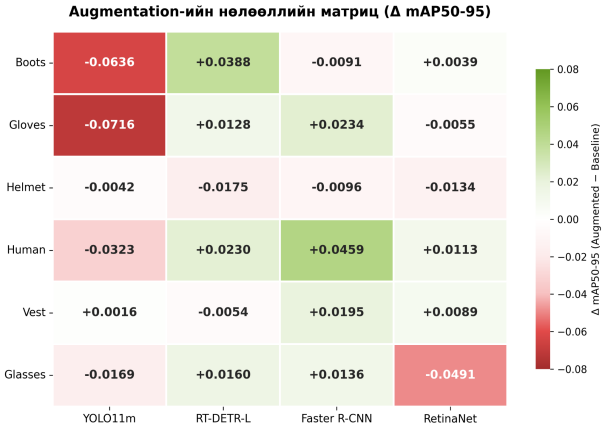
ХҮСНЭГТ 6. AUGMENTED ӨГӨГДЛИЙН БАГЦ ДАХЬ АНГИЛАЛ ТУС БҮРИЙН mAP50-95

Ангилал	YOLO11m	RT-DETR-L	Faster R-CNN	RetinaNet
Boots	0.6060	**0.6430**	0.5391	0.5653
Gloves	0.6640	**0.7760**	0.7025	0.7157
Helmet	0.6910	**0.6960**	0.6088	0.6254
Human	0.7620	**0.8020**	0.6494	0.6680
Vest	0.8200	**0.8250**	0.7216	0.7395
Glasses	0.6280	**0.6390**	0.5375	0.5537

- Augmented нөхцөлд RT-DETR-L загвар 6 ангилал бүгдэд хамгийн өндөр mAP50-95 үзүүлж, baseline нөхцөлтэй харьцуулахад ерөнхий тэргүүлэгч болсон.
- Жижиг объектуудад RT-DETR-L загвар (gloves: 0.776, glasses: 0.639) үзүүлж transformer архитектур augmentation-тай хослохдоо илт давуу гарсан.
- Vest ангилал бүх загварт хамгийн өндөр (0.72–0.83), glasses ангилал хамгийн бага үзүүлэлттэй (0.54–0.64) гарсан нь baseline-тай

ижил хэв маяг хадгалж, объектын хэмжээ илрүүлэлтийн нарийвчлалд ихээхэн нөлөөлдгийг нотлож байна.

- Faster R-CNN болон RetinaNet augmentation-аас ашиг хүртсэн ч ерөнхий гүйцэтгэлээрээ one-stage YOLO болон transformer RT-DETR-L-ийг давж чадаагүй.



Зураг 11. Ангилалууд дээрх augmentation нөлөөллийн матриц. Augmented – baseline mAP50-95 өөрчлөлтийг дөрвөн загвар болон зургаан ангилалаар харьцуулсан. Ногоон өнгө augmentation-ийн эерэг, улаан өнгө сөрөг нөлөөг илэрхийлнэ.

C. Test өгөгдөл дахь илрүүлэлтийн жишээ зураг



Зураг 12. Илрүүлэлт

D. Мониторинг систем

Туршилтын үр дүнд augmentation-ийн дараа хамгийн тогтвортой гүйцэтгэл үзүүлсэн RT-DETR-L загварт тулгуурлан XXX мониторинг систем боловсруулсан.

Боловсруулсан мониторинг системийн архитектурыг Зураг 13-т үзүүлэв.

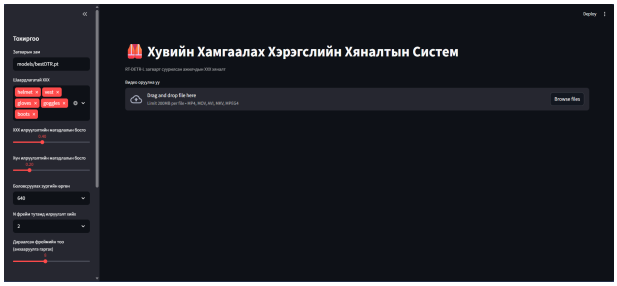


Зураг 13. Системийн архитектур

Эхний үе шатанд видео эх оруулгыг фрейм тус бүрээр задлан боловсруулна. Human detection болон PPE detection үе шатуудад RT-DETR-L загварыг ашиглан хүн болон XXX-ийн объектуудыг илрүүлж bounding box-оор тэмдэглэнэ. Association logic үе шатанд илрүүлсэн XXX-ийн bounding box нь хүний bounding box дотор байгаа эсэхийг IoU threshold ашиглан шалгаж, тухайн XXX-ийг харгалзах хүнд хамааруулна. Missing PPE check үе шатанд хүн тус бүрийн илрүүлсэн XXX-ийг шаардлагатай XXX-ийн жагсаалттай харьцуулж дутуу байгаа эсэхийг шалгана. Эцэст нь дутуу XXX илрүүлсэн тохиолдолд Alert & Logging үе шатанд анхааруулга үүсгэж бүртгэлд хадгална.

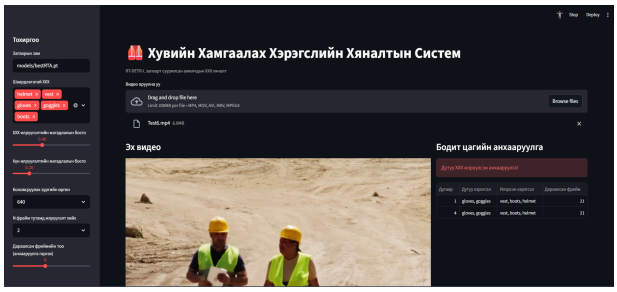
E. Хэрэглэгчийн интерфейс

Зураг 14-т системийн хэрэглэгчийн интерфейсийг үзүүлэв. Зүүн талын тохиргооны самбараас шаардлагатай XXX-ийн ангиллуудыг сонгох, PPE болон human илрүүлэлтийн confidence threshold тохируулах, боловсруулах зургийн өргөн болон inference хийх фреймийн давтамжийг тохируулах боломжтой. Мөн дутуу XXX хэдэн дараалсан фреймд илрүүлэгдсэний дараа анхааруулга үүсгэхийг "Alert after consecutive missing frames" тохиргоогоор тодорхойлно. Баруун талд видео файл оруулах хэсэг байрлах бөгөөд MP4, MOV, AVI, MKV, MPEG4 форматыг дэмждэг.



Зураг 14. Нүүр хуудас

Зураг 15-т системийн ажиллагааны үр дүнг үзүүлэв. Видео оруулсны дараа зүүн талд эх видео, баруун талд бодит цагийн анхааруулгын самбар харагдана. Туршилтанд Test5.mp4 файл ашиглаж, processing width=960, inference every 2 frames, confidence threshold=0.40 тохиргоогоор ажиллуулсан. Live Alert Preview самбарт "Confirmed missing PPE alert" анхааруулга гарч, track_id=1 болон track_id=4 дугаартай ажилтнуудын хувьд gloves болон goggles дутуу байгааг илрүүлж, boots, helmet, vest илэрсэн гэж бүртгэсэн байна.



Зураг 15. Системийн ажиллагаа



Зураг 16. Загварын илрүүлэлтийн үр дүн

V. Дүгнэлт

Энэхүү судалгаанд барилга, үйлдвэрлэлийн орчинд ажилтнуудын хувийн хамгаалах хэрэгслийн хэрэглээг компьютерийн хараат ашиглан автоматаар илрүүлж хянах систем боловсруулсан. Судалгааны хүрээнд Roboflow платформоос цуглуулсан 4999 зургийн өгөгдлийн санг ашиглан YOLO11m, RT-DETR-L, Faster R-CNN MobileNetV3 FPN болон RetinaNet ResNet50 FPN гэсэн дөрвөн объект илрүүлэлтийн загварыг сургаж, харьцуулан үнэлсэн. Загваруудын харьцуулалтад one-stage болон two-stage detector, мөн transformer архитектурт суурилсан загваруудын гүйцэтгэлийг нэгэн нөхцөлд үнэлэх боломж олдсон нь судалгааны нэг чухал хувь нэмэр болсон.

Baseline туршилтын үр дүнд YOLO11m загвар mAP50=0.917, mAP50-95=0.726 үзүүлж нарийвчлал, хурдны хувьд тэнцвэртэй гүйцэтгэл үзүүлсэн (205.54 FPS). RT-DETR-L загвар mAP50=0.905 үзүүлж YOLO11m-тэй ойролцоо нарийвчлалтай гарсан нь transformer архитектур объект илрүүлэлтийн чиглэлд CNN-д суурилсан загваруудтай өрсөлдөх чадвартай болсныг харуулж байна. RetinaNet загвар recall=0.942, F1-score=0.911 үзүүлж бусад загваруудыг давсан нь focal loss ашигласнаар ангилал хоорондын тэнцвэргүй байдлын асуудлыг үр дүнтэй шийдэж байгааг нотолсон. Faster R-CNN загвар бүх үзүүлэлтээр харьцангуй бага гүйцэтгэл үзүүлсэн нь two-stage архитектурын онцлогоос шалтгаалсан гэж үзэж болно.

Ангилал тус бүрийн үнэлгээнээс харахад vest ангилал бүх загварт хамгийн өндөр mAP50-95 үзүүлсэн бол glasses болон boots ангиллуудад харьцангуй бага үзүүлэлт гарсан. Энэ нь уг ангиллуудын объект харьцангуй жижиг хэмжээтэй, бусад объектод хэсэгчлэн далдлагдах тохиолдол их байдагтай холбоотой бөгөөд цаашдын судалгаанд анхаарал хандуулах шаардлагатай чиглэл болохыг харуулж байна.

Augmentation-ийн нөлөөллийн шинжилгээнээс харахад гадна augmentation-ийн стратеги нь загварын төрлөөс хамаарч ялгаатай үр дүн өгсөн. RT-DETR-L загвар augmentation-аас хамгийн их ашиг хүртэж mAP50=0.918, mAP50-95=0.732 үзүүлсэн бол Faster R-CNN, RetinaNet загварууд мөн эерэг нөлөөг хүлээн авсан. Харин YOLO11m загвар augmentation-тай хослохдоо гүйцэтгэл нь буурсан нь уг framework-ийн дотоод augmentation механизмтай давхцсаны үр дагавар гэж дүгнэгдсэн. Эдгээр олдвор нь дотоод augmentation-тай framework ашиглаж буй загваруудад гадна augmentation-ийг болгоомжтой хэрэглэх шаардлагатайг харуулж байна. Тиймээс augmentation-ийн дараа хамгийн тогтвортой, өндөр гүйцэтгэл үзүүлсэн RT-DETR-L загварыг эцсийн мониторинг системд сонгон ашигласан.

Судалгааны гол үр дүн болох мониторинг систем нь бодит видео дамжуулалт дээр дутуу XXX-ийг амжилттай илрүүлж, анхааруулга өгөх боломжтойг туршилтаар баталгаажуулсан. Streamlit ашиглан боловсруулсан хэрэглэгчийн интерфейс нь шаардлагатай XXX-ийн ангиллуудыг уян хатан тохируулах, confidence threshold тохируулах болон бодит цагийн анхааруулга харуулах боломжийг олгодог тул практик хэрэглээнд нэвтрүүлэхэд тохиромжтой. Ирээдүйн судалгаанд объект tracking алгоритмыг сайжруулах, edge device дээр системийг байршуулах боломжийг судлах, мөн өгөгдлийн баяжуулалтын стратегийг загвар тус бүрийн онцлогт тохируулан оновчтой болгох, ялангуяа gloves болон glasses зэрэг жижиг объектын илрүүлэлтийн нарийвчлалыг дээшлүүлэх зорилтуудыг тавьж байна.

Ном зүй

[1] ILO, "World Statistics on Occupational Safety and Health," International Labour Organization, Geneva, 2023. [Online]. Available: https://www.ilo.org/global/topics/safety-and-health-at-work/lang--en/index.htm

[2] Z. Zhao et al., "Object Detection with Deep Learning: A Review," IEEE Trans. Neural Netw. Learn. Syst., vol. 30, no. 11, pp. 3212–3232, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8627998

[3] N. D. Nath, A. H. Behzadan, and S. G. Paal, "Deep learning for site safety: Real-time detection of personal protective equipment," Automation in Construction, vol. 112, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580519308325

[4] Y. Li et al., "Deep Learning-Based Safety Helmet Detection in Engineering Management Based on Convolutional Neural Networks," Advances in Civil Engineering, 2020. [Online]. Available: https://www.hindawi.com/journals/ace/2020/9703560

[5] B. Lu et al., "Weighted Loss Function for PPE Detection with Class Imbalance," IEEE Access, vol. 9, 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9352513

[6] N. D. Nath and A. H. Behzadan, "Deep Convolutional Networks for Construction Object Detection Under Different Visual Conditions," Frontiers in Built Environment, vol. 6, p. 97, 2020. [Online]. Available Available:

https://www.frontiersin.org/journals/built-environment/articles/10.3389/fbuil.2020.00097/full

[7] J. Redmon et al., "You Only Look Once: Unified, Real-Time Object Detection," in Proc. IEEE CVPR, 2016, pp. 779–788. [Online]. Available: https://arxiv.org/abs/1506.02640

[8] G. Jocher et al., "Ultralytics YOLO11," GitHub, 2024. [Online]. Available: https://github.com/ultralytics/ultralytics

[9] S. Ren et al., "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE TPAMI, vol. 39, no. 6, pp. 1137–1149, 2017. [Online]. Available: https://arxiv.org/abs/1506.01497

[10] T.-Y. Lin et al., "Focal Loss for Dense Object Detection," in Proc. IEEE ICCV, 2017, pp. 2980–2988. [Online]. Available: https://arxiv.org/abs/1708.02002

[11] N. Carion et al., "End-to-End Object Detection with Transformers," in Proc. ECCV, 2020, pp. 213–229. [Online]. Available: https://arxiv.org/abs/2005.12872

[12] Y. Zhao et al., "DETRs Beat YOLOs on Real-time Object Detection," arXiv, 2023. [Online]. Available: https://arxiv.org/abs/2304.08069

[13] Y. Zhao et al., "Automated non-PPE detection on construction sites using YOLOv10 and transformer architectures for surveillance and body worn cameras with benchmark datasets," *Scientific Reports*, vol. 15, 2025. [Online]. Available: https://www.nature.com/articles/s41598-025-12468-8

[14] Y. Liu et al., "Personal Protective Equipment Detection for Industrial Environments: A Lightweight Model Based on RTDETR for Small Targets," *Research Square*, 2024. [Online]. Available: https://www.researchsquare.com/article/rs-5352653/v1

[15] D. Kim and S. Xiong, "Enhancing Worker Safety: Real-Time Automated Detection of Personal Protective Equipment to Prevent Falls from Heights at Construction Sites Using Improved YOLOv8 and Edge Devices," *Journal of Construction Engineering and Management*, vol. 151, no. 1, 2024. [Online]. Available: https://ascelibrary.org/doi/abs/10.1061/JCEMD4.COENG-14985

[16] S. Li et al., "MRC-DETR: A High-Precision Detection Model for Electrical Equipment Protection in Power Operations," *Sensors*, vol. 25, no. 13, p. 4152, 2025. [Online]. Available: https://www.mdpi.com/1424-8220/25/13/4152

[17] C. Li et al., "SD-YOLOv5: a rapid detection method for personal protective equipment on construction sites," *Frontiers in Built Environment*, vol. 11, 2025. [Online]. Available: https://www.frontiersin.org/journals/built-environment/articles/10.3389/fbuil.2025.1563483/full

Хүүхдийн аюулгүй байдлыг бодит цаг хугацаанд хянах систем

Очирын Оюунжаргал
Компьютерын ухааны тэнхим
ШУТИС, Мэдээлэл холбоо технологийн сургууль
Улаанбаатар хот
B232270019@must.edu.mn

Сүрэнжавын Азжаргал
Компьютерын ухааны тэнхим
ШУТИС, Мэдээлэл холбоо технологийн сургууль
Улаанбаатар хот
B232270052@must.edu.mn

Бямбадоржийн Золзаяа
Мэдээллийн технологийн тэнхим
ШУТИС, Мэдээлэл холбоо технологийн сургууль
Улаанбаатар хот
b.zolzaya@must.edu.mn

Хураангуй*— Сүүлийн жилүүдэд бага насны хүүхдийг хараа хяналтгүй орхисны улмаас гарч буй ахуйн осол, гэмтлийн тоо нэмэгдэж байна. Эдгээр осол нь хүүхдийн бие махбод, сэтгэл зүй, хөгжлийн хэвийн үйл явцад сөрөг нөлөө үзүүлж болзошгүй бөгөөд гэр ахуйн орчинд аюулгүй байдлын асуудлыг сайжруулах шаардлагатай байна. Энэхүү судалгаанд бид хүүхдийн аюулгүй байдлыг хангах зорилгоор дараах дөрвөн үе шаттай боловсруулалтыг хийсэн. (1) Хүүхэд илрүүлэх загварын хувьд Detectron2, RF-DETR, YOLOv5, YOLOv8, YOLOv11 гэсэн 5 ангиллын загварыг сургаж тестийн өгөгдөл дээр гүйцэтгэлийг нь харьцууланы үнэллээ. Туршилтын үр дүнгээс үзэхэд өгөгдлийн хязгаардлагдмал нөхцөлд YoloV8 загвар нь бусад загваруудаас илүү өндөр нарийвчлал (87%), гүйцэтгэлтэй гарч хүүхдийг илүү үр дүнтэй таних чадварыг харуулсан. (2) Хүүхдийн хөдөлгөөнийг тасралтгүй хянах зорилгоор ByteTrack, DeepSORT, BoT-SORT гэсэн мөрдөлтийн гурван алгоритм дээр туршилт хийсэн ба гүйцэтгэлийг нь харьцуулан үзэхэд BoTSORT нь бусад алгоритмаас илүү өндөр нарийвчлалтай MOTA(0.813), MOTP(0.818), F1(0.915) байв. (3). Гэрийн аюултай бүсийг тодорхойлохдоо хоёр төрлийн аргаар буюу аюултай объект таних, аюултай объект тэмдэглэх аргыг ашигласан. Аюултай объект танихдаа YOLOv8 дээр 4 төрлийн аюултай объектыг (шарах шүүгээ, богино долгионы зуух, плитка, залгуур) нарийн тохируулга (fine-tune) хийхэд тестийн өгөгдөл дээр гүйцэтгэл (80.25%) байв. Аюултай бүсийг тэмдэглэхдээ суурин камерын хувьд аюултай гэж үзсэн бүсээ, гараар цэгүүдийг нь сонгож тэмдэглэх боломжтой. (4) Хүүхэд аюултай бүс болон объекттой ойртвол мессежээр, мэйлээр, application-р дамжуулан гурван төрлийн мэдэгдэл илгээнэ. Уг судалгаа нь гэр ахуйн орчинд осол гэмтэлд өртсөн хүүхдийн тоог бууруулах, урьдчилан сэргийлэх, аюулгүй байдлыг хангах, асран хамгаалагчийн хяналтыг дэмжих зорилготой.

Түлхүүр үг— Yolo, BoTsort, Мэдэгдэл, Суурин камер

I.Удиртгал

Хүүхдийн аюулгүй байдлыг хангах асуудал нийгэмд анхаарал татсаар олон жилийг ардаа үдлээ. Учир нь нийгмийн амьдрал илүү завгүй болохын хэрээр эцэг эх, асран хамгаалагчдын хараа хяналт багасаж түүнчлэн хүүхэд богино хугацаанд ахуйн осол, гэмтэлд өртөх эрсдэл нэмэгдсээр байна.

Гэмтэл согог судлалын үндэсний төвийн мэдээллээр, сүүлийн гурван жилийн хугацаанд нийт 157 мянган хүүхэд осол, гэмтэлд өртөж эмнэлгийн тусламж үйлчилгээ авсан байна. Тэдний 53 хувь буюу 83 мянга нь гэртээ байхдаа осолд өртсөн гэх судалгаа гаржээ [1]. Энэхүү статистик нь гэр ахуйн орчин нь хүүхдэд аюулгүй байх ёстой хэдий ч, халуун гадаргуу, цахилгааны залгуур зэрэг эрсдэлтэй орчинтой харьцахаас үүдэлтэй түлэгдэлт мөн өндрөөс унаж бэртэх гэх мэт ноцтой гэмтэл бэртэлд өртөх нь их байгааг харуулж байна. Хэдийгээр эцэг эхчүүд хүүхдээ байнга хянахыг хичээдэг ч гэр ахуйн янз бүрийн ажлын улмаас тасралтгүй анхаарал хандуулах боломж хязгаарлагдмал байдаг. Гэвч тэрхүү хэдхэн секундын анхаарал сулрах нь буцалтгүй үр дагаварт хүргэх эрсдэлтэй байдаг.

Иймээс хүний хяналтыг дэмжих, нөхөх зорилготой бодит цагийн ухаалаг хяналтын системийг хөгжүүлэх нь нийгмийн тулгамдсан асуудлыг шийдвэрлэх чухал чиглэл болж байна. Компьютерын хараа, хиймэл оюун ухаанд суурилсан шийдлүүд нь хүүхдийн хөдөлгөөнийг автоматаар хянах, аюултай нөхцөл байдлыг эрт илрүүлэх, шуурхай мэдэгдэл хүргэх боломжийг бүрдүүлж байна.

Энэхүү судалгааны ажлаар дараах үндсэн хувь нэмрүүдийг оруулсан болно. Нэгдүгээрт, хүүхдийн аюулгүй байдлыг хянах зориулалттай хосолсон (hybrid) системийг санал болгосон бөгөөд энэ нь суурин камерын дүрсэн дээр аюултай бүсийг гараар полигон хэлбэрээр тэмдэглэх болон YOLOv8 загвараар аюултай объектыг автоматаар илрүүлэх аргуудыг нэгтгэсэн. Ийнхүү хэрэглэгчийн оролцоотой болон автомат илрүүлэлтийн давуу талыг хослуулснаар практик хэрэглээнд илүү уян хатан шийдэл болсон.

Хоёрдугаарт, хүүхэд илрүүлэх болон мөрдөх үе шатанд орчин үеийн гүн сургалтын таван өөр архитектур болон мөрдөлтийн гурван алгоритмын гүйцэтгэлийг нарийвчлан харьцуулж, бодит цагийн хэрэглээнд хамгийн оновчтой YOLOv8 болон BoT-

SORT алгоритмын хослолыг тодорхойлсон. Энэ нь системийн оновчлол болон туршилтын үр дүнгийн хувьд ач холбогдолтой.

Гуравдугаарт, аюултай нөхцөл илэрсэн үед SMS, Email, Push notification зэрэг олон сувгаар мэдэгдэл илгээх механизмыг хэрэгжүүлснээр интернэт холболтгүй нөхцөлд ч хэрэглэгчид шуурхай мэдээлэл хүргэх боломжийг бүрдүүлсэн. Энэ нь системийн найдвартай байдал, бодит хэрэглээний үнэ цэнийг нэмэгдүүлж байна.

II.Ижил төстэй ажлын судалгаа

Орчин үеийн технологийн дэвшлийг ашиглан хүүхдийн аюулгүй байдлыг хангах чиглэлээр олон судалгаа, бүтээлүүд хийгдэж байгаа бөгөөд эдгээрийн ихэнх нь компьютерын хараа болон гүн сургалтын аргуудыг ашиглан бодит цагийн орчныг хянахад чиглэдэг. Судалгаанаас харахад хүүхдийн аюултай бүсийг тодорхойлох үйл явц ихэвчлэн хоёр үндсэн аргаар хийгддэг. Үүнд:

1. Гараар аюултай бүсийг тодорхойлох арга
2. Объект таних загвар ашиглах автомат арга

Эхний аргын хувьд хэрэглэгч системийн интерфэйс дээр суурин камерын хувьд аюултай бүсийг гараар тодорхойлж, тэмдэглэсний дараа тухайн бүсэд хүүхэд орсон эсэхийг хянадаг зарчмаар ажилладаг. Энэ нь хэрэглэгчийн оролцоонд тулгуурладаг тул уян хатан, хэрэглэхэд хялбар боловч автоматаар шинэ нөхцөл байдалд дасан зохицох чадваргүй бөгөөд аюултай объектыг өөрөө илрүүлэх боломжгүй сул талтай.

Харин хоёр дахь арга буюу объект таних загвар ашигладаг системүүд нь орчныг автоматаар хянаж, гүн сургалтын загварын тусламжтайгаар аюултай объектуудыг илрүүлдэг. Жишээлбэл, “A Child Safety System for Pool Area” судалгаанд Raspberry Pi камер болон YOLO загварыг ашиглан усан сангийн орчинг бодит цагийн горимоор хянаж, хүүхэд усан сангийн бүсэд орсон тохиолдолд анхааруулга илгээдэг шийдэл боловсруулсан байдаг [2]. Энэ арга нь бүрэн автомат ажиллагаатай боловч хэрэглэгч өөрийн хүссэн бүсийг гараар тодорхойлох боломжгүй хязгаарлагдмал талтай.

Нөгөөтэйгүүр, CuboAI зэрэг худалдаанд гарсан ухаалаг хүүхэд хяналтын системүүд нь голчлон хөдөлгөөн, дуу авиа, нүүр царай илрүүлэх технологид тулгуурладаг [3]. Эдгээр нь тодорхой нөхцөлд анхааруулга илгээж чаддаг ч орчны аюултай бүсийг гараар тэмдэглэх болон гүн сургалтын аргаар автоматаар тодорхойлох функцүүдийг хослуулсан цогц шийдэл одоогоор бүрэн хөгжөөгүй байна.

Иймээс бид судалгаандаа дээрх хоёр аргыг хослуулсан, бодит цагийн хэрэглээнд тохиромжтой ухаалаг хяналтын системийг боловсруулахыг зорьж байна.

III.Ашигласан арга

Энэхүү систем нь гурван үндсэн үе шаттайгаар ажилладаг бөгөөд эдгээрийг дараах байдлаар тодорхойлж болно.

1. Хүүхэд илрүүлэх
2. Хүүхдийн хөдөлгөөнийг мөрдөх
3. Аюултай бүсийн илрүүлэлт ба мэдэгдэл илгээх

A. Өгөгдлийн багц

Бид судалгааны ажилдаа дараах өгөгдлийн багцуудыг бэлтгэж ашигласан болно. Үүнд:

1. Сургалтын өгөгдлийн багц

1.1. DataSet1

DataSet1 өгөгдлийн багц нь бидний өөрсдийн бэлтгэсэн, нэг хүүхдийн 864 зурган өгөгдлөөс бүрдсэн.



1-р зураг. DataSet1 өгөгдлийн багц

1.2. DataSet2

DataSet2 өгөгдлийн багцыг Roboflow сайтаас цуглуулсан ба олон хүүхдийн өгөгдлөөр баяжуулж нийт 4041 зурган өгөгдлөөс бүрдсэн.



2-р зураг. DataSet2 өгөгдлийн багц

1.3. DataSet3

DataSet3 өгөгдлийн багц нь Roboflow сайтаас цуглуулсан ба насны хувьд ялгаатай өгөгдлөөр баяжуулж 8686 зурган өгөгдлөөс бүрдсэн.



3-р зураг. DataSet3 өгөгдлийн багц

2. Тестийн өгөгдлийн багц

2.1. TestSet 1

TestSet1 нь том хүн болон хүүхдийн өгөгдлөөс бүрдсэн 15 зурган өгөгдөл юм.



4-р зураг. TestSet1 өгөгдлийн багц

2.2. TestSet 2

TestSet2 нь өөрсдийн бэлтгэсэн нэг хүүхдийн 149 зургаас бүрдсэн өгөгдөл юм.



5-р зураг. TestSet2 өгөгдлийн багц

2.3. TestSet 3

TestSet3 нь өөр өөр хүүхдийн, сургалтын өгөгдөлд огт ороогүй 150 зурган өгөгдлөөс бүрдсэн.



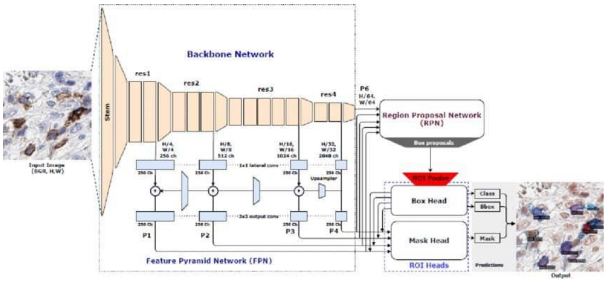
6-р зураг. TestSet3 өгөгдлийн багц

B. Хүүхэд илрүүлэх (Object Detection)

Бид судалгаандаа дараах гурван түгээмэл хэрэглэгддэг объект илрүүлэлтийн загварыг харьцуулан судалсан.

1. Detectron2

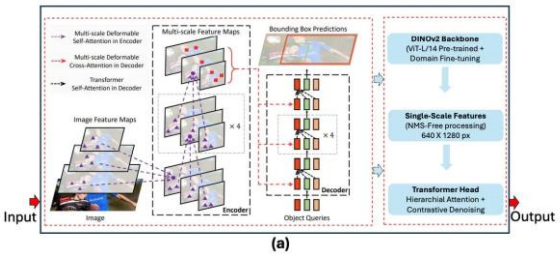
Facebook AI Research (FAIR)-ийн боловсруулсан Mask R-CNN framework бөгөөд bounding box, instance segmentation, keypoint detection зэрэг олон төрлийн илрүүлэлтийг дэмждэг. Энэ загвар нь Region Proposal Network (RPN) ашиглан боломжит объектын байрлалыг урьдчилан таамаглаж, дараа нь нарийн ангилал болон хил хязгаарын (bounding box) сайжруулалт хийдэг. Detectron2 нь өндөр нарийвчлалтай ч бодит цагийн хэрэглээнд боловсруулах хурд бага байдаг [4].



7-р зураг. Detectron2 архитектур

2. RF-DETR

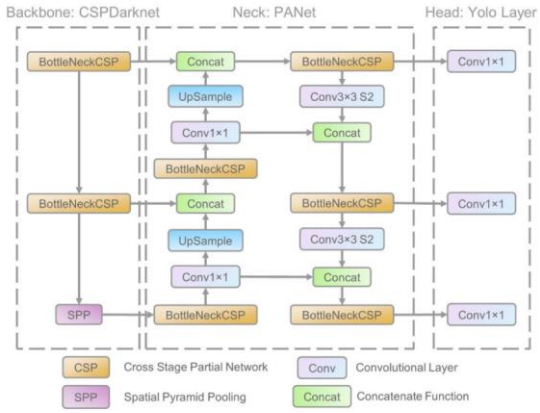
RF-DETR нь transformer encoder-decoder архитектур дээр суурилсан, DETR (Detection TRansformer) загварыг сайжруулсан хувилбар юм. Энэ загвар нь анхны DETR-ийн удаан суралцах асуудлыг шийдэхийн тулд refinement module болон IoU-aware query selection зэрэг механизмыг ашиглаж, илүү хурдан convergence ба илүү нарийн илрүүлэлтийг хангадаг [5].



8-р зураг. RF-DETR архитектур

3. YOLOv5

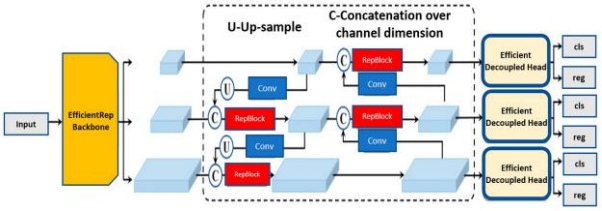
YOLOv5 нь CSPDarknet-д суурилсан Backbone-ийг ашиглан онцлог мэдээллийг үр дүнтэй ялгаруулж, PANet (Path Aggregation Network)-ийг Neck болгон ашиглаж, олон масштабын онцлогуудыг (features) нэгтгэн дамжуулдаг. Энэ загвар нь хурд ба нарийвчлалын хоорондын тэнцвэрийг сайн барьдаг бөгөөд жижиг хэмжээний объектуудыг илрүүлэхэд үр дүнтэй. Гэвч илүү шинэ загваруудтай харьцуулахад нарийн төвөгтэй дүрс (жишээ нь, хөдөлгөөн бүхий хүүхдийн дүрс) дээр тодорхой хэмжээнд алдаа гаргадаг *[6]*.



9-р зураг. YOLOv5 архитектур

4. YOLOv8

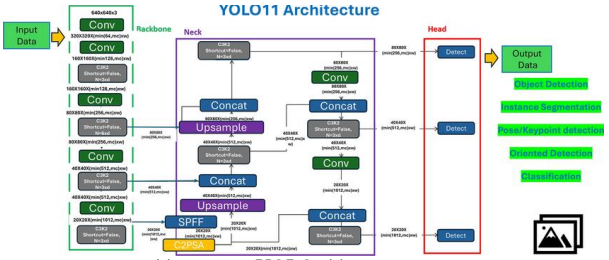
YOLOv8 нь Ultralytics-аас гаргасан хамгийн түгээмэл хэрэглэгдэж буй YOLO загвар бөгөөд anchor-free илрүүлэлтийн аргачлалтай. YOLO нь ганцхан үе шаттай (single-shot) алгоритм бөгөөд зөвхөн нэг мэдрэлийн сүлжээг ашиглан оролтын зургийг бүхэлд нь боловсруулж объектыг танидаг тул бодит цагийн объект илрүүлэлт хийхэд тохиромжтой. Мөн feature extraction-ийн шинэ C2f module, илүү үр ашигтай loss function болох TaskAlignedAssigner, мөн multi-scale training аргуудыг ашиглан илрүүлэлтийн хурд, нарийвчлал хоёуланг нь сайжруулсан [7].



10-р зураг. YOLOv8 архитектур

5. YOLOv11

YOLOv11 нь өмнөх YOLO загвараас ялгаатай нь бодит цагийн гүйцэтгэл болон нарийвчлалыг нэмэгдүүлэх зорилгоор архитектурын томоохон шинэчлэлүүдийг хийсэн. Тухайлбал, уламжлалт C2f блокыг илүү үр ашигтай C3k2 блокоор сольж, тооцооллыг оновчтой болгосон. Хамгийн гол нь, attention mechanism ашигласан C2PSA блокыг нэвтрүүлснээр загварын анхаарлыг зураг доторх хамгийн чухал хэсгүүдэд төвлөрүүлж, илрүүлэлтийн нарийвчлалыг мэдэгдэхүйц нэмэгдүүлсэн [8].



11-р зураг. YOLOv11 архитектур

C. Хүүхдийн хөдөлгөөнийг мөрдөх

Объект илрүүлэлтийн загвараар (YOLOv8) хүүхдийг таньсны дараа, тухайн хүүхдийн байрлалыг фреймээс фреймд тасралтгүй хянах нь аюултай бүсэд нэвтэрсэн эсэхийг тогтоохын тулд зайлшгүй шаардлагатай. Энэхүү алхамыг Object Tracking алгоритмууд гүйцэтгэдэг. Судалгаанд гурван түгээмэл мөрдөлтийн алгоритмыг (DeepSORT, ByteTrack, BoT-SORT) туршиж, тэдгээрийн нарийвчлал, илрүүлэлтийн чадвар болон бодит цагийн гүйцэтгэлийг харьцуулан үнэлэв.

1. DeepSORT

DeepSORT (Deep Simple Online and Realtime Tracking) нь SORT алгоритмын сайжруулсан хувилбар бөгөөд Калманы шүүлтүүр (Kalman Filter) ашиглан объектын дараагийн байрлалыг таамагладаг. Тухайн илрүүлэлтийг өмнөх мөрдөлтийн объекттой холбохдоо хоёр үе шаттай тохирох (matching) механизм хэрэглэдэг [9] [10].

- Махаланобисын зай (Mahalanobis distance)-г ашиглан хамгийн ойр илрүүлэлтүүдийг шүүнэ.

$$d_M(x, y) = \sqrt{(x-y)^T S^{-1}(x-y)} \qquad (1)$$

- Косинусын зай (Cosine distance)-г ашиглан объектуудын дүрсийн төстэй байдлыг хэмжинэ.

$$d_C(a, b) = 1 - \frac{a \cdot b}{||a||||b||} \quad (2)$$

Ийнхүү DeepSORT нь объектын identity буюу ялгагдах байдлыг хадгалан, урт хугацааны мөрдөлтөд тогтвортой ажилладаг. Гэсэн хэдий ч илрүүлэлтийн чанар муудсан нөхцөлд андуурах магадлал бага зэрэг нэмэгддэг.

2. BoT-SORT

BoT-SORT (Bag of Tricks for SORT) нь DeepSORT-ийн суурин дээр боловсронгуй болгосон алгоритм бөгөөд Re-identification (ReID) загвар ашиглан дүрсний онцлогийг илүү нарийвчлан ялгадаг [11].

BoT-SORT-ийн гол онцлогууд:

- Camera motion compensation - камерууд хөдөлж байгаа нөхцөлд дүрсийг тогтворжуулж, илрүүлэлтийн координатыг засдаг.
- IOU-based matching - объектуудын давхцлын хэмжээг (Intersection over Union) ашиглан хамгийн боломжит хослолыг олдог.
- Appearance embedding - дүрсний өнгө, хэлбэр, бүтэц зэрэг онцлогийг вектор хэлбэрээр хадгалж, ижил объектуудыг илүү найдвартай ялган таних боломж олгодог.

3. ByteTrack

ByteTrack нь объект илрүүлэлтийн оноонд суурилсан уламжлалт хандлагыг сайжруулж, өндөр (high-score) болон бага (low-score) итгэлцэлтэй илрүүлэлтүүдийг хоёуланг нь ашигладаг.

Ингэснээр бага итгэлцэлтэй ч бодит объектуудыг хаяхгүй, илүү бүрэн гүйцэт мөрдөлт хийх боломжтой болдог. Энэ нь track fragmentation буюу мөрдөлтийн тасралтыг багасгадаг [12].

D. Аюултай бүсийн илрүүлэлт ба мэдэгдэл илгээх

Бидний хөгжүүлсэн энэхүү систем нь хүүхдийн аюултай бүсийг тодорхойлохдоо хоёр аргыг хослуулсан байдлаар ажилладаг. Үүнд:

1. Аюултай бүс тэмдэглэх

Хэрэглэгч интерфэйс дээр суурин камерын хувьд өөрийн хүссэн байрлалд олон цэг сонгон холбож, аюултай бүсийг полигон хэлбэрээр тодорхойлдог.



12-р зураг. Аюултай бүсийг гараар тэмдэглэх аргын хэрэгжүүлэлт

Энэ арга нь гэртээ байгаа тогтмол байршилтай хөдөлгөөнгүй дурын обьектуудыг аюултай бүс хэмээн тодорхойлж өгөхөд тохиромжтой. Жишээ : цонх, зуух, зурагт, шат гэх мэт

2. Аюултай объект таних арга

Объект таних загварыг ашиглан ахуйн орчны аюултай объектуудыг автоматаар илрүүлж, тухайн объект орчмын бүсийг аюултай бүсээр тэмдэглэдэг.



13-р зураг. Аюултай объектыг загвар ашиглан таних арга

Ингэхдээ бид туршилтаар плитка, шарах шүүгээ, богино долгионы зуух, разетка уртасгагч гэх дөрвөн объектуудыг сонгож авч өөрсдийн бэлтгэсэн өгөгдлөөр өмнө дурдагдсан YOLOv8 загварыг fine-tune хийн сургасан. Өгөгдлөө бэлдэхдээ өгөгдлийн баяжуулалт хийсэн ба доорх хүснэгтэд өмнөх болон дараах үр дүнг харуулав.

- Flip: Horizontal, Vertical
- 90° Rotate: Clockwise, Counter-Clockwise, Upside Down
- Rotation: Between -15° and +15°
- Shear: ±10° Horizontal, ±10° Vertical
- Blur: Up to 1.0px

1-р хүснэгт. Өгөгдлийн баяжуулалт хийсэн үр дүн

Төрөл	**Өмнөх өгөгдөл**	**Data augmentation**
Шарах шүүгээ	167	835

Богино долгионы зуух	136	680
Плитка	96	480
Залгуур	189	945
Нийт	588	2940

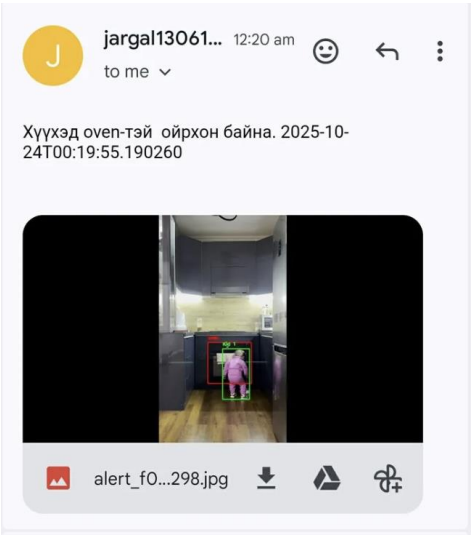
Дээрх 2 аргыг хослуулан аюултай бүсийг илрүүлэлт хийгдсэний дараа систем тухайн объект орчмын координатыг аюултай бүсийн хил болгон тэмдэглэж, хүүхдийн хөдөлгөөний мөрдөлттэй (tracking trajectory) харьцуулдаг. Хэрэв хүүхдийн bounding box ба аюултай бүсийн polygon хоорондын огтлолцол (Intersection over Union, IoU) ба хугацааны хувьд босго утгыг (IoU = 0.3, хугацаа = 5 секунд) давбал аюул илэрсэн гэж үзэж эцэг эхэд мэдэгдэл илгээдэг. Ингэхдээ доорх гурван аргаар илгээж болно.

$$IoU = \frac{Area_{int}}{Area_A + Area_B - Area_{int}} \quad (3)$$

$Area_A$ - аюултай бүсийн bounding box
$Area_B$ – хүүхдийн bounding box
$Area_{int}$ – аюултай бүс болон хүүхдийн bounding box огтлолцол

1. E-Mail

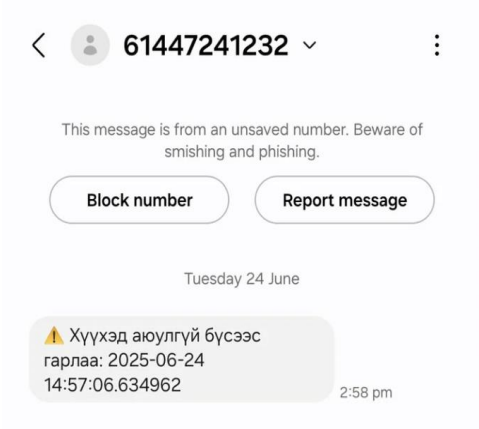
Тухайн үеийн дүрс болон огноо, цагийн мэдээлэл, анхааруулах зурвас бүхий мэйл эцэг эхэд илгээгдэнэ.



14-р зураг. Имэйлээр мэдэгдэл илгээх арга

2. Мессеж

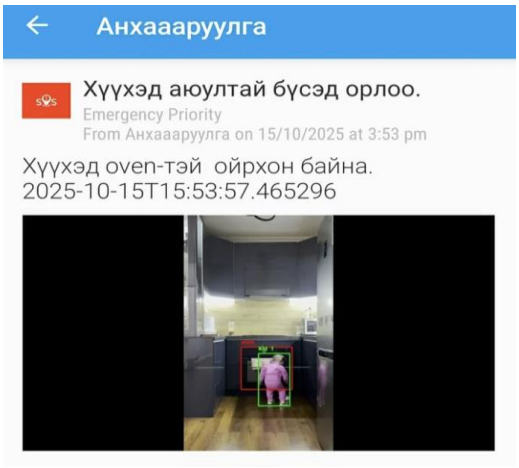
Энэхүү арга нь аюул илэрсэн үед автоматаар ClickSend API-г ашиглан хэрэглэгчийн бүртгэлтэй гар утас руу SMS мэдэгдэл илгээдэг. Энэхүү API нь cloud-based SMS үйлчилгээ бөгөөд HTTP хүсэлтээр дамжуулан илгээх боломжтой тул найдвартай, шуурхай ажиллагаатай. Уг арга нь эцэг эх интернэтэд холбогдоогүй үед ч ирдгээрээ давуу талтай ба анхааруулах зурвас, огноо, цагийн мэдээлэл бүхий мессеж илгээгдэнэ.



15-р зураг. Мессежээр мэдэгдэл илгээх арга

3. Application

Мөн эцэг эхэд гар утасны аппликейшнээр SOS alert илгээдэг ба энэхүү төрлийн мэдэгдэл нь илүү яаралтай дохиоллын зориулалттай ба гар утсанд чанга чимээ гаргаж, богино хугацаанд чичирхийлж эцэг эхийн анхаарлыг татна. Ингэхдээ Pushover API-г ашиглан систем нь серверээс хэрэглэгчийн token болон user key-г дамжуулж мэдэгдэл илгээдэг. Илгээгдсэн мэдэгдэл нь хэрэглэгчийн төхөөрөмж дээр real-time push notification хэлбэрээр гарч ирдэг [13].



16-р зураг. Application-р мэдэгдэл илгээх арга

IV. Үр дүн

Судалгааны хүрээнд гүн сургалтын загваруудыг боловсруулах, сургах болон туршилтын үр дүнг тооцоход 1.4 GHz Quad-Core Intel Core i5 процессор, 8 GB 2133 MHz LPDDR3 санах ой болон Intel Iris Plus Graphics 645 (1536 MB) график карт бүхий тооцоолуурыг ашигласан. Загваруудыг сургахдаа 640x640 пикселийн хэмжээтэй оролтын дүрс дээр, тус бүр нийт 100 epoch сургалтыг гүйцэтгэсэн.

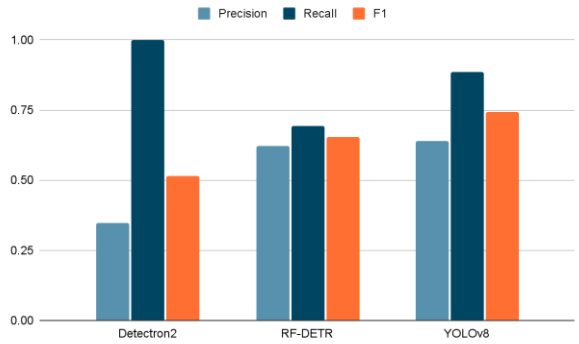
1. *Хүүхэд таних загвар*

a. Архитектурын харьцуулалт

Системийн эхний шатанд бид хүүхдээ танихдаа гурван өөр архитектурыг (RF-DETR, Detectron2, YOLOv8) *TestSet1* дээр туршсан үр дүнг доор харуулав.

2-р хүснэгт. Архитектурын харьцуулалт хийсэн үр дүн

Model	Detectron2	RF-DETR	YOLOv8
TP	26	18	23
FP	49	11	13
FN	0	8	3
Precision	0.347	0.621	0.639
Recall	1	0.692	0.885
F1	0.515	0.655	0.742

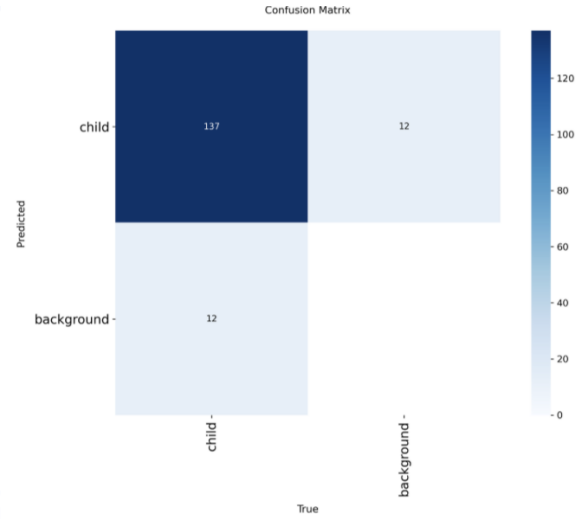


17-р зураг. Архитектурын харьцуулалтын график үр дүн

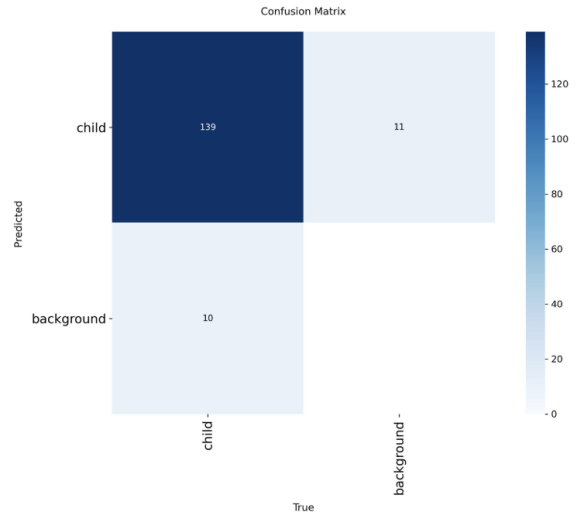
Туршилтын үр дүнгээс харахад Detectron2 нь Recall 1.0 буюу бүх объектыг илрүүлсэн боловч Precision маш бага (0.347) буюу хүүхдийг том хүнтэй андуурах тохиолдол их байсан. RF-DETR нь үзүүлэлтийн хувьд дундаж гүйцэтгэлтэй байсан. YOLOv8 нь Precision (0.639), Recall (0.885) гэх үзүүлэлтээр уг гурван архитектураас хамгийн өндөр байна. Иймд бид YOLOv8-ыг хүүхэд илрүүлэх үндсэн архитектур болгон сонгосон.

b. YOLO цувралын харьцуулалт

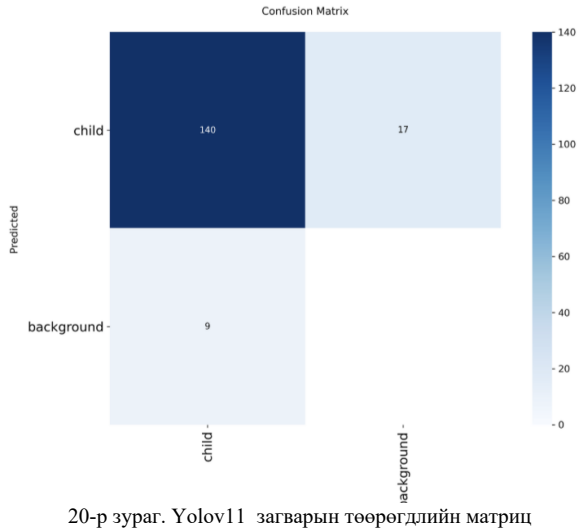
Архитектураа сонгосны дараа YOLO-ийн өөр хувилбарууд (v5, v8, v11)-ийг нарийвчлан харьцуулж, аль нь хамгийн тохиромжтойг туршсан. Ингэхдээ өөрсдийн бэлдсэн 864 өгөгдлийг буюу *DataSet1* багцыг ашиглан сургасан.



18-р зураг. Yolov5 загварын төөрөгдлийн матриц



19-р зураг. Yolov8 загварын төөрөгдлийн матриц



20-р зураг. Yolov11 загварын төөрөгдлийн матриц

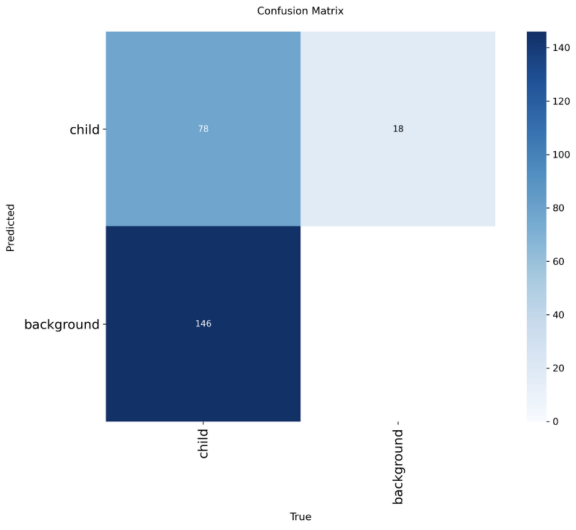
Загваруудын *TestSet2* багц дээрх нарийвчлалыг тус тус тооцон гаргахад Yolov5 хувилбар 85%, Yolov8 хувилбар 87%, Yolov11 хувилбар 84% гарсан. Бүх хувилбарын үзүүлэлтийн хооронд ялгаа бага байсан.

Эцсийн байдлаар судалгааны ажилдаа ашиглах хүүхэд таних загвараар үзүүлэлт хамгийн өндөр буюу 87%-ийг үзүүлэлттэй байсан Yolov8-ийг сонгов.

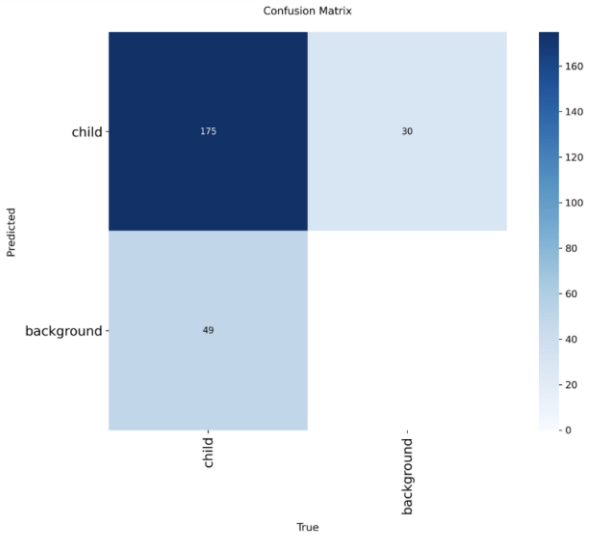
с. Өгөгдлийн харьцуулалт

Сонгосон Yolov8 загвар дээр *TestSet3* ашиглан туршилт хийсэн ба загварын гүйцэтгэл 87% - с 32% нарийвчлалтай болж буурсан. Загварыг нэг хүүхдийн өгөгдлөөр сургах нь гүйцэтгэлийг бууруулж, олон хүүхдийн зурагтай өгөгдлөөр баяжуулахаас гадна өгөгдлийн хэмжээг нэмэгдүүлж туршиж үзсэн. *Dataset2* (4041 зураг)-ыг ашиглан сургалт явуулахад *TestSet3* багц дээр 68%-ийн гүйцэтгэл үзүүлсэн бол *Dataset3* (8686 зураг)-ыг ашиглан сургалт хийснээр гүйцэтгэл 92% хүртэл өссөн байна. Үүнээс харахад өгөгдлийн хэмжээг нэмэгдүүлснээр загварын гүйцэтгэл сайжирч байна.

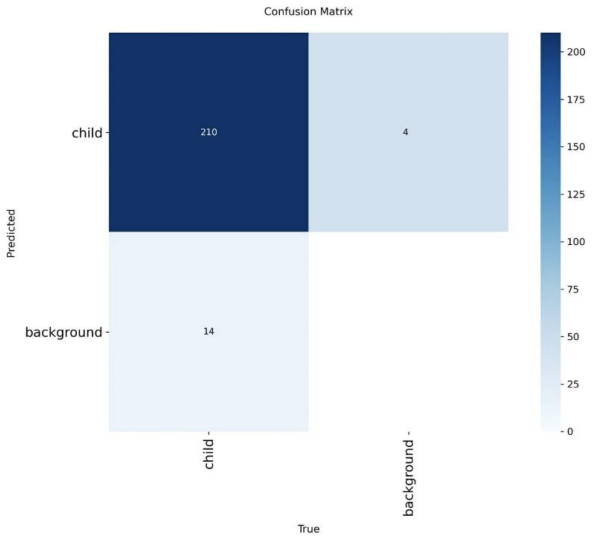
Загвар	Dataset1	Dataset2	Dataset3
YoloV8	32%	68%	92%



21-р зураг. DataSet1 (864) өгөгдөл бүхий загварын төөрөгдлийн матриц



22-р зураг. DataSet2 (4041) өгөгдөл бүхий загварын төөрөгдлийн матриц



23-р зураг. DataSet3 (8686) өгөгдөл бүхий загварын төөрөгдлийн матриц

2. *Object tracking алгоритмуудын харьцуулалт*

Судалгааны хүрээнд ByteTrack, DeepSORT, BoT-SORT алгоритмуудыг харьцуулан туршсан.
Object tracking алгоритмуудын гүйцэтгэлийг дараах үзүүлэлтүүдээр хэмждэг [14]. Үүнд:

- True Positive (TP) - Зөв илрүүлсэн объектуудын тоо
- False Positive (FP) - Буруу илрүүлэлт (байхгүй объект илэрсэн)
- False Negative (FN) - Илрүүлж чадаагүй бодит объектууд
- Precision - Илрүүлсэн объектуудаас хэд нь зөв байсан эсэх
- Recall - Бодит объектуудаас хэдийг илрүүлсэн эсэх
- IDSW - ID солигдлын тоо
- F1-score - Precision баRecall дундаж үзүүлэлт
- C_t – тухайн $frame$ илрүүлсэн объект ба GT хоорондын зай

- $d_{t,i}$ – тухайн $frame$ илрүүлсэн объектын тоо
- GT_t – тухайн $frame$ байгаа бодит объектын тоо

$$F1 = 2 \times \frac{\Pr e\,cision \times Recall}{\Pr e\,sion + Recall} \quad (4)$$

- MOTA - Нийт хянах чадварыг хэмжинэ.

$$MOTA = 1 - \frac{\Sigma_t(FN_t + FP_t + IDSW_t)}{\Sigma_t GT_t} \quad (5)$$

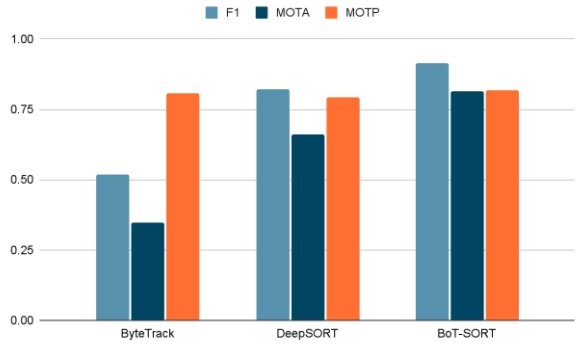
- MOTP - Хянасан объектуудын байрлал хэр зэрэг оновчтой байгааг илтгэнэ.

$$MOTP = \frac{\Sigma_{t,i} d_{t,i}}{\Sigma_t C_t} \quad (6)$$

3-р хүснэгт. Tracking алгоритмуудын харьцуулалтын үр дүн

Алгоритм	ByteTrack	DeepSORT	BoT-SORT
TP	134	278	321
FP	2	20	2
FN	245	101	58
Precision	0.985	0.933	0.994
Recall	0.354	0.734	0.847
IDSW	1	7	11
F1	0.52	0.821	0.915
MOTA	0.346	0.662	0.813
MOTP	0.808	0.793	0.818

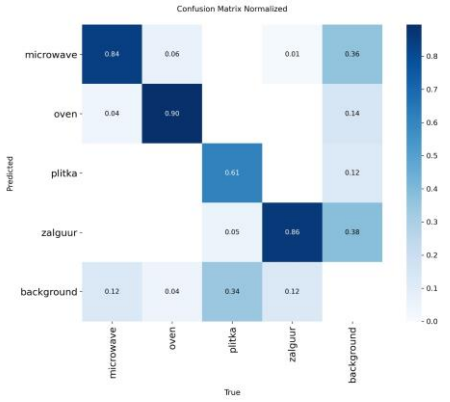
Туршилтын үр дүнгээс харахад BoT-SORT алгоритм нь хамгийн өндөр Precision (0.994) болон Recall (0.847)-тэй гарч, нийт үнэлгээ болох MOTA = 0.813 үзүүлэлтээр бусад аргуудаас илүү байв.



24-р зураг. Tracking алгоритмуудын гүйцэтгэлийг харьцуулсан график үзүүлэлт

3. *Аюултай бусийн илрүүлэлт*

Бидний өөрсдийн бэлтгэсэн өгөгдөл дээр YOLOv8 загварыг fine-tune хийсний үр дүнд шарах шүүгээ (0.90), богино долгионы зуух (0.84), залгуур (0.86) гэх үзүүлэлттэй байна. Харин плитка (0.61) ангиллын хувьд үзүүлэлт харьцангуй муу гарсан ба зарим тохиолдолд background гэж андуурагдсан байдал ажиглагдаж байна. Тестийн дундаж нарийвчлал (0.8025) байв.



25-р зураг. Аюултай объект таних загварын төөрөгдлийн матриц

Эдгээр туршилтын явцад дараах алдаа ажиглагдсан.

- Том хүнтэй андуурах - Зарим загварууд, ялангуяа Detectron2 нь тонгойж суугаа, бөхийсөн эсвэл мөлхөж буй том хүнийг хүүхэдтэй андуурч илрүүлэх тохиолдол гарсан. Энэ нь дүрсэн дэх биеийн хэмжээ болон хэлбэрийн төстэй байдлаас шалтгаалж байна.

- Хэсэгчлэн халхлагдсан байдал - Хүүхэд тавилга болон бусад объектын ард орж биеийн хэсэг нь харагдахгүй үед илрүүлэлтийн нарийвчлал буурч, зарим фреймд танигдахгүй байсан. BoT-SORT алгоритм нь хөдөлгөөнийг таамаглан мөрдөх боломжтой боловч удаан хугацааны халхлалтын үед ID солигдох алдаа ажиглагдсан.

- Гүний мэдээлэл дутмаг байдал (Depth limitation): 2D дүрсэнд суурилсан IoU тооцоолол нь хүүхэд аюултай объектын наана эсвэл цаана байрлаж буй эсэхийг нарийн ялгаж чадахгүй тохиолдол гарсан. Үүний улмаас бодит хүрэлцээгүй нөхцөлд хуурамч анхааруулга (False Alarm) үүсэх боломжтой байв.

V. Дүгнэлт

Энэхүү судалгааны ажлаар бага насны хүүхдийн аюулгүй байдлыг хангах зорилгоор бодит цагийн ухаалаг хяналтын системийн шийдэл санал болголоо.

Судалгаанд объект илрүүлэлтийн загвар, мөрдөлтийн алгоритм, аюултай бүс илрүүлэх загваруудыг харьцуулан, хамгийн оновчтой загварыг сонгов.

Хүүхэд илрүүлэх загварын хувьд Detectron2, RF - DETR, YoloV5, YoloV8, YoloV11 загваруудыг ашиглан туршилт хийсэн ба тестийн өгөгдөл дээр YoloV8 нь 87% нарийвчлалтай хамгийн оновчтой ажиллаж байв.

Анхны өгөгдлийн туршилт үр дүн: Анхны сургалтын 864 зурган өгөгдөл дээр олон төрлийн хүүхдийн өгөгдлөөр туршилт хийхэд 32% болж буурсан.

Өгөгдлийн хэмжээг нэмэгдүүлсэн туршилт:

Олон хүүхдийн зургаар сургалтын өгөгдлөө баяжуулсны дараа тестийн багц дээр 68% болж сайжирсан. Өгөгдлийн хэмжээ нэмэгдэх тусам гүйцэтгэл нь 92% болж сайжирсан. Энэ нь олон ялгаатай хүүхдийн өгөгдлөөр загварыг сургах нь загварын гүйцэтгэлийг сайжруулж байгааг харуулж байна. Мөн өгөгдлийн хэмжээг нэмэгдүүлэхэд загварын нарийвчлал сайжрав.

Хүүхэд мөрдөх алгоритмын хувьд BoT-SORT бодит цагт хүүхдийг тасралтгүй мөрдөж гүйцэтгэлийн хувьд хамгийн үр дүнтэй буюу MOTA(0.813), MOTP(0.818), F1(0.915) алгоритм байв.

Аюултай бүс илрүүлэлтийн хувьд шарах шүүгээ, богино долгионы зуух, залгуур, плитка гэсэн 4 төрлийн объектоос бүрдсэн 2940 зурган өгөгдлийг ашиглаж YoloV8 дээр fine-tune хийж сургасан. Тестийн дундаж нарийвчлал (0.8025) байв. Үүнээс ганцхан объектоор нь танилт хийх нь уян хатан байдлыг бууруулж байсан учраас суурин камерын хувьд гар аргаар аюултай бүсийг нэмж тодорхойлж өгөх нь загварын нарийвчлалыг сайжруулж байв.

Мэдэгдлийг зөвхөн нэг аргаар буюу мессежээр илгээх уян хатан байдлыг бууруулж байсан ба нэмэлтээр email хаяг болон pushover ашиглан илгээх нь мэдэгдлийг анзаарахгүй өнгөрөх эрсдэлийг бууруулж, найдвартай байдлыг нэмэгдүүлсэн.

Судалгааны үр дүнгээс харахад YoloV8 загвар нь өгөгдлийн хэмжээ болон төрлөөс хамааран үр дүнтэй ажиллах чадвартай нь батлагдлаа. BoT-SORT алгоритм нь бодит цагийн мөрдөлтийг тогтвортой гүйцэтгэж чадсан бөгөөд аюултай бүсийг илрүүлэх системтэй хослуулснаар хүүхдийн аюулгүй байдлыг хангахад үр дүнтэй шийдэл болохыг харуулсан.

Бид энэхүү судалгааны ажлаараа бага насны хүүхдийн ахуйн орчны аюулгүй байдлыг хангах, осол гэмтлээс урьдчилан сэргийлэх зорилгоор ухаалаг хяналтын системийг санал болголоо. Энэхүү систем нь хүүхдийг илрүүлж, аюултай бүсэд нэвтэрсэн тохиолдолд бодит цагийн анхааруулга илгээх замаар эцэг эхийн хяналтыг дэмжиж, ахуйн орчны аюулгүй байдлыг сайжруулахад хувь нэмэр оруулах технологийн шийдэл юм.

VI. Системийн хязгаарлалт ба цаашид хийгдэх ажлууд

Энэхүү судалгааны хүрээнд боловсруулсан систем нь хүүхдийн аюулгүй байдлыг хангахад үр дүнтэй болох нь туршилтаар батлагдсан боловч бодит хэрэглээний нөхцөлд дараах хязгаарлалтууд байна.

Нэгдүгээрт, систем нь камерын дүрс дэх орон зайн гүнийг (depth) мэдрэх чадваргүй юм. Энэ нь хүүхэд аюултай объектын наана эсвэл цаана, ямар нэгэн биет хүрэлцэлгүйгээр зөрөх үед bounding box-ууд давхцсанаар хуурамч дохиолол (False Alarm) үүсэх үндсэн шалтгаан болж байна. Хоёрдугаарт, туршилтын өгөгдөл болон мөрдөлтийн алгоритм нь одоогоор нэг хүүхдийг хянахад илүү оновчтой бөгөөд олон хүүхэд нэгэн зэрэг хөдөлгөөн үйлдэх, бие биеийг халхлах нөхцөлд ID солигдох болон хяналтын нарийвчлал буурах магадлалтай байна. Гуравдугаарт, систем нь хөдөлгөөнгүй (fixed) камер дээр суурилан аюултай бүсийг тодорхойлж байгаа тул камерын өнцөг өөрчлөгдөх эсвэл гэрэлтүүлэг огцом буурах үед (жишээ нь, шөнийн цагт) илрүүлэлтийн чанар муудах эрсдэлтэй. Түүнчлэн, одоогийн үр дүн нь богино хугацааны хяналттай орчинд хийгдсэн туршилт дээр суурилсан тул бодит гэрийн нөхцөлд урт хугацааны (long-term deployment) ажиллагааг баталгаажуулах туршилт хараахан хийгдээгүй байна.

Цаашид гүний мэдрэгч бүхий камер (RGB-D) ашиглан орон зайн нарийвчлалыг сайжруулах, олон хүүхэдтэй болон гэрэлтүүлэг багатай орчинд загварыг оновчлох, хүүхдийн хөдөлгөөнөөс гадна үйлдлийг (action recognition) таньж мэдэх нь бидний ирээдүйн судалгааны гол чиглэл болох юм.

НОМ ЗҮЙ

[1] Н. Б., “Хүүхэд ахуйн осолд өртөж гэмтэх нь нэмэгдэж байна,” 2025. [Холбогдсон]. Available: https://news.mn/r/2775344/. [Нэвтэрсэн 09 01 2025].

[2] P. H. L. Yujia Zhang, “A SMART CHILD SAFETY SYSTEM FOR ENHANCED POOL SUPERVISION USING COMPUTER VISION AND MOBILE APP INTEGRATION”.

[3] C. [Холбогдсон]. Available: https://us.getcubo.com/?srsltid=AfmBOor70l_5HX5ZMykpBLE7nPuGbdtWu3iOySQQQJ2JjsJH5r10ODTC.

[4] G. Boesch, “Detectron2: A Rundown of Meta’s Computer Vision Framework,” 11 2 2024. [Холбогдсон]. Available: https://viso.ai/deep-learning/detectron2/.

[5] Peter Robicheaux, James Gallagher, Joseph Nelson, Isaac Robinson, “RF-DETR: A SOTA Real-Time Object Detection Model,” 20 03 2025. [Холбогдсон]. Available: https://blog.roboflow.com/rf-detr/.

[6] T. M, “YOLO Object Detection: An Introduction,” 19 6 2025. [Холбогдсон]. Available: https://blog.roboflow.com/yolo-object-detection/.

[7] “What is YOLOV8? A Step-By-Step Guide,” 2024. [Холбогдсон]. Available: https://www.folio3.ai/blog/what-is-yolov8-architecture/.

[8] Ranjan Sapkota , Marco Flores-Calero, Rizwan Qureshi,Chetan Badgujar, “YOLO advances to its genesis: a decadal and comprehensive,” p. 19, 2025.

[9] A. Kouidri, “Mastering Deep Sort: The Future of Object Tracking Explained,” 11 10 2023. [Холбогдсон]. Available: https://www.ikomia.ai/blog/deep-sort-object-tracking-guide.

[10] “Multi-Object Tracking with DeepSORT,” [Холбогдсон]. Available: https://www.mathworks.com/help/vision/ug/multi-object-tracking-with-deepsort.html.

[11] KEECHIN, “Implementing Object Tracking for Computer Vision,” 21 7 2022. [Холбогдсон]. Available: https://datature.io/blog/implementing-object-tracking-for-computer-vision.

[12] “An Introduction to BYTETrack: Multi-Object Tracking by Associating Every Detection Box,” [Холбогдсон]. Available: https://datature.io/blog/introduction-to-bytetrack-multi-object-tracking-by-associating-every-detection-box.

[13] “Pushover,” [Холбогдсон]. Available: https://pushover.net/.

[14] “MOTChallenge: A Benchmark for Single-Camera Multiple Target Tracking,” 2020. [Холбогдсон]. Available: https://link.springer.com/article/10.1007/s11263-020-01393-0.

[15] A. V. “Enhancing Child Safety with Computer Vision,” 2023.

[16] S. T. “Child Monitoring and Data Analysis via Computer Vision,” 2024.

[17] K.-L. L. “An Image-Based Fall Detection System for the Elderly,” 2018.

Веб камераар, нүүрний цэгүүдэд суурилан харааны чиглэл таамаглах машин сургалтын загвар сургалт, тохируулга хийх нь

Лхагвадоржийн Сэргэлэн
Мэдээлэл, компьютерын ухааны тэнхим
Монгол улсын их сургууль
Улаанбаатар, Монгол улс
22B1NUM0796@stud.num.edu.mn

Чагнаагийн Алтангэрэл
Мэдээлэл, компьютерын ухааны тэнхим
Монгол улсын их сургууль
Улаанбаатар, Монгол улс
altangerel@num.edu.mn

Очирбатын Билгүүн*
Мэдээлэл, компьютерын ухааны тэнхим
Монгол улсын их сургууль
Улаанбаатар, Монгол улс
bilguun.o@num.edu.mn

***Хураангуй*—Энэхүү судалгааны ажлаар ердийн веб камерын тусламжтайгаар хүний харааны чиглэлийг (gaze) тодорхойлох машин сургалтын загвар хөгжүүлэхийг зорьсон. MediaPipe FaceMesh гүн сургалтын загварыг ашиглаж нүүрний гол цэгүүдийг (facial landmarks) илрүүлж, тэдгээрийг загвар сургах шинж чанар болгон ашигласан. Өгөгдөл цуглуулах зорилгоор веб аппликейшн хөгжүүлж, 96 оролцогчоос нийт 4,416 түүвэр өгөгдөл цуглуулсан. Судалгаанд Ridge Regression, Random Forest, Extreme Gradient Boosting болон PyTorch MLP загваруудыг сургаж, тэдгээрийг хооронд нь харьцуулж үнэлэв. Загваруудын үнэлгээнд Mean Absolute Error (MAE) ашигласан бөгөөд 10-fold cross validation туршилтын өгөгдөл дээр хамгийн сайн загвар нь 141.3 пиксель алдаатай байсан. Загварын таамаглалыг сайжруулахын тулд 16 цэгийн тохируулгын (calibration) процесс болон аффин хувиргалтыг хэрэгжүүлсэн. Энэхүү судалгааны ажил нь хүний зан төлөв, анхаарал төвлөрлийн судалгаанд ашиглагдах боломжтой бөгөөд бага өртөгтэй "eye-tracking" шийдэл хэрэгжүүлэх боломжийг судаллаа.**

***Түлхүүр үг*—харааны чиглэл, нүдний хараа дагах, eye gaze, eye tracking, MediaPipe, нүүрний цэгүүд, тохируулга, веб камер, машин сургалт**

I. Удиртгал

Хүний нүд нь гадаад орчноос мэдээлэл хүлээн авах гол суваг төдийгүй, хүний анхаарал төвлөрөл, сонирхол, танин мэдэхүйн үйл ажиллагааг илэрхийлэгч чухал үзүүлэлт юм. Харааны чиглэлийг дэлгэцийн координат руу бодит хугацаанд буулгах нь хүн-компьютерын харилцаа (HCI), хөдөлгөөний хязгаарлалттай хэрэглэгчдэд зориулсан туслах технологи (assistive technology), хэрэглэгчийн зан төлөвийн шинжилгээ болон зах зээлийн судалгаа зэрэг олон салбарт практик ач холбогдолтой.

Уламжлалт нүдний хараа дагах системүүд (eye trackers) нь хэт улаан туяаны мэдрэгч зэрэг тусгай тоног төхөөрөмж ашиглан өндөр нарийвчлалтай үр дүн гаргадаг боловч өндөр өртөгтэй, ашиглахад нарийн тохируулга шаарддаг, хэрэглээний хувьд хязгаарлагдмал байдаг. Үүнийг шийдвэрлэх орчин үеийн арга нь энгийн RGB веб камер ашиглан "Appearance-Based" буюу дүрсэд суурилсан харааны чиглэл таамаглах систем юм. Энэхүү арга нь тусгай техник хангамж шаардахгүйгээр дурын төхөөрөмж дээр ажиллах боломжтой гэдгээрээ давуу талтай.

Иймд энэхүү судалгааны ажлаар бид энгийн веб камерт суурилсан, бага өртөгтэй бөгөөд хүртээмжтэй харааны чиглэл тодорхойлох машин сургалтын загварыг хөгжүүлэхийг зорьсон. Бид Google-ийн MediaPipe FaceMesh загварыг ашиглан нүүрний 478 цэгийн (3D landmarks) координатыг илрүүлж, тэдгээрээс харааны чиглэлийг таамаглахад шаардлагатай шинж чанаруудыг (features) ялган авсан.

II. Холбогдох Судалгаа

A. Төхөөрөмжид суурилсан нүд дагах систем

Анхны нүд дагах системүүд эвэрлэгийн тусгалыг үүсгэх хэт улаан туяаны мэдрэгч, өндөр фрэймийн хурдтай камер, тусгай оптик зэрэг тоног төхөөрөмж ашигладаг байсан [1]. Tobii зэрэг арилжааны төхөөрөмжүүд нь нарийвчлал сайн боловч өндөр өртөгтэй, мөн хүн тус бүр дээр нарийн тохируулга шаарддаг. Эдгээр асуудал нь хөгжүүлэлтийг зөвхөн лабораторийн орчноор нь хязгаарладаг.

B. Apprearance-Based харааны чиглэл таамаглах

Appearance-Based аргууд нь тусгай тоног төхөөрөмжгүйгээр нүүр эсвэл нүдний зургаас харааны чиглэлийг тооцоолдог [2]. Энэ арга нь их хэмжээний өгөгдөл цуглуулалт болон веб камерын зурагт тулгуурлан нарийвчлалтай загвар сургах боломжтой. Нүдний хэсгийн зураг дээр шууд ажилладаг, толгойн байрлал оруулсан CNN ашиглаж, тоног төхөөрөмжид суурилсан нүд хянагчтай өрсөлдөхүйц нарийвчлалтай байх боломжтойг [3] харуулсан. Мөн синтетик сургалтын өгөгдөл нэмэлтээр ашиглаж таамаглалыг сайжруулж болохыг харуулсан [4] бол iOS хамтын оролцоот программаар 1,400 гаруй оролцогчийн өгөгдөл цуглуулж, таблет дээр 1.71 см-ийн дундаж алдаа гаргадаг multi-stream CNN сургасан [5] байна.

C. Нүүрний цэгэн өгөгдөлд суурилсан арга

Зурган өгөгдөл дээр шууд ажиллахын оронд хэд хэдэн судалгааны ажлууд нүүрний онцлог цэгүүдийг (facial

landmarks) ашигласан байдаг. Онцлог цэгүүд нь өгөгдлийн хэмжээ бага буюу авсаархан, зургийн нягтаршлаас хамааралгүй байдгаараа давуу талтай. Lian et al. [6] толгойн байрлалыг оновчтой тодорхойлохын тулд гүн (depth) болон RGB онцлог цэгүүдийг хослуулан ашигласан. MediaPipe FaceMesh [7] гарч ирснээр нүдний солонгон бүрхүүлийн (iris) цэгүүдийг багтаасан 478 цэг бүхий нүүрний цэгүүдийг хэрэглэгчийн төхөөрөмж дээр бодит хугацаанд илрүүлэн, ашиглах боломжтой болсон.

D. Тохируулга (Calibration)

Тохируулга нь ерөнхийлөх зорилгоор сурсан загвар болон тухайн хэрэглэгчийн бие болон хэрэглэж буй төхөөрөмжийн онцлог хоорондын зөрүүг арилгахад чухал үүрэгтэй. Affine calibration, polynomial calibration болон Гауссын процесс дээр суурилсан арга нь өмнө нь судлагдаж байсан [9]. Affine calibration нь хамгийн энгийн арга бөгөөд хэмжээ (scale), шилжилт (offset) болон хазайлтын (skew) зөрүүг арилгахад тохиромжтой.

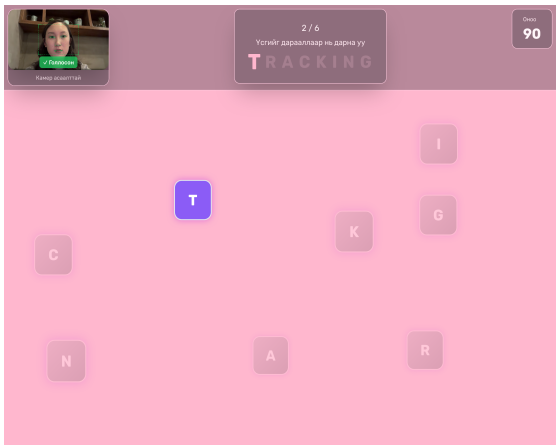
III. Өгөгдөл Цуглуулах

A. Өгөгдөл цуглуулах интерфейс

Харааны өгөгдлийг цуглуулахын тулд веб аппликейшн хөгжүүлсэн. Хэрэглэгч дэлгэц дээр санамсаргүй байрлалаар гарч ирэх үсгүүдийг дараалаар нь дарж, нийт зургаан үгийг бүтээх даалгавар гүйцэтгэнэ. Зургаан үг нь нийт 46 үсгээс бүтэх тул энэ тооны цэгийг оролцогчид дэлгэц дээр харж дарна. Ингэснээр хэрэглэгчийн хараа дэлгэц дээрх тодорхой координатуудад (x,y) төвлөрөх нөхцөлийг бүрдүүлнэ. Үсэг бүрийг дарах үйлдэлтэй уялдуулан дараах дарааллаар цуглуулсан.

- Үсэг дээр дарахаас өмнөх 5 фрэйм;
- Дарах агшны 1 фрэйм;
- Дарсны дараах 2 фрэйм.

Хэрэглэгчийн хараа тухайн цэг рүү шилжих болон тогтох хөдөлгөөнийг илүү оновчлох үүднээс үсэг даралт бүр дээр нийт 8 фрэйм бүхий дарааласан зургуудыг авсан.



1: Өгөгдөл цуглуулах веб аппликейшн

B. Цуглуулсан өгөгдөл

Өгөгдөл цуглуулахад 114 оролцогчийн сэшн өгөгдөл цугларсан. Өгөгдөл цэвэрлэгээний (Хэсэг III-C) дараа 96 сэшн үлдсэн ба сэшн бүр нь 46 фрэймээс гаргаж авсан нүүрний цэгэн мэдээллийг агуулах бөгөөд үүнээс нийт 4,416 (96 $\times$ 46) түүвэр өгөгдөл үүснэ. Цугларсан өгөгдөл нь 1280×720-оос 2240×1400 хүртэлх дэлгэцийн 18 өөр хэмжээсийг хамарсан. Дэлгэц дээрх цэгүүдийн байршил нь нормчлогдсон координатаар хэвтээ тэнхлэгт [0.04, 0.99], босоо тэнхлэгт [0.26, 0.99] мужийг хамарч байгаа бөгөөд энэ нь дэлгэцийн ашиглагдах боломжтой талбайн ихэнх хэсгийг хамарсан болно.

C. Өгөгдөл цэвэрлэгээ

Өгөгдлийг цэвэрлэх шатанд дараах алхмуудыг хэрэгжүүлсэн:

- **Алдаатай өгөгдлийг хасах:** Харааны координатууд $[0, 1]$ хэвийн хүрээнээс гадагш (олон дэлгэц ашигласан) 3 сэшн хассан.
- **Гар утасны сэшн:** Мобайл төхөөрөмж дээр бичигдсэн (дэлгэцийн өргөн $\leq$ 767 пиксель) сэшн хассан.
- **Царай илрүүлэх:** 8 фрэймийн 5-аас багад нь нүүр илэрсэн бол тухайн сэшнийг хасна. Энэ шалгуур нь чанаргүй, бүдэг эсвэл царай харагдахгүй зургуудаас үүдэлтэй шуугианыг бууруулна.

Сэшн бүрд 8 фрэйм бүгд нүүр илрүүлсэн нь өгөгдлийн чанар сайн байгааг харуулж байна.

IV. Шинж Чанар Ялгах

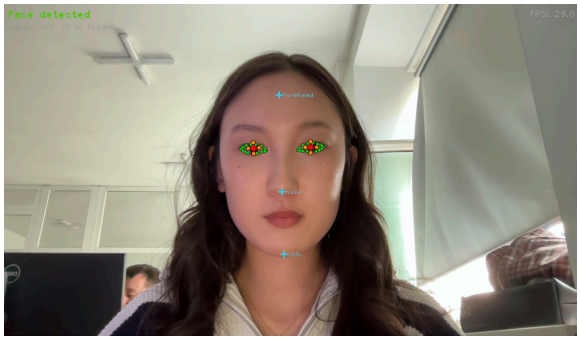
A. MediaPipe FaceMesh

MediaPipe FaceMesh [7]-ийг түүвэр өгөгдөл бүрийн 8 фрэйм дээр `refine_landmarks=True` тохиргоотой ажиллуулсан бөгөөд энэ нь царайны 468 цэг дээр нэмээд нүдний солонгон бүрхүүлийн 10 цэгийг илрүүлнэ. Эдгээр цэгүүд нь хүний нүүрний зургийн хэмжээтэй харьцангуй нормчлогдсон: $x \in [0, 1]$ (зургийн өргөний хувь), $y \in [0, 1]$ (зургийн өндрийн хувь), z нь цэгийн хамрын үзүүрт харьцуулсан гүнийг илэрхийлнэ, $z \in [-0.1, 0.1]$ завсарт байдаг.

B. Цэгэн шинж чанар сонголт

Шинж чанар болгох цэгүүдээр нүдний ойр орчмын цэгүүдийг сонгосон:

- **Нүдний хүрээ** (нэг нүдэнд 16 цэг): Нүдний булан, дээд болон доод зовхины цэгүүд. Эдгээр нь нүдний хөдөлгөөн дээшлэх, доошлох болон анивчих үеийн зовхины хэлбэрийг тодорхойлно.
- **Солонгон бүрхүүлийн тойрог** (нэг нүдэнд 4 цэг): Эдгээр цэгүүд нь солонгон бүрхүүлийн хэмжээ болон чиглэлийг илэрхийлнэ.
- **Солонгон бүрхүүлийн төв** (нэг нүдэнд 1 цэг): Харааны чиглэлийг тооцоолоход ашиглана.
- **Толгойн байрлалын цэгүүд** (3 цэг): Хамрын үзүүр, эрүү, дух. Эдгээр нь нүдний байрлалаас дэлгэцийн координат руу буулгахад нөлөөлдөг толгойн хазайлт болон эргэлтийг тодорхойлно.



2: Ашигласан нүүрний цэгүүд

C. Шинж чанарын вектор үүсгэх

Дэлгэц дээр дарагдсан цэгийн 8 фрэйм бүр дээр MediaPipe FaceMesh ажиллуулж, тэдгээрийн цэгүүдийг дундажлана. Дараа нь Хүснэгт I-д үзүүлсэнчлэн 141 хэмжээст шинж чанарын вектор үүсгэнэ.

Харааны чиглэлийн вектор:

$$\mathbf{d} = \frac{(\mathbf{c}_L - \bar{\mathbf{e}}_L) + (\mathbf{c}_R - \bar{\mathbf{e}}_R)}{2} \tag{1}$$

Энд $\mathbf{c}$ нь солонгон бүрхүүлийн төвийг, $\bar{\mathbf{e}}$ нь нүдний хүрээний цэгүүдийн дундаж байрлалыг илэрхийлнэ. Энэ вектор нь толгойн хазайлтаас үл хамааран зураг дээр нүдний алимын эргэлтийг тодорхойлно.

D. Цэгэн өгөгдлүүдийг ашиглан нэмэлтээр гаргаж авсан шинж чанарууд

Машин сургалтын загварт түүхий өгөгдөл (raw landmarks)-ийг шууд өгөхөөс илүүтэйгээр тэдгээрийг геометр болон физик утга бүхий шинж чанаруудад хөрвүүлэх нь загварын суралцах явцыг хурдасгаж, нарийвчлалыг нэмэгдүүлнэ гэж үзэж дараах шинж чанарыг гарган авсан.

1) **Нүдний солонгон бүрхүүлийн диаметр:** Энэ хэмжээг мэдсэнээр хүн дэлгэцээс хол эсвэл ойр сууж байгааг мэдэх боломжтой. Солонгон бүрхүүлийн диаметр их бол дэлгэцэд ойрхон, бага бол дэлгэцээс хол байна. Энэ утгыг олохдоо нэг нүдний солонгон бүрхүүлийн зүүн болон баруун цэгийн координатыг ашиглан олж, зургийн өргөнийг ашиглан нормчилсон. Нүдний диаметрийг олохдоо Euclidean зай олох аргыг ашигласан.

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

2) **Дэлгэц хүртэлх зайг тооцоолох:** Хүний нүдний солонгон бүрхүүлийн хэвтээ диаметр нь хүн амын дунд дунджаар 11.7±0.5 мм [10] тогтмол байдагт үндэслэн Pinhole Camera Model ($D = f \times W/P$)-ийг урьдчилан боловсруулах (preprocessing) шат болгон ашигласан. Хэрэглэгч бүрийн камерын үзүүлэлт (Field of View) өөр байдаг тул фокусын зайг (f) сэшн бүрд автоматаар тодорхойлох bootstrapping аргыг ашигласан. Хэрэглэгчийг дэлгэцээс дунджаар 500 мм зайд сууж байна гэж таамаглан, тухайн сэшний туршид бүртгэгдсэн солонгон бүрхүүлийн дундаж пиксель хэмжээгээр фокусын зайг дараах байдлаар тооцоолсон:

$$f_{\text{session}} = \frac{\overline{iris_{px}} \times 500}{11.7}$$

3) **Харааны чиглэлийн өнцгүүд (gaze_dir_h/v)** MediaPipe-аас ирж буй 3 хэмжээст векторыг (dx,dy,dz) хэвтээ болон босоо чиглэлийн өнцөг (градус) руу хөрвүүлсэн. Вектор утгууд нь дэлгэцийн координаттай шугаман бус хамааралтай байдаг бол өнцгийн утга нь дэлгэцийн байршилтай шууд хамааралтай. Энэ нь загварт X болон Y координатыг таамаглахад шаардлагатай мэдээллийг өгөх ач холбогдолтой үзүүлэлт болно.
4) **Толгойн эргэлтийн өнцгүүд** Хүний хараа нь зөвхөн нүдний хөдөлгөөн бус, толгойн хөдөлгөөнөөс хамаарна.
   - **Yaw (Хэвтээ эргэлт):** Хамар болон духны хоорондох векторын хэвтээ өөрчлөлтөөр тооцоолно.
   - **Pitch (Босоо хазайлт):** Хамар болон эрүүний хоорондох векторын босоо өөрчлөлтөөр тооцоолно.

   Толгой болон нүдний хөдөлгөөний хамаарлыг загварт өгнө. Хэрэглэгч толгойгоо эргүүлж харж байна уу, эсвэл зөвхөн нүдээ хөдөлгөж байна уу гэдгийг таамаглахад тусална.
5) **Хоёр хүүхэн хараа хоорондын зай:** Хоёр нүдний төв цэг хоорондын пиксель зайг тооцоолж, зургийн өргөнд харьцуулан нормчилсон. Энэ нь солонгон бүрхүүлийн хэмжээнээс үл хамаарах гүний (depth) хоёрдогч үзүүлэлт болно. Мөн толгой эргэх үед проекцийн улмаас хоёр нүд хоорондын зай багасаж харагдах тул зай болон толгойн эргэлтийн (yaw) давхар баталгаажуулалт болж өгнө.

E. Шинж чанар стандартчилах

Өгөгдлийн урьдчилсан боловсруулалтын нэг хэсэг болгон StandardScaler (Z-score normalization) ашиглан шинж чанар бүрийг дундаж нь 0, стандарт хазайлт нь 1 байхаар хувиргасан. Өгөгдөл нь 3 бүлэг өөр хэмжээс, нэгж бүхий (нүдний цэгэн өгөгдөл, дэлгэцийн хэмжээ, геометрийн утгууд) шинж чанаруудаас бүрдэж байгаа тул сургалтын явцад аль нэг үзүүлэлт хэт давамгайлах эсвэл загвар удаан суралцах эрсдэлийг бий болгоно гэж үзээд тус бүр ялгаатай **scaler** ашигласан.

F. Сургалт/Баталгаажуулалт/Туршилт өгөгдөл хуваалт

Загварын алдаар эцсийн байдлаар үнэлэхдээ k-fold cross validation (k=10) аргыг ашигласан. Энэ нь цуглуулсан өгөгдлийн хэмжээ хязгаарлагдмал нөхцөлд загварыг өгөгдлийн бүх хэсэг дээр давтан шалгаж, хэт сурах (overfitting) эрсдэлийг багасгахын зэрэгцээ ерөнхийлөх чадварыг (generalization) илүү бодитой тодорхойлох боломжийг олгоно гэж үзсэн. Харин загварын

I: Шинж чанарын вектор бүтэц

Индекс	Агуулга	Хэмжээ	Scaler
0–47	Зүүн нүдний хүрээ (16 × xyz)	48	eye
48–95	Баруун нүдний тоо (16 × xyz)	48	eye
96–107	Зүүн солонгон бүрхүүлийн тойрог (4 × xyz)	12	eye
108–119	Баруун солонгон бүрхүүлийн тойрог (4 × xyz)	12	eye
120–122	Зүүн солонгон бүрхүүлийн төв (xyz)	3	eye
123–125	Баруун солонгон бүрхүүлийн төв (xyz)	3	eye
126–128	Харааны чиглэл (dx, dy, dz)	3	eye
129–137	Нүүрний байрлалын цэгүүд	9	eye
138–140	Дэлгэцийн өргөн, өндөр, DPR	3	screen
141	солонгон бүрхүүлийн хэмжээ (px)	1	geom
142	Дэлгэц хүртэлх зай (mm)	1	geom
143	Харааны хэвтээ чиглэлийн өнцөг (degree)	1	geom
144	Харааны босоо чиглэлийн өнцөг (degree)	1	geom
145	Толгойн хэвтээ эргэлт (degree)	1	geom
146	Толгойн босоо хазайлт (degree)	1	geom
147	Хүүхэн хараа хоорондын зай (px)	1	geom
	Нийт	**148**	

чадварыг дүрслэн харуулах зорилгоор өгөгдлийг сургалт/баталгаажуулалт/туршилт гэсэн уламжлалт бүтцээр хуваасан. Ингэхдээ өгөгдлийг мөрөөр биш сэшнээр хуваасан. Хэрэв мөрөөр хуваавал нэг хүний өгөгдөл сургалтын болон турших өгөгдөлд хоёуланд нь орж, загвар тухайн хүний онцлогийг ”цээжилсэн” мэт өндөр үр дүн үзүүлэх боловч бодит хэрэглээнд муу ажиллана (data leakage). Өгөгдлийг дараах байдлаар хуваасан:

$$96 \text{ сэшн} \xrightarrow{\text{random shuffle (seed=42)}} \begin{cases} \text{Сургалт: 72 сэшн} \\ \text{Баталгаажуулалт: 5 сэшн} \\ \text{Туршилт: 5 сэшн} \end{cases}$$

V. Сургалт

A. Ridge Regression

Энэхүү судалгаанд Ridge Regression шугаман загварыг суурь болгон ашигласан. Тус загвар нь дараах функцийн утгыг хамгийн бага байлгах замаар шугаман хамаарлыг тодорхойлдог:

$$\mathcal{L}(\mathbf{w}) = \sum_{i=1}^{N} (y_i - \mathbf{w}^T \mathbf{x}_i)^2 + \alpha \|\mathbf{w}\|^2$$

Энд α=1.0. Харах чиглэлийг тодорхойлохын тулд дэлгэцийн тэнхлэг бүрд (хэвтээ ба босоо) харгалзах хоёр тусдаа регрессорыг сургасан. Тус загварт ашиглагдаж буй L2 нь аль нэг шинж чанарт хэт өндөр жин (weight) оноохоос сэргийлдэг бөгөөд ингэснээр загварын ерөнхийлөн дүгнэх чадварыг сайжруулж, тооцооллын хувьд хялбар ”closed-form” шийдлийг гаргадаг давуу талтай.

B. Random Forest

Судалгаанд 300 шийдвэрийн модноос (decision trees) бүрдэх Random Forest загварыг ашигласан бөгөөд үүнийг олон гаралттай (multi-output) горимоор сургасан. Загварын гипер-параметрүүдийг тохируулахдаа навчны хамгийн бага түүвэрлэлтийг min_samples_leaf=2, модны гүнийг хязгааргүй (max_depth=None) байхаар сонгосон. Мод бүрийг сургалтын өгөгдлийн ”bootstrap” түүвэр дээр үндэслэн үүсгэхдээ салаа (split) бүрд шинж чанаруудын санамсаргүй дэд олонлогийг (random feature subset) сонгон авсан. Эцсийн таамаглал нь бүх 300 модны дунджаар тодорхойлогдоно. Энэ загвар нь өгөгдлийн өөр өөр дэд хэсгүүд дээр ”хэт сурсан” (overfit) олон төрлийн моднуудын үр дүнг нэгтгэн дундажлах замаар системийн нийт хазайлтыг бууруулж, загварын тогтвортой байдлыг хангах ач холбогдолтой.

C. Extreme Gradient Boosting (XGBoost)

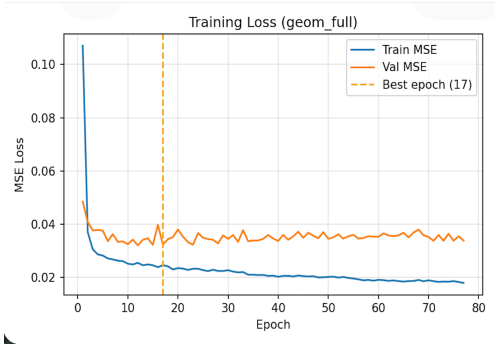
Судалгаанд тэнхлэг бүрд (хэвтээ ба босоо) харгалзах хоёр тусдаа XGBRegressor загварыг сургасан бөгөөд сургалтын давталтыг (n_estimators) 300, сургалтын хурдыг (learning_rate) 0.05 байхаар тус тус тогтоосон.

Gradient Boosting арга нь шийдвэрийн модуудыг дараалсан хэлбэрээр үүсгэдэг бөгөөд мод бүр нь өмнөх ансамблийн (ensemble) үлдэгдэл алдааг (residuals) засах замаар сурдаг. Түүнчлэн, загварын ерөнхий үзүүлэлтийг сайжруулах болон хэт сурахаас (overfitting) сэргийлэх зорилгоор max_depth=6, subsample=0.8, colsample_bytree=0.8 параметрүүдийг тохируулж, L1 (reg_alpha=0.1) болон L2 (reg_lambda=1.0) зэрэг аргуудыг ашигласан.

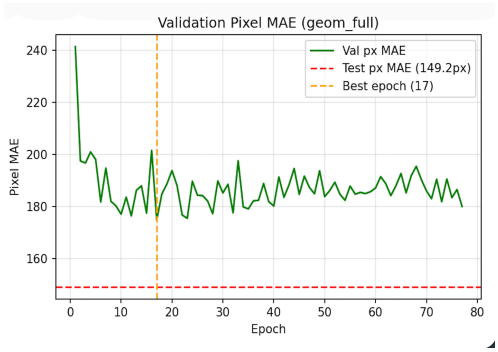
D. Multilayer Perceptron (Гүн сургалтын загвар)

Нийт 3 ялгаатай хэмжээстэй шинж чанар байгаа тул 3 өөр салаа болгон сургаж, сүүлд нь нэгтгэсэн. Загварын үр дүнг сайжруулах, байгаа өгөгдлийг хамгийн үр ашигтайгаар ашиглах зорилгоор хоёр үе шаттай сургалтын хийсэн.

1) **1-р шат:** Параметр тохируулга ба Эрт зогсолт (Optimization & Early Stopping)
   Энэ шатанд нийт өгөгдлийг сургалтын (Train=3956) болон баталгаажуулах (Val=230) хэсэгт хуваж сургалтыг явуулсан. Үүний зорилго нь загвар хэт сурахаас сэргийлэх, хамгийн бага алдаатай үеийг (best_epoch) тодорхойлох. Баталгаажуулах хэсгийн алдаа 60 үе (epochs) дараалан сайжрахгүй бол сургалтыг зогсоож, хамгийн сайн үзүүлэлттэй үеийн жинг хадгална.
2) **2-р шат:** Нэгтгэсэн сургалт
   Загварын эцсийн хувилбарыг гаргахдаа баталгаажуулах хэсэгт ашигласан өгөгдлийг сургалтад нэмж оруулсан. Зорилго нь загварыг илүү олон өгөгдлөөр сургаж, таамаглах чадварыг нэмэгдүүлэх.
   Загварын жинг дахин шинээр эхлүүлж, сургалт болон баталгаажуулах өгөгдлийг нэгтгэн (нийт=4186 мөр) сургасан.
   Энэ шатанд эрт зогсолт ашиглахгүй бөгөөд 1-р шатанд тодорхойлогдсон best_epoch сургалтыг явуулна.



3: MLP loss curve



4: MLP Validation Pixel MAE

VI. Туршилт ба Үр Дүн

A. Үнэлгээний хэмжүүр

Бүх загварын үр дүнг MAE (Mean Absolute Error) хэмжүүрээр үнэлсэн:

$$MAE_{px} = \frac{1}{N}\sum_{i=1}^{N}\frac{|\hat{x}_i - x_i| \cdot W + |\hat{y}_i - y_i| \cdot H}{2}$$

Энд W нь дэлгэцийн өргөн, H нь өндөр, $\hat{x}, \hat{y}$ нь таамагласан координат, x, y нь бодит координат юм. Загвар нормчилсон (0–1) координат дээр сурч, үр дүнг пиксельд хөрвүүлэн харьцуулсан.

B. Загвар харьцуулалт

Сургалтад нийт 93 сешн өгөгдлийг ашигласан бөгөөд загварын таамаглах чадварыг бодитойгоор шалгахын тулд 3 сэшнийг сургалтын өгөгдлөөс салгаж, сургалтын үйл явцыг $k = 10$ хуваалтаар гүйцэтгэв.

Хүснэгт II-д Ridge Regression, Random Forest, XGBoost болон MLP загваруудын гүйцэтгэлийг дундаж үнэмлэхүй алдаа (MAE)-аар харьцуулан харуулав.

C. Дундаж пиксель алдаа

Загварыг үнэлэхэд гарсан дундаж алдаа (MAE) нь хэрэглэгч бүрийн ашиглаж буй дэлгэцийн физик хэмжээ болон нягтралаас хамааран харилцан адилгүй утгыг илэрхийлнэ.

- **Харьцангуй алдаа:** Пиксель алдаа нь дэлгэцийн нягтрал (resolution) өндөр байх тусам физик зайны хувьд багасна. Жишээ нь, 4K дэлгэц дээрх 141 пиксель нь Full HD дэлгэцтэй харьцуулахад илүү бага талбайг эзэлнэ.
- **Төхөөрөмжийн дасан зохицох чадвар (Cross-device Compatibility):** Туршилтад олон төрлийн дэлгэц (13–27 инч) хамруулсан нь загвар төхөөрөмжийн ялгаатай байдлаас үл хамааран тогтвортой ажиллах боломжтойг харуулж байна.

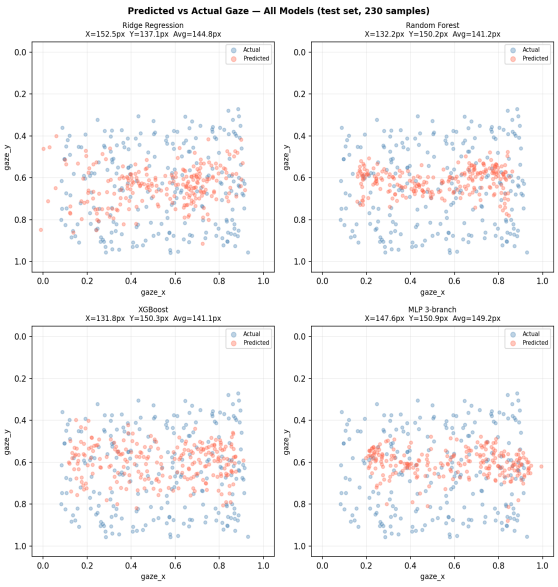
II: Загваруудын гүйцэтгэлийн харьцуулалт (k-fold cross-validation, k=10)

Загвар	Avg. MAE (px)	X MAE (px)	Y MAE (px)
Ridge Regression	146.1	153.3	138.9
Random Forest	143.2	134.4	151.9
XGBoost	**141.3**	135.7	146.8
MLP	141.5	130.8	152.1

Пиксель алдаа нь хэрэглэгчийн дэлгэцийн үзүүлэлтээс хамааран өөр өөр байх боловч одоогоор тооцоолсон дундаж утга нь системийн ерөнхий нарийвчлалыг харуулна.

D. Алдааны тархалтын шинжилгээ

Хүснэгт II нь 10-fold cross-validation аргыг ашиглан загваруудын тогтвортой байдал болон дундаж гүйцэтгэлийг үнэлсэн болно. Үүнээс гадна, загваруудын бодит өгөгдөл дээрх таамаглалын алдааны (error behavior) болон хазайлтыг харах зорилгоор 86/5/5 (train/val/test) харьцаатай сургалт дээрх үр дүнг доорх зурагт дүрслэв.



5: Scatter Plots

VII. Бодит Туршилт

Хүн бүрийн нүүрний анатоми, нүдний онцлог, орчны гэрэлтүүлэг, камерын өмнө суусан байрлал, мөн төхөөрөмжийн камерын байрлал зэрэг ялгаатай тул судалгаанд оролцогчдын өгөгдөл дээр сургасан загвар бодит туршилт

дээр зөрөх, алдаа гаргах магадлалтай. Иймд загварыг бодит орчинд ажиллуулахын өмнө тохируулга (calibration) хэсэг нэмж ажиллуулснаар энэ алдааг засах боломжтой бөгөөд тохируулгыг хэрэгжүүлэхдээ нүд дагах загвараар MLP загварыг ашигласан болно.

A. Тохируулга ашиглан загварын түүхий таамаглалын алдааг багасгах нь

Энэхүү процесс нь тухайн хэрэглэгчийн нүдний онцлог болон бусад ялгаатай хүчин зүйлд зориулан машин сургалтын загварын таамаглалын алдааг бууруулах хувиргалтын процесс юм.

1) **Тохируулгын 16 цэг:**
   Дэлгэцийн талбайг бүхэлд нь хамарсан 4×4 хэмжээтэй цэгүүдийг ашигласан. Цэгүүдийг дэлгэцийн нормчилсон координат дээр дараах байдлаар байрлуулсан:

   - **X тэнхлэг:** 0.08, 0.36, 0.64, 0.92
   - **Y тэнхлэг:** 0.12, 0.40, 0.67, 0.93

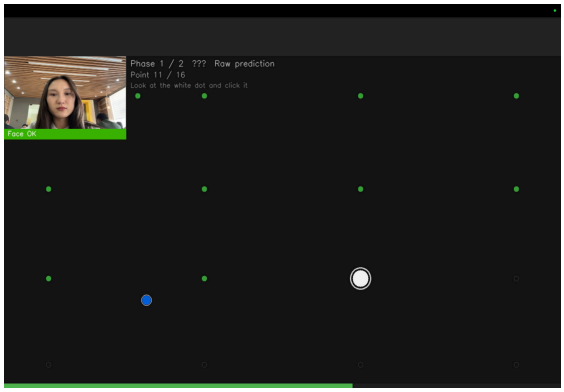
   Зүүн дээд өнцгийн цэгийг (0.08, 0.12) веб камерын цонхтой давхцуулахгүй байх үүднээс (0.24, 0.12) координат руу шилжүүлж байрлуулсан.
2) **Загварыг үнэлэх**

   - **Түүхий алдааг хэмжих (Raw Error):** 16 цэг дэлгэц дээр дарааллаар гарч ирэх бөгөөд хэрэглэгч цэг бүр дээр дарна. Цэг дарах бүрд тухайн цэг дээрх загварын 10 фрэймийн таамаглалыг дундажлаж, түүхий алдааны MAE тооцно.
   - **Аффин хувиргалт:** Өмнөх 16 хос цэгийг ашиглан хамгийн бага квадратын аргаар (Least Squares) 2×3 хэмжээтэй аффин матрицыг тодорхойлно. Энэхүү матриц нь загвараас гарч буй таамаглалын зөрүү (scale), шилжилт (translation) болон хазайлтыг (skew) засахад ашиглагдана.

$$\begin{bmatrix} x_{true} \\ y_{true} \end{bmatrix} = \mathbf{A} \times \begin{bmatrix} x_{pred} \\ y_{pred} \\ 1 \end{bmatrix} \tag{2}$$

   - **Зассан алдааг хэмжих (Calibrated Error):** 16 цэгийг дахин харуулж, энэ удаад загварын таамаглалыг аффин матрицаар зассаны дараа алдааг дахин хэмжинэ.

B. Бодит хугацаанд алдааг үнэлэх



6: Загварыг алдааг үнэлэх интерфейс

Туршилтыг 1280x832, device pixel ratio=2.0 төхөөрөмж дээр хийсэн болно.

III: Calibration үр дүн

Үзүүлэлт	MAE X (px)	MAE Y (px)	MAE Avg (px)
Тохируулгын өмнө	93.9 (7.3%)	231.1 (27.8%)	162.5 (15.4%)
Тохируулгын дараа	**64.8 (5.1%)**	**121.9 (14.6%)**	**93.4 (8.8%)**
Сайжруулалт	**31%**	**47%**	**43%**

- **Y тэнхлэгийг хазайлт:** Аффин хувиргалт хийснээр Y тэнхлэгийн алдаа хамгийн их буюу 47%-аар буурсан. Энэ нь аффин хувиргалт шилжилтийн алдааг (translation bias) болон тогтмол алдааг (systematic bias) засдагийг харуулж байна.
- **Нарийвчлал:** Нийт дундаж алдаа 162.5 px-ээс 93.4 px болж, бараг хоёр дахин багассан.
- **Тохируулгын ач холбогдол:** Туршилтаар хэрэглэгч бүрийн онцлогт тохирсон тохируулга хийх нь загварын нарийвчлалыг нэмэгдүүлэх гол хүчин зүйл болох нь харагдаж байна.

10-fold туршилтын үед Y тэнхлэгийн алдаа бага гарсан ч бодит цагийн туршилтад их байгаа нь загвар сургалтын болон баталгаажуулах өгөгдөлд байгаа адилхан системийн хэлтийлт (systematic bias) сурсантай холбоотой гэж үзэж байна. Энэ нь бидний ашигласан аффин хувиргалтын арга яагаад Y тэнхлэг дээр 47% сайжирсныг тайлбарлаж байна. Мөн X тэнхлэгийн алдаа нь Y тэнхлэгээс харьцангуй бага байгаа нь хэвтээ тэнхлэг нь илүү жигд, нүдний цэгүүдийн хөдөлгөөнтэй харьцангуй шууд хамааралтай байгаатай холбоотой гэж үзэж байна.

VIII. Дүгнэлт

Энэхүү судалгааны ажлаар MediaPipe FaceMesh болон машин сургалтын аргуудыг ашиглан энгийн веб камераар бага өртөгтэй, хүртээмжтэй харааны чиглэл тодорхойлох загваруудыг сургаж, тэдгээрийг үнэлэв.
II хүснэгтээс загваруудын алдаа 141-146 px MAE хооронд буюу харьцангуй ижил үр дүн үзүүлсэн нь харагдаж байна. Эндээс загваруудын үр дүн тэдгээрийн архитектураас

бус сургалтын өгөгдөлд агуулагдах системийн хэлтийлт (systematic bias) болон шуугиантай холбоотой гэж үзэж байна. Ялангуяа Ridge Regression шугаман загварын үзүүлэлт бусад загваруудтай ойролцоо байгаа нь үүнийг баталж байна. Мөн MLP загварын баталгаажуулалтын алдаа тогтвортой байсан ч сургалтын алдаа нь буурсаар байгаа нь загвар өгөгдлийн шуугиан дээр хэт сурч байгааг илтгэж байна. Гэсэн хэдий ч XGBoost, MLP загварууд нь шугаман бус хамаарлыг илүү үр дүнтэй сурснаар бага зэрэг давуу талтай байгаа нь ажиглагдав. Эцэст нь, харааны чиглэл тодорхойлох загваруудын хувьд алгоритмын сонголтоос илүүтэйгээр өгөгдлийн чанарыг сайжруулах нь илүү чухал ач холбогдолтой гэж дүгнэж байна.

IX. Талархал

Энэхүү судалгааны ажлын хүрээнд ашиглагдсан өгөгдөл цуглуулахад тусалсан бүх багш, оюутнууддаа талархал илэрхийлье.

X. Ном Зүй

[1] A. T. Duchowski, *Eye Tracking Methodology: Theory and Practice*, 2nd ed. Springer, 2007.
[2] Y. Sugano, Y. Matsushita, and Y. Sato, “Learning-by-synthesis for appearance-based 3D gaze estimation,” in *Proc. IEEE CVPR*, 2014, pp. 1821–1828.
[3] X. Zhang, Y. Sugano, M. Fritz, and A. Bulling, “Appearance-based gaze estimation in the wild,” in *Proc. IEEE CVPR*, 2015, pp. 4511–4520.
[4] E. Wood *et al.*, “Learning an appearance-based gaze estimator from one million synthesised images,” in *Proc. ACM ETRA*, 2016, pp. 131–138.
[5] K. Krafka *et al.*, “Eye tracking for everyone,” in *Proc. IEEE CVPR*, 2016, pp. 2176–2184.
[6] D. Lian, L. Li, J. Liang, L. Liu, C. Hu, and S. Gao, “RGBD-based real-time gaze detection system,” in *Proc. IEEE FG*, 2019.
[7] C. Lugaresi *et al.*, “MediaPipe: A framework for building perception pipelines,” *arXiv preprint arXiv:1906.08172*, 2019.
[8] A. Papoutsaki, P. Sangkloy, J. Laskey, N. Daskalova, J. Huang, and J. Huang, “WebGazer: Scalable webcam eye tracking using user interactions,” in *Proc. IJCAI*, 2016, pp. 3839–3845.
[9] D. W. Hansen and Q. Ji, “In the eye of the beholder: A survey of models for eyes and gaze,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 3, pp. 478–500, 2010.
[10] J. P. G. Bergmanson and J. G. Martinez, “Size does matter: what is the corneo-limbal diameter?” *Clinical and Experimental Optometry*, vol. 100, no. 5, pp. 522–528, 2017.

Гүн сургалтад суурилсан цээжний рентген зургаас эмнэлзүйн тайлан үүсгэх аргын судалгаа

Мэндбаярын Энхболор
Мэдээллийн технологи
ШУТИС, МХТС
Монгол улс, Улаанбаатар
B221890016@must.edu.mn

Бямбадоржийн Золзаяа
Мэдээллийн технологи
ШУТИС, МХТС
Монгол улс, Улаанбаатар
b.zolzaya@must.edu.mn

***Хураангуй*— Цээжний рентген зураг нь уушги, зүрх судасны өвчнийг оношлоход өргөн хэрэглэгддэг дүрс оношилгооны нэг бөгөөд эмч мэргэжилтнүүд цээжний рентген зургийг үзэн шалгаж, ажиглалтын дагуу тайлан бичих үйл явц нь цаг их шаарддаг, нарийн ажил юм. Энэхүү судалгааны зорилго нь хиймэл оюун ухаанд суурилсан цээжний рентген зургаас оношилгооны тайлан үүсгэх систем боловсруулан эмчийн ажлыг хөнгөвчлөх, эмч нарын нөөц бага, ачаалал ихтэй эмнэлгүүдэд дэмжлэг үзүүлэх юм. Судалгааны ажлыг дараах 3 үе шаттайгаар гүйцэтгэсэн. (1) Indiana University (IU) Chest X-ray өгөгдлийн багц дээр дөрвөн харааны энкодер загвар (Swin, ViT, DEiT, BEiT) болон хоёр хэлний декодер загвар (DistilGPT, BioGPT)-ийн нийт найман хослолыг сурган гүйцэтгэлийг үгийн түвшний хэмжүүр (BLEU, ROUGE, METEOR), эмнэлзүйн хэмжүүр (CheXbert F1, RadGraph F1)-ээр үнэлсэн. Туршилтын үр дүнд Swin+BioGPT болон DEiT+BioGPT хослолууд хамгийн сайн гүйцэтгэл үзүүлсэн. (2) Indiana University (IU) Chest X-ray өгөгдөл дээр MIMIC-CXR өгөгдлийг алхамчилсан байдлаар (+2000 хүртэлх зураг, +4000 хүртэлх зураг) нэмж Swin+BioGPT болон DEiT+BioGPT загваруудыг сурган, гүйцэтгэлийг үнэлэхэд үгийн түвшний хэмжүүрүүдэд мэдэгдэхүйц өөрчлөлт гараагүй ч эмнэлзүйн хэмжүүр 74%-аас 80% болж өссөн. Мөн MIMIC-CXR өгөгдлийг цэвэрлэн рентген зурагт ажиглагдаж буй олдворуудтай хамааралгүй өгүүлбэрүүдийг устгахад эмнэлзүйн хэмжүүр 82% болж сайжирсан. (3) Indiana University (IU) Chest X-ray + MIMIC CXR нийлмэл өгөгдөл дээр on-the-fly data augmentation хийж загваруудыг сурган гүйцэтгэлийг үнэлэхэд үгийн түвшний хэмжүүр бага хэмжээгээр сайжирсан бөгөөд эмнэлзүйн хэмжүүр тогтвортой хадгалагдсан. Үүний үр дүнд үгийн түвшний болон эмнэлзүйн хэмжүүр хоёулаа сайжирснаар загвар нь найруулгын хувьд чанартай бөгөөд эмнэлзүйн хувьд илүү үнэн зөв тайлан үүсгэх чадвартай болсон.**

Түлхүүр үг*— *Vision Encoder, Language Decoder, Swin, ViT, DEiT, BEiT, DistilGPT, BioGPT

I. Удиртгал

Цээжний рентген зураг нь уушгины үрэвсэл, сүрьеэ, зүрхний томрол, уушгины хавдар зэрэг олон төрлийн эмгэгийг илрүүлэхэд өргөн хэрэглэгддэг дүрс оношилгооны арга юм. Дэлхий даяар жилд ойролцоогоор 830 сая гаруй удаа цээжний рентген зургийн оношилгоо хийгддэг гэж Хойд Америкийн Радиологийн Нийгэмлэг (RSNA) мэдээлсэн [1].

Монгол Улсын хувьд сүрьеэгийн өндөр өвчлөлтэй орны нэг хэвээр байгаа бөгөөд 2024 оны судалгаагаар 100,000 хүн тутамд 446 шинэ сүрьеэгийн тохиолдол бүртгэгдсэн [2]. Энэ нь цээжний рентген зургийн оношилгооны хэрэгцээ шаардлага их байх нөхцөлийг бүрдүүлж байна. Түүнчлэн цээжний рентген зургийг үнэлж, оношилгооны тайлан бичих нь нарийвчилсан ажиглалт болон цаг хугацаа их шаарддаг үйл явц бөгөөд эмч нарын ачаалал нэмэгдэх үед тайлан гаргах хугацаа улам удаашрах асуудал тулгарч байна.

Нөгөө талаас дүрс оношилгооны эмч нарын хүний нөөцийн хомсдол нь бодит сорилт хэвээр байсаар байна. 2025 оны судалгаагаар Монгол улсад нийт 670 дүрс оношилгооны эмч ажиллаж байгаа бөгөөд 100,000 хүн тутамд 19.1 дүрс оношилгооны эмч ногдож байгаа нь эмч нарын ажлын ачаалал ихсэх, улмаар оношилгооны тайлан гаргах хугацаа удаашрах шалтгаан болж байна [3].

Иймд оношилгооны тайлан гаргах процессыг автоматжуулах, эмч нарт туслах систем боловсруулах шаардлага тулгарч байна.

Сүүлийн жилүүдэд гүн сургалтын арга, ялангуяа харааны трансформер (Vision Transformer) болон том хэлний загварууд (Large Language Model) хөгжсөнөөр зурган өгөгдлөөс нарийн орон зайн хэв шинжийг илрүүлэх болон түүнээс уялдаа холбоотой, контекстэд тохирсон текст үүсгэх өндөр чадамжтай болсон. Энэ технологи нь эмнэлгийн дүрсээс автоматаар оношилгооны тайлан үүсгэх боломжийг нээж байна.

Гэсэн хэдий ч зургаас текст үүсгэх технологийг цээжний рентген зурган өгөгдлөөс тайлан үүсгэх чиглэлд ашигласан судалгаанууд нь харьцангуй сүүлд эхэлсэн. Харин түүнээс өмнө нь хийгдсэн ихэнх судалгаанууд нь рентген зургаас “өвчтэй/өвчингүй” гэдгийг ангилах ангиллын түвшинд төвлөрсөн байдаг бол бодит практикт илүү их цаг шаарддаг хэсэг болох оношилгооны тайлан бичих процессыг шийдэж чаддаггүй.

Энэхүү судалгааны зорилго нь цээжний рентген зургаас эмнэлзүйн хувьд үнэн зөв, найруулгын хувьд чанартай оношилгооны тайлан үүсгэх хиймэл оюун ухаанд суурилсан систем боловсруулж, эмч нарын ажлыг хөнгөвчилж тэдэнд туслах, эмчийн нөөц бага эмнэлгүүдэд дэмжлэг үзүүлэхэд оршино.

Судалгааны шинэлэг тал нь цээжний рентген зургаас оношилгооны тайлан үүсгэх загварын гүйцэтгэлийг зөвхөн текстийн төстэй байдлын хэмжүүрээр бус, эмнэлзүйн утга агуулгын үнэн зөв байдлаар давхар үнэлсэнд оршино. Мөн орчин үеийн харааны трансформер болон том хэлний загваруудыг хослуулсан олон encoder-decoder архитектуруудыг нэг ижил нөхцөлд системтэйгээр харьцуулан судалсан. Түүнчлэн IU Chest X-ray болон MIMIC-CXR өгөгдлүүдийг нэгтгэн ашиглаж загварын ерөнхийлөх чадварыг сайжруулснаараа давуу талтай.

II. Ижил төстэй ажлуудын судалгаа

Оношилгооны тайлан үүсгэхэд хамгийн өргөн ашиглагддаг гүн сургалтад суурилсан арга нь энкодер-декодер фреймворк юм.

Цээжний рентген зургаас тайлан үүсгэх эхэн үеийн судалгаануудад ашиглагдсан арга нь Convolutional Neural Network (CNN) загварыг дүрсний онцлог шинж чанарыг гаргагч буюу энкодерт, Recurrent Neural Network (RNN) болон Long Short-Term Memory (LSTM) загваруудыг текст үүсгэгчид буюу декодерт ашиглах нь элбэг байсан. Жишээлбэл, Xue нар [4] нь цээжний рентген зургаас тайлан үүсгэх даалгаварт CNN-ийг дүрсний онцлог гаргагч болгон ашиглаж, RNN-д суурилсан текст үүсгэгчтэй attention механизм хослуулсан загвар санал болгосон.

Сүүлийн үеийн судалгаанд Transformer архитектурт суурилсан загваруудыг оношилгооны тайлан автоматаар үүсгэх чиглэлд өргөн хэрэглэх болсон. Тухайлбал, Raminedi нар [5] харааны болон хэлний трансформеруудыг нэгтгэсэн архитектур санал болгосон. Уг судалгаанд ViT болон түүний хувилбарууд болох BEiT, DEiT загваруудыг энкодерт, GPT2 загварыг декодерт ашиглан тэдгээрийн хослол архитектуруудын гүйцэтгэлийг харьцуулсан. Гүйцэтгэлийг BLEU, ROUGE зэрэг үгийн түвшний хэмжүүрээр үнэлж, өмнөх CNN-RNN суурь аргуудаас илүү гүйцэтгэл үзүүлснийг харуулсан.

Prateek нар [6] Swin Transformer-ийг харааны онцлог гаргагч болгон, DistilGPT-ийг текст үүсгэгч болгон хослуулсан sequence-to-sequence multimodal архитектур санал болгосон. Загварын гүйцэтгэл нь үгийн түвшний хэмжүүрүүдээр CNN-RNN аргуудаас илүү үзүүлэлттэй байсан. Гэсэн хэдий ч уг судалгаа нь нэг эх үүсвэрээс бүрдсэн IU Chest X-ray өгөгдлийн багцад тулгуурласан нь загварын ерөнхийлөх чадварт тодорхой хязгаарлалт үүсгэх боломжтой буюу зарим өвчлөлийн хэлбэрүүдийг дутуу төлөөлөх магадлалтай гэдгийг дурдсан.

Дээрх 3 судалгаа нь загварын гүйцэтгэлийг үгийн түвшний хэмжүүрүүдээр үнэлсэн бөгөөд эмнэлзүйн үнэн зөв байдлыг илэрхийлэх CheXbert F1, RadGraph F1 зэрэг хэмжүүрүүдийг онцолж авч үзээгүй. Мөн зөвхөн нэг төрлийн өгөгдлийн багц дээр загваруудыг сургасан тул ерөнхийлөх чадвар сул байж болзошгүй байна.

Эдгээр судалгааны ажлуудыг үндэслэн, бид цээжний рентген зургийн тайлан үүсгэх олон энкодер-декодер хослолыг IU Chest X-ray болон MIMIC-CXR өгөгдлийн багцуудын нийлмэл өгөгдлөөр сурган харьцуулан судалж, гүйцэтгэлийг үгийн түвшний болон эмнэлзүйн хэмжүүрүүдээр цогцоор нь үнэлнэ.

III. АШИГЛАСАН АРГА

A. Өгөгдлийн багц

Бид судалгааны ажилдаа дараах өгөгдлийн багцуудыг ашигласан болно. Үүнд:

1. IU Chest X-ray өгөгдлийн багц

Индиана Их Сургуулийн эмнэлгийн сүлжээний Open-I цуглуулгад хамаарах IU Chest X-ray нь оношилгооны тайлан үүсгэх судалгаанд хамгийн өргөн ашиглагддаг өгөгдлийн багц юм [7]. Энэхүү өгөгдлийн багц нь 7466 цээжний рентген зураг болон түүнд харгалзах 3826 дүрс оношилгооны тайланг агуулдаг. Өгөгдлийн багц дахь тайлан бүр нь урдаас болон хажуугаас (PA, lateral view) буюу 2 өнцгөөс авсан зурагтай. Нэг тайланд харгалзах рентген зургуудын тоо 1-ээс 5 хүртэл хэлбэлздэг. Нэг зурагтай тайлан 11.6%, хоёр зурагтай тайлан 83.3%, гурван зурагтай 4.7%, дөрвөн зурагтай 0.39%, таван зурагтай тайлан 0.03%-ийг тус тус эзэлж байна.

2. MIMIC-CXR өгөгдлийн багц

Medical Information Mart for Intensive Care-Chest X-ray (MIMIC-CXR) [8] өгөгдөл нь нийт 200 мянга гаруй тайлан, 370 мянга гаруй зурагтай Beth Israel Deaconess Medical Center эмнэлгийн өгөгдөл юм.

Бид судалгаандаа MIMIC-CXR өгөгдлийн 3636 тайлан, түүнд харгалзах 3971 зургийг ашигласан. PhysioNet сайтаас өгөгдлийн багцыг татахдаа 1 зурагтай тайлан 90%, 2 зурагтай тайлан 10% байхаар шүүн сонгосон. Учир нь тооцооллын нөөцийн хязгаарлалтаас үүдэн олон зураг, цөөн тайлантай өгөгдлийг ашиглахаас илүүтэйгээр өвчний олон тохиолдлуудыг багтаасан олон тайлан, цөөн зураг бүхий өгөгдөл ашиглахыг зорьсон. Өгөгдлийн бүтцийг хүснэгт 1-д харуулав.

ХҮСНЭГТ 1. ӨГӨГДЛИЙН БҮТЭЦ

Багана	**Тайлбар**
Indication	Шинжилгээ хийх болсон эмнэлзүйн шалтгаан, өвчтөний шинж тэмдэг
Technique	Зургийг аль өнцгөөс авсан тухай мэдээлэл
Impression	Оношилгооны дүгнэлт
Comparison	Өмнөх оношилгоотой харьцуулж одоогийн тайланг бичсэн эсэх
Findings	Рентген зургийн ажиглалт, анатомийн бүтэц болон эмгэг өөрчлөлтийн дэлгэрэнгүй тайлбар

B. Өгөгдлийн урьдчилсан боловсруулалт

1. Зурган өгөгдлийн урьдчилсан боловсруулалт

- Бүх зургуудыг ижил хэмжээтэй (224×224) болгосон.
- Зургийн pixel утгууд 0-255-ын хооронд байсныг ToTensor() функц ашиглан 0-1 хооронд болгосон.
- ImageNet стандартчиллын дагуу normalize хийсэн. ImageNet стандартчилал гэдэг нь зургийн RGB сувгийн пикселийн утгуудыг ImageNet өгөгдлийн багц дээр ашиглагдсан дундаж болон стандарт хазайлтын утгаар стандартчилахыг хэлнэ.

2. Текст өгөгдлийн урьдчилсан боловсруулалт

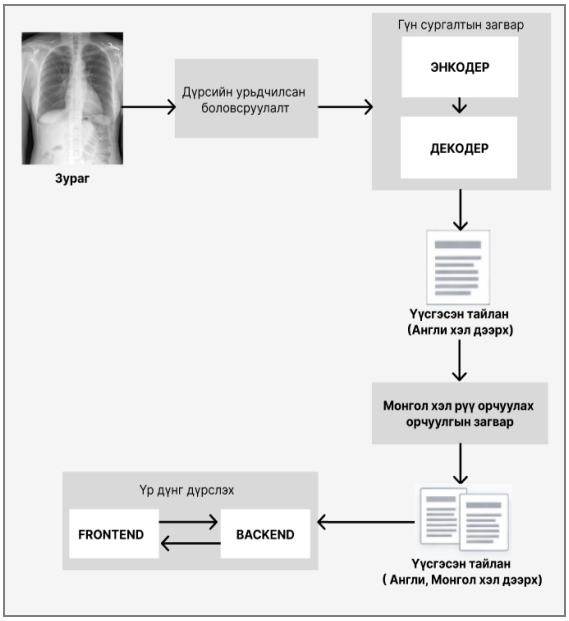
- Бүх цифр болон тусгай тэмдэгтүүдийг устгасан.
- Жижиг үсэг болгосон.
- “өмнөх тайлантай харьцуулахад”, “сайжирсан/муудсан байна” гэх мэт зурагт шууд харагдахгүй мэдээлэл байсан тул “comparison, compared, prior, previous, earlier” гэх түлхүүр үгсийг агуулсан өгүүлбэрүүдийг хассан.
- Мөн өвчтөнд санал болгосон эмчилгээ, “CT авхуулах хэрэгтэй” гэх мэт зөвлөмж агуулсан өгүүлбэрүүдийг хассан.
- Зурагтай огт холбоогүй “email, contact, call, telephone” гэх мэт үгийг агуулсан өгүүлбэрүүдийг устгасан.
- Эмнэлзүйн ач холбогдол хамгийн өндөр “Findings” болон “Impression” баганууудыг нэгтгэн

"Summary" нэртэй шинэ багана үүсгэж, загварын зорилтот багана болгон ашигласан.

С. Загварын архитектур

Энэхүү цээжний рентген зургаас оношилгооны тайлан үүсгэх систем нь энкодер-декодер архитектурт суурилсан бөгөөд дүрсний онцлог гаргах энкодер модуль болон текст үүсгэх декодер модулиас бүрдэнэ. Мөн энкодер-декодер загвараас гарсан англи текстийг монгол хэл рүү орчуулах модулийг нэмж өгснөөр загварын үүсгэсэн тайланг хоёр хэл дээр харах боломжийг бүрдүүлж, эмнэлзүйн практикт хэрэглэх цар хүрээг өргөжүүлсэн.

Зураг 1-т системийн үйл ажиллагааны диаграммыг дүрслэн харуулав.



Зураг 1. Системийн үйл ажиллагааны диаграмм

1. Энкодер

Бид судалгаандаа цээжний рентген зургаас онцлог шинжийг илүү үр дүнтэй ялган авах энкодерийг тодорхойлох зорилгоор дөрвөн өөр харагны архитектурыг туршсан. Үүнд:

1.1. Vision Transformer (ViT)

ViT нь дүрсний шинжилгээнд зориулсан шинэ төрлийн трансформер загвар бөгөөд зургуудыг self-attention механизм ашиглан боловсруулдаг [10]. Оролтын зургийг тогтмол хэмжээтэй жижиг хэсгүүд болгон хуваадаг. Ингэснээр загвар нь зургийн жижиг локал хэсгүүдийг бие даан боловсруулах боломжтой болдог. Дараа нь зургийн жижиг хэсэг бүрийг шугаман проекцын (linear projection) давхаргаар дамжуулж, бага хэмжээст вектор төлөөлөл болгон хувиргана. Энэхүү embedding процесс нь зургийн жижиг хэсэг бүрийн чухал мэдээллийг хадгалж авдаг.

Embedding хийгдсэн жижиг хэсгүүдийг (patch) трансформер энкодерийн давхаргууд руу оруулна. Эдгээр давхаргууд нь self-attention механизм ашиглан жижиг хэсгүүдийн хоорондын хамаарлыг тодорхойлж, зургийн глобал контекстийг барьж авдаг.

Бид судалгаандаа ViT загварын "vit_base_patch16 224" хувилбарыг ашигласан.

1.2. Bidirectional Encoder Representations from Image Transformers (BEiT)'

BEiT нь BERT загвартай адил урьдчилсан сургалтыг дүрсний трансформерт хэрэгжүүлсэн аргачлал юм. BERT өгүүлбэр доторх үгийг mask хийдэгтэй адил BEiT нь зураг доторх patch-уудыг mask хийдэг [11]. Дараа нь загвар нь mask хийгдсэн хэсгийг үлдсэн харагдаж буй patch-ууд дээр тулгуурлан таамагладаг. Энэ процесс нь зургийн хэсгүүдийн хоорондын хамаарлыг ойлгоход тусалдаг. Стандарт ViT нь нэг чиглэлт энкодер ашигладаг бол BEiT нь хоёр чиглэлт транформер энкодер ашигладаг.

Бид судалгаандаа "beit_base_patch16_224" хувилбарыг ашигласан.

1.3. Data Efficient Image Transformer (DEiT)

DEiT нь зураг ангилал болон бусад компьютер харааны даалгаварт зориулагдсан Vision Transformer (ViT)-ийн сайжруулсан хувилбар бөгөөд шошготой өгөгдөл бага үед ч үр дүнтэй суралцах зорилготой архитектур юм. Стандарт ViT-тэй адил транформер энкодерийг feature extractor болгон ашигладаг [5].

DEiT нь Transformer encoder-ийн бага зэрэг өөрчлөгдсөн feed-forward давхаргуудыг ашигласнаар тооцооллын үр ашгийг сайжруулсан [5].

Бид судалгаандаа "deit_base_patch16_224" хувилбарыг ашигласан.

1.4. Swin трансформер

Swin Transformer нь шаталсан бүтэцтэй энкодер архитектур бөгөөд shifted window self-attention механизмыг ашиглан локал болон глобал орон зайн хамаарлыг үр дүнтэй загварчилдаг. Энэ нь уламжлалт Vision Transformer (ViT)-тэй харьцуулахад илүү тооцооллын үр ашигтай бөгөөд өндөр нягтаршилтай дүрсний онцлогийг шаталсан бүтэцтэйгээр гарган авах чадвартай. [12]

Оролтын зургийг жижиг хэсгүүдэд (patch) хуваан токен болгон хувиргаж, window-based self-attention ашиглан локал мэдээллийг боловсруулна. Харин shifted window механизм нь цонхнуудын хооронд мэдээлэл солилцох боломжийг бүрдүүлдэг. Мөн patch merging ашиглан шаталсан бүтэц үүсгэж, олон түвшний дүрсний төлөөллийг гарган авдаг.

2. Декодер

Энкодероос гарсан дүрсний төлөөлөл нь projection буюу conditioning механизмаар декодерийн token embedding орон зайтай уялдуулан холбогдож, декодер нь дараалсан байдлаар оношилгооны тайланг үүсгэнэ. Бид нөөцийн хязгаарлалтын асуудлыг харгалзан үзэн GPT2 архитектурын хөнгөн хувилбар болох DistilGPT загварыг декодер болгон ашигласан. Мөн анагаах ухааны домэйн текст өгөгдөл дээр урьдчилан сургагдсан трансформер хэлний загвар болох BioGPT загварыг декодерт туршсан.

2.1. DistilGPT

DistilGPT нь GPT-2 хэлний загварын чадварыг хадгалсан боловч параметрийн тоо багатай, санах ойн хэрэглээ багасгасан хөнгөн загвар юм.

DistilGPT нь авторегрессив буюу дараагийн токенийг таамаглах 6 трансформер блок, 82 сая гаруй параметртэй, 50,257 токентой загвар юм.

2.2. BioGPT

BioGPT нь био-анагаах ухааны салбарт зориулсан генератив трансформер хэлний загвар бөгөөд PubMed онлайн мэдээллийн сангийн анагаахын шинжлэх ухааны 15 сая орчим өгүүлэлд дээр урьдчилан сургагдсан.

Энэхүү загвар нь үгийн сангаа үүсгэхдээ byte pair encoding (BPE) аргыг ашигласан бөгөөд нийт үгийн хэмжээ 42,384 байна.

BioGPT нь GPT-2 архитектур дээр суурилсан трансформер загвар бөгөөд оролтын дарааллыг Query(Q), Key(K), Value(V) вектор болгон хувиргаж, attention тооцоолол хийдэг [14]. Multi-head attention механизм нь хэд хэдэн attention head-үүдийг зэрэг ажиллуулж, тэдгээрийн үр дүнг нэгтгэн гаргадаг. Үүнийг (1)-д үзүүлсэн бөгөөд тус бүрийн head-ийн attention тооцооллыг (2) илэрхийлнэ.

$$MultiHead(Q,K,V) = Concat(head_1, head_2, \dots head_h)W \quad (1)$$

$$head_i = softmax\left(\frac{Q_i K_i T}{\sqrt{d}}\right) V_i \quad (2)$$

Энд:

- h — attention head-ийн тоо
- d — feature хэмжээ
- W — шугаман хувиргалтын параметр

3. Орчуулгын загвар

Энэхүү системд загварын үүсгэсэн англи хэл дээрх оношилгооны тайланг монгол хэл рүү орчуулах зорилгоор Gemini API-д суурилсан орчуулгын модулийг багтаасан.

Орчуулгын модулийг Google-ийн Gemini 2.5 Pro загварыг API хэлбэрээр ашиглан хэрэгжүүлсэн. Gemini нь Google DeepMind-ийн боловсруулсан том хэмжээний multimodal генератив хэлний загвар бөгөөд текст, зураг болон бусад төрлийн мэдээллийг ойлгож боловсруулах чадвартай. Мөн олон хэлний дэмжлэгтэй тул мэргэжлийн болон эмнэлзүйн нэр томьёог контекстийн дагуу орчуулах боломжтой [15].

D. Туршилтын тохиргоо

1. Өгөгдлийн хуваалт

Энэхүү судалгааны эхний туршилтыг зөвхөн IU Chest X-ray өгөгдөл дээр хийж, дараа нь MIMIC-CXR өгөгдлийг алхамчилсан байдлаар нэмэх замаар загварын гүйцэтгэлийн өөрчлөлтийг судалсан.

Туршилтыг дараах дарааллаар гүйцэтгэсэн.

- Туршилт 1– зөвхөн IU Chest X-ray өгөгдөл дээр загварыг сургаж туршсан.
- Туршилт 2 – IU Chest X-ray өгөгдөл дээр MIMIC-CXR өгөгдлөөс 2000 зураг нэмэн загварыг дахин сургаж туршсан. Ингэхдээ текст өгөгдөлд ямар нэгэн цэвэрлэгээ хийгээгүй.
- Туршилт 3 – MIMIC-CXR өгөгдлөөс 2000 зурагт харгалзах тайлан буюу текст өгөгдөлд цэвэрлэгээ хийж, рентген зурагтай холбоогүй үг болон өгүүлбэрүүдийг хасахад 1938 зурагтай өгөгдлийн багц болсон. Үүнийг IU Chest X-ray өгөгдөл дээр нэмж загварыг дахин сургасан.
- Туршилт 4 – IU Chest X-ray өгөгдөл дээр MIMIC-CXR өгөгдлөөс дахин 2033 зураг, түүнд харгалзах тайлангуудыг нэмэн загварыг сургаж туршсан.

Хүснэгт 2-д туршилт бүрд хэрэгжүүлсэн өгөгдлийн хуваалтыг харуулав.

ХҮСНЭГТ 2. ӨГӨГДЛИЙН ХУВААЛТ

ТУРШИЛТ	Өгөгдөл	Нийт	Сургалт (80%)	Баталгаажуулалт (10%)	Тест (10%)
1	IU	3826 тайлан, 7466 зураг	3062 тайлан, 5954 зураг	382 тайлан, 758 зураг	382 тайлан, 754 зураг
2	IU + MIMIC (цэвэрлээгүй)	5639 тайлан, 9466 зураг	4523 тайлан, 7555 зураг		
3	IU + MIMIC (цэвэрлэсэн)	5589 тайлан, 9404 зураг	4505 тайлан, 7514 зураг		
4	IU + MIMIC (цэвэрлэсэн)	7462 тайлан, 11437 зураг	6698 тайлан, 9925 зураг		

2. Параметрүүдийн тохиргоо

Эхний туршилтад 8 энкодер-декодер загварыг дараах параметрийн тохируулгатайгаар сургасан.

Хүснэгт 3-т загварын параметрүүдийн тохиргоог харуулав.

ХҮСНЭГТ 3. ТОХИРГОО

Бүлэг	Параметрийн тохиргоо	
Vision Encoder	Оролтын зургийн хэмжээ	224×224
	Patch size	16x16
	Hidden dimension	768
Language Decoder	Max token length	128
	Hidden size	768 (DistilGPT), 1024 (BioGPT)
Training	Batch size	4
	Learning rate	1e-5
	Optimizer	AdamW
	Epoch	4
	Loss function	Cross-entropy

3. Гиперпараметрийн тохируулга

Загварын гүйцэтгэлийг сайжруулах зорилгоор сургалтын гиперпараметрүүд болох learning rate, batch size болон epoch-ийн утгуудыг өөрчлөн туршив.

Хүснэгт 4-д гиперпараметрийн туршсан утгыг харуулав.

ХҮСНЭГТ 4. ГИПЕРПАРАМЕТРИЙН ТОХИРУУЛГА

Параметр	**Туршсан утгууд**
Learning rate	1e-4, 5e-5, 1e-5
Batch size	2, 4, 8
Epoch	4, 8

Гиперпараметрийн тохируулгыг эхний туршилтад хамгийн сайн үр дүн үзүүлсэн Swin + BioGPT, DEiT + Biogpt загварууд дээр туршсан. Туршилтын үр дүнгээс харахад learning rate = 5e-5, batch size= 4, epoch = 4 байх үед загвар хамгийн сайн гүйцэтгэл үзүүлсэн. Иймд дараагийн туршилтуудад сургалтын гиперпараметрийн тохиргоог learning rate = 5e-5 болгон шинэчилж ашиглав.

Е. Үнэлгээний хэмжүүрүүд

1. Үгийн түвшний үнэлгээний хэмжүүр

- BLEU (Bilingual Evaluation Understudy)

BLEU нь машин орчуулгын системийг үнэлэх зорилгоор Papineni [17] нарын санал болгосон автомат үнэлгээний арга бөгөөд загварын үүсгэсэн текст болон reference текстийн n-gram давхцалыг хэмжих зарчимд тулгуурладаг.

- METEOR (Metric for Evaluation of Translation with Explicit ORdering)

METEOR хэмжүүрийг Banerjee болон Lavie нар [18] BLEU-ийн сул талуудыг сайжруулах зорилгоор санал болгосон. Энэ хэмжүүр нь precision болон recall хоёрыг хамтад нь авч үздэг бөгөөд үгийн дараалал, утгын ойролцоо байдал зэргийг харгалзан үздэг.

METEOR нь unigram тохирлыг хэмжиж, precision болон recall-ийг хослуулах бөгөөд stemming болон synonym matching-ийг дэмждэг.

Иймээс BLEU-тэй харьцуулахад хүний үнэлгээтэй илүү сайн уялддаг гэж үздэг.

- ROUGE (Recall-Oriented Understudy for Gisting Evaluation)

ROUGE хэмжүүрийг Lin [19] текстийн автомат хураангуйллын системийг үнэлэх зорилгоор санал болгосон бөгөөд reference болон үүсгэсэн текстийн n-gram recall-ийг хэмждэг. ROUGE-L нь үгийн дарааллын бүтцийг илүү сайн тусгадаг тул текстийн семантик бүтцийн төстэй байдлыг үнэлэхэд ашиглагддаг.

2. Эмнэлзүйн үнэлгээний хэмжүүр

Цээжний рентген зургаас оношилгооны тайлан үүсгэх даалгаварт зөвхөн үгийн түвшний төстэй байдлыг хэмжих нь хангалтгүй байдаг. Учир нь үгсийн төстэй байдал өндөр байсан ч эмнэлзүйн мэдээлэл буруу байх боломжтой. Иймээс судалгаанд үүсгэсэн тайлангийн эмнэлзүйн зөв байдлыг үнэлэх зорилгоор CheXbert болон RadGraph-д суурилсан хэмжүүрүүдийг ашигласан.

- CheXbert F1

CheXbert нь цээжний рентген зургийн тайлангаас эмнэлзүйн шинж тэмдгүүдийг автоматаар тодорхойлох зорилготой нейрон сүлжээнд суурилсан ангилагч юм. Энэ арга нь үүсгэсэн тайлан болон бодит тайлангаас 14 төрлийн ажиглалтыг илрүүлэн харьцуулж F1-score хэлбэрээр үнэлгээ өгдөг [20]. Хүснэгт 5-д тэдгээр 14 төрлийн ажиглалтыг харуулав.

ХҮСНЭГТ 5. CHEXBERT АЖИГЛАЛТУУД

Ажиглалтууд	
Pneumonia	Уушгины хатгалгаа
Fracture	Ясны хугарал
Consolidation	Уушгины нягтаршил
Enlarged Cardiom	Зүрх болон медиастины томрол
No finding	Эмгэг өөрчлөлт илрээгүй
Pleural Other	Гялтангийн бусад эмгэг
Cardiomegaly	Зүрх томрол
Pneumothorax	Уушгины хийн хуралдалт
Atelectasis	Уушги агшилт
Support Devices	Эмнэлгийн төхөөрөмж
Edema	Хаван
Pleural Effusion	Гялтангийн шингэн
Lung Lesion	Уушгины гэмтэл
Lung Opacity	Уушгины сүүдэр

- RedGraph F1

RadGraph нь тайлан дахь эмнэлзүйн энтити болон тэдгээрийн хоорондын харилцааг тодорхойлох зорилготой мэдлэгийн граф дээр суурилсан арга юм [21]. Энэ арга нь тайлангийн доторх эмнэлзүйн ойлголтуудыг граф бүтэц болгон илэрхийлж, бодит болон үүсгэсэн тайлангийн хоорондын семантик төстэй байдлыг хэмждэг.

RadGraph арга нь тайлан дахь эмнэлзүйн ойлголтууд (жишээ нь: эмгэг, анатомийн бүтэц, ажиглалт) болон тэдгээрийн хоорондын холбоог entity–relation граф хэлбэрээр илэрхийлдэг.

Энд:

- entity - эмнэлзүйн ойлголт (opacity, effusion гэх мэт)
- relation - ойлголтуудын хоорондын холбоо (located_at)

IV. ҮР ДҮН

А. Энкодер-Декодер хослолын харьцуулалт

Эхний туршилтад IU Chest X-ray өгөгдлийн багц дээр 4 харааны энкодер болох Swin, ViT, DEiT, BeiT болон 2 хэлний декодер болох DistilGPT, BioGPT-ийн нийт 8 хослолыг сурган гүйцэтгэлийг үнэлэв.

Хүснэгт 6-д үзүүлсэн үгийн түвшний хэмжүүрүүдийн үр дүнг харьцуулахад энкодер-декодер архитектурын сонголт нь тайлан үүсгэх загварын гүйцэтгэлд тодорхой нөлөө үзүүлж байгааг харж болно.

- Эдгээр энкодер-декодер хослолоос Swin + BioGPT загвар хамгийн сайн үр дүн үзүүлсэн

байна. Энэхүү загвар нь BLEU-1 (0.7266), BLEU-4 (0.5028), METEOR (0.404), ROUGE-1 (0.7416) зэрэг бүх үгийн түвшний хэмжүүрүүд дээр хамгийн өндөр оноог үзүүлсэн. Энэ нь Swin Transformer-ийн шаталсан бүтэц нь зургийн орон зайн нарийн мэдээллийг илүү сайн хадгалж, BioGPT decoder нь эмнэлзүйн текстийн хэлбэрийг илүү сайн үүсгэж байгааг харуулж байна.

- Дараагийн хамгийн сайн гүйцэтгэлтэй загвар нь DEiT + BioGPT байна. Энэ загвар BLEU-2 (0.6024), BLEU-3 (0.5316), ROUGE-2 (0.5357) зэрэг хэмжүүрүүд дээр өндөр гүйцэтгэл үзүүлсэн. Энэ нь DEiT encoder нь рентген зургийн дүрсийн онцлогийг илүү сайн ялган сурч, текст decoder-д илүү чанартай feature дамжуулж байгааг илтгэнэ.
- BioGPT декодер ашигласан загварууд DistilGPT декодертой загваруудтай харьцуулахад илүү өндөр үзүүлэлт үзүүлсэн байна. Энэ нь BioGPT загвар нь biomedical болон clinical текст дээр урьдчилан сургасан тул оношилгооны тайлан үүсгэх даалгаварт илүү тохиромжтой байж болохыг харуулж байна.

Хүснэгт 6-д туршилтын үр дүнгийн үгийн түвшний хэмжүүрүүдийг харууллаа.

ХҮСНЭГТ 6. ҮГИЙН ТҮВШНИЙ ХЭМЖҮҮРҮҮД

Загвар	**Үгийн түвшний хэмжүүрүүд**							
	BLEU-1	***BLEU-2***	***BLEU-3***	***BLEU-4***	***METEOR***	***ROUGE-1***	***ROUGE-2***	***ROUGE-L***
Swin+ DistilGPT	0.6955	0.5879	0.5164	0.4637	0.378	0.7185	0.5265	0.6827
ViT+ DistilGPT	0.6949	0.589	0.5177	0.4659	0.3799	0.7159	0.529	0.6833
DEiT+ DistilGPT	0.70	0.5948	0.5246	0.4737	0.3817	0.7203	0.5365	0.6866
BEiT+ DistilGPT	0.6948	0.5892	0.5205	0.4673	0.3787	0.7177	0.5311	0.6846
ViT+BioGPT	0.698	0.578	0.4993	0.4422	0.3794	0.7152	0.5057	0.6757
BEiT+ BioGPT	0.7104	0.5993	0.5241	0.4708	0.3888	0.7269	0.53	0.6894
DEiT+ BioGPT	**0.71**	**0.6024**	**0.5316**	**0.4791**	**0.3891**	**0.7253**	**0.5357**	**0.6901**
Swin+ BioGPT	**0.7266**	**0.6208**	**0.5527**	**0.5028**	**0.404**	**0.7416**	**0.5555**	**0.7056**

ХҮСНЭГТ 7. ЭМНЭЛЗҮЙН ХЭМЖҮҮРҮҮД

Загвар	**Эмнэлзүйн хэмжүүр**	
	CheXbert F1	***RG_ER F1***
Swin+ DistilGPT	0.6754	0.5364
ViT+ DistilGPT	0.7016	0.5397
DEiT+ DistilGPT	0.678	0.5456
BEiT+ DistilGPT	0.699	0.5521
ViT+BioGPT	0.759	0.5297
BEiT+BioGPT	0.759	0.5473
DEiT+BioGPT	**0.7592**	**0.5464**
Swin+BioGPT	**0.7408**	**0.5657**

Хүснэгт 7-д үзүүлсэн эмнэлзүйн хэмжүүрүүд нь энкодер-декодер ялгаатай хослолуудаас хамааран өөр үзүүлэлттэй байна.

1. Эмнэлзүйн хэмжүүрийн хувьд BioGPT декодер ашигласан загварууд илүү өндөр гүйцэтгэл үзүүлж байгаа буюу эмнэлзүйн хувьд илүү үнэн зөв тайланг үүсгэж байна.
2. CheXbert F1 хэмжүүрийн хувьд DEiT + BioGPT загвар хамгийн өндөр гүйцэтгэл (0.7592) үзүүлсэн нь энэхүү загвар эмнэлзүйн энтитиг илүү зөв илэрхийлсэн тайланг үүсгэж байгааг илтгэнэ.
3. RG_ER F1 хэмжүүрийн хувьд Swin + BioGPT загвар хамгийн өндөр үзүүлэлт (0.5657) үзүүлсэн нь энэхүү эмнэлзүйн энтити болон тэдгээрийн хоорондын харилцааг илүү сайн гаргаж ирж байгааг илтгэнэ.

Эдгээр үр дүнгээс харахад DEiT + BioGPT болон Swin + BioGPT загварууд үгийн түвшний болон эмнэлзүйн хэмжүүрүүд дээр хамгийн өндөр гүйцэтгэлтэй байгаа тул дараагийн өгөгдөл нэмэгдүүлэх туршилтыг энэ хоёр загвар дээр гүйцэтгэсэн.

B. Статистикийн ач холбогдлын шинжилгээ

Хүснэгт 6-д үзүүлсэн үр дүнгээс хамгийн өндөр гүйцэтгэлтэй Swin+BioGPT болон түүнтэй ойролцоо гүйцэтгэлтэй DEiT+BioGPT загваруудыг сонгон авч, гүйцэтгэлийн ялгаа статистикийн хувьд ач холбогдолтой эсэхийг шалгав.

Энэхүү шинжилгээнд paired bootstrap resampling аргыг ашигласан бөгөөд тестийн өгөгдлийн багц дахь зураг бүрд загварын үүсгэсэн тайлангийн BLEU-4, METEOR, ROUGE-L, CheXbert F1, RG_ER F1 хэмжүүрүүдийг ашиглан 1,000 удаагийн bootstrap давталт хийж, дундаж ялгааны 95% confidence interval болон p-value-г тооцоолсон. Хэрэв p-value нь 0.05-аас бага, мөн confidence interval нь 0 утгыг агуулаагүй тохиолдолд ялгааг статистикийн хувьд ач холбогдолтой гэж үзэв.

ХҮСНЭГТ 8. СТАТИСТИКИЙН АЧ ХОЛБОГДОЛ

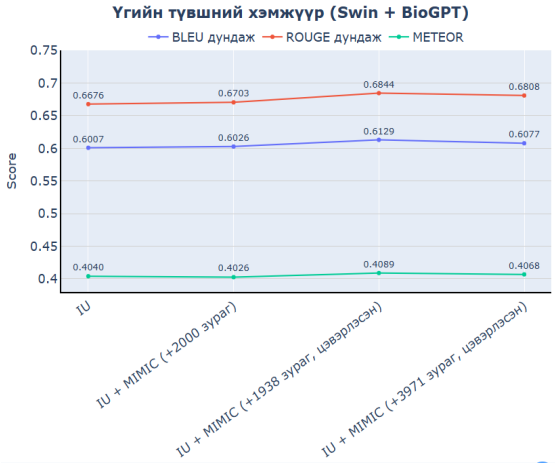
Хэмжүүр	Ялгаа	95% CI	p-value	Дүгнэлт
BLEU-4	+0.0237	[0.0112, 0.0368]	0.003	Significant
METEOR	+0.0149	[0.0061, 0.0245]	0.008	Significant
ROUGE-L	+0.0155	[0.0043, 0.0271]	0.021	Significant
CheXbert F1	-0.0184	[-0.0345, -0.0042]	0.017	Significant
RG_ER F1	+0.0193	[0.0067, 0.0328]	0.006	Significant

Хүснэгт 8-д үзүүлсэн үр дүнгээс харахад Swin+BioGPT болон DEiT+BioGPT загваруудын хоорондын ялгаа бүх хэмжүүрүүд дээр статистикийн хувьд ач холбогдолтой гарсан байна (p-value < 0.05). Иймд Swin+BioGPT болон DEiT+BioGPT загваруудын гүйцэтгэлийн ялгаа нь санамсаргүй хэлбэлзлээс үүсээгүй харин загварын архитектурын ялгаанаас үүдэлтэй бодит ялгаа болохыг илтгэнэ.

C. Өгөгдлийн хэмжээг нэмэгдүүлэх туршилт

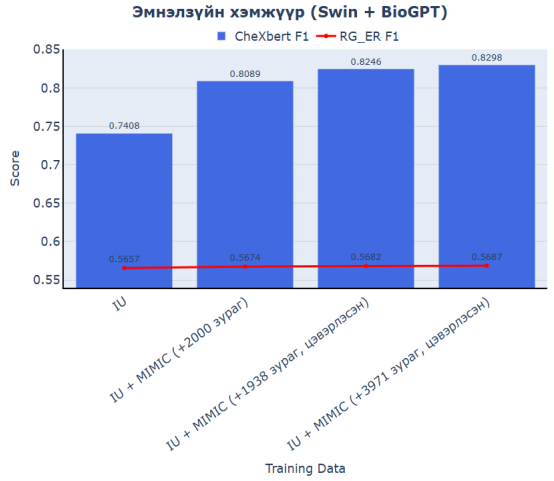
1. Swin + BioGPT

Өгөгдлийн хэмжээг нэмэгдүүлэх туршилтыг гүйцэтгэн Swin + BioGPT загварын гүйцэтгэлийн үгийн түвшний хэмжүүрийг зураг 2-т, эмнэлзүйн хэмжүүрийг зураг 3-д үзүүлэв.



Зураг 2. Үгийн түвшний хэмжүүр (Swin + BioGPT загвар)

Зураг 2-т үзүүлснээр анхны IU өгөгдөл дээр сургасан загварын BLEU дундаж 0.60 орчим, ROUGE дундаж 0.66 орчим, METEOR 0.40 байсан байна. Харин MIMIC өгөгдлийг нэмснээр үгийн түвшний хэмжүүрүүд мэдэгдэхүйц биш ч бага зэрэг өсөх хандлагатай байна. Ялангуяа MIMIC-CXR өгөгдлийн текстийг цэвэрлэснээр гүйцэтгэл нэмэгдэж байна. Энэ нь өгөгдлийн хэмжээ нэмэгдэх тусам гүйцэтгэл тодорхой хэмжээнд сайжирдаг боловч өгөгдлийн чанар мөн адил чухал нөлөөтэй болохыг харуулж байна.



Зураг 3. Эмнэлзүйн хэмжүүр (Swin + BioGPT загвар)

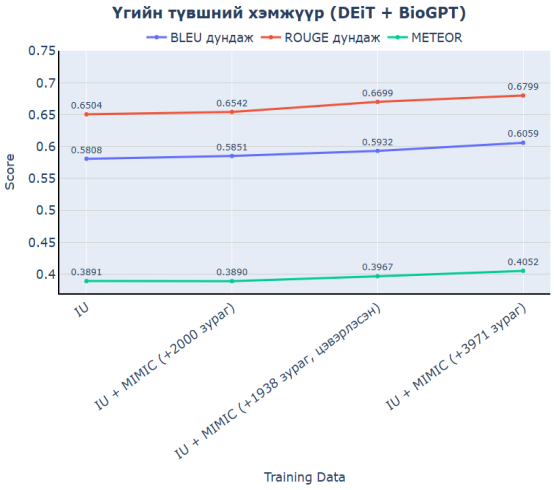
Зураг 3-т үзүүлснээр анхны IU өгөгдөл дээр сургасан загварын CheXbert F1 оноо 0.7408 байсан бол MIMIC өгөгдлийг нэмснээр энэ үзүүлэлт мэдэгдэхүйц өссөн. Тухайлбал:

- +2000 MIMIC зураг нэмэхэд CheXbert F1 = 0.8089 (0.06-аар нэмэгдсэн),
- +1938 зураг, цэвэрлэсэн тайлан бүхий MIMIC өгөгдөл нэмэхэд 0.8246 (0.02-аар нэмэгдсэн),
- 3971 зураг, цэвэрлэсэн тайлан бүхий MIMIC өгөгдөл нэмэхэд CheXbert F1 хэмжүүр 0.8298 хүртэл өссөн байна.

Эдгээр үр дүнгээс харахад MIMIC-CXR өгөгдлийг IU Chest X-ray өгөгдөл дээр нэмэх нь загварын гүйцэтгэлийг сайжруулж байна. Ялангуяа эмнэлзүйн хэмжүүрүүд дээр мэдэгдэхүйц өсөлт ажиглагдсан бөгөөд энэ нь илүү олон төрлийн эмнэлзүйн мэдээлэл бүхий өгөгдөл ашиглах нь загварын үүсгэсэн тайлангийн эмнэлзүйн үнэн зөв байдлыг нэмэгдүүлэх боломжтойг харуулж байна.

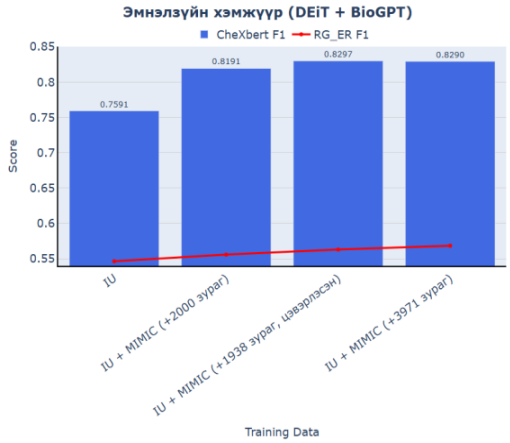
2. DEiT + BioGPT

Өгөгдлийн хэмжээг нэмэгдүүлэх туршилтыг гүйцэтгэн DEiT + BioGPT загварын гүйцэтгэлийн үгийн түвшний хэмжүүрийг зураг 4-т, эмнэлзүйн хэмжүүрийг зураг 5-т үзүүлэв.



Зураг 4. Үгийн түвшний хэмжүүр (DEiT + BioGPT загвар)

Үгийн түвшний хэмжүүрүүдийн хувьд MIMIC өгөгдөл нэмэгдэх тусам BLEU дундаж, ROUGE дундаж болон METEOR хэмжүүрүүд тогтвортой өсөх хандлага ажиглагдсан. Анхны IU өгөгдөл дээр BLEU дундаж ойролцоогоор 0.58, ROUGE дундаж 0.65, METEOR 0.39 байсан бол MIMIC өгөгдлийг нэмснээр эдгээр хэмжүүрүүд аажмаар нэмэгдсэн байна. Ялангуяа MIMIC өгөгдлийн 3971 зураг нэмсэн туршилтад BLEU болон ROUGE хэмжүүрүүд хамгийн өндөр утгыг үзүүлсэн бөгөөд энэ нь өгөгдлийн хэмжээ нэмэгдэх нь үүсгэсэн тайлангийн үгийн түвшний чанарыг сайжруулж байгааг харуулж байна.



Зураг 5. Эмнэлзүйн хэмжүүр DEiT + BioGPT загвар

Зураг 5-т үзүүлснээр анхны IU өгөгдөл дээр CheXbert F1 0.7591 байсан бол MIMIC өгөгдлийг нэмснээр энэ үзүүлэлт 0.8290 хүртэл өссөн байна. Энэ нь үүсгэсэн тайлан нь эмнэлзүйн энтитиг илүү зөв илэрхийлэх чадвар нэмэгдсэнийг харуулж байна.

3. DEiT + BioGPT болон Swin + BioGPT харьцуулалт

Хүснэгт 9-д IU+MIMIC нийлмэл өгөгдөл дээр сургагдсан DEiT + BioGPT болон Swin + BioGPT загварын үгийн түвшний хэмжүүрийг, хүснэгт 10-д эмнэлзүйн хэмжүүрийн харьцуулалтыг харуулав.

ХҮСНЭГТ 9. ҮГИЙН ТҮВШНИЙ ХЭМЖҮҮРҮҮД

Загвар	Үгийн түвшний хэмжүүрүүд							
	BLEU-1	***BLEU-2***	***BLEU-3***	***BLEU-4***	***METEOR***	***ROUGE-1***	***ROUGE-2***	***ROUGE-L***
DEiT+ BioGPT	0.7206	0.6271	0.5625	0.5133	0.4052	0.7444	0.5773	0.7179
SWIN +BioGPT	**0.7226**	**0.6287**	**0.564**	**0.5154**	**0.4068**	**0.7461**	**0.578**	**0.7182**

ХҮСНЭГТ 10. ЭМНЭЛЗҮЙН ХЭМЖҮҮРҮҮД

Загвар	Эмнэлзүйн хэмжүүр	
	CheXbert F1	***RG_ER F1***
DEiT+BioGPT	0.829	0.5684
SWIN+BioGPT	**0.8298**	**0.5687**

Өгөгдлийн хэжээг нэмэхээс өмнө SWIN нь үгийн түвшний хэмжүүрүүдийн хувьд илүү бол, DEiT нь эмнэлзүйн Chexbert хэмжүүрийн хувьд илүү байсан.

MIMIC өгөгдөл нэмснээр SWIN нь үгийн түвшний хэмжүүрүүдийн хувьд илүү хэвээрээ ба эмнэлзүйн Chexbert хэмжүүрийн хувьд DEiT-тэй бараг ижил болсон.

Иймд Swin + BioGPT нь илүү тэнцвэртэй сайн загвар болсон гэж үзээд дараагийн data augmentation хийх туршилтыг Swin + BioGPT загвар дээр гүйцэтгэсэн.

D. Data augmentation туршилт

Энэхүү туршилтад сургалтын өгөгдлийн хэмжээг нэмэгдүүлэх болон загварын ерөнхийлөх чадварыг сайжруулах зорилгоор on-the-fly data augmentation-ийг дараах тохиргоотойгоор хэрэгжүүлсэн.

Геометрийн өөрчлөлт:

- ±5° эргүүлэх: Зураг бага зэрэг хазайж орно.
- ±2% шилжүүлэх: Зургийг бага зэрэг дээш/доош/зүүн/баруун хөдөлгөнө.
- 0.95–1.05 масштаблах: 5% орчим томруулах/жижигсүүлэх.

Гэрэлтүүлгийн өөрчлөлт:

- Гэрэлтүүлэг ±8%: Зураг бага зэрэг бүдэг эсвэл тод болно.
- Контраст ±12%: Уушгины нягтрал илүү тод/зөөлөн харагдана.

Геометр болон гэрэлтүүлгийн өөрчлөлтийн магадлалуудыг 70%, 80%, 90% гэсэн гурван өөр утгаар туршиж, баталгаажуулалтын багц дээрх гүйцэтгэлийг харьцуулан үнэлсэн.

Туршилтын үр дүнгээс харахад augmentation хэрэгжих магадлал (p=0.9) 90% байх үед загвар хамгийн сайн гүйцэтгэл үзүүлсэн. Энэ нь augmentation илүү олон удаа хэрэгжих үед загвар илүү олон төрлийн дүрсийн хувилбар дээр суралцаж, ерөнхийлөх чадвар нь сайжирч байгааг харуулж байна.

Хүснэгт 11-т Swin+BioGPT загварын data augmentation хийхээс өмнөх болон хийсний дараах үгийн түвшний хэмжүүрийг харуулав.

ХҮСНЭГТ 11. ҮГИЙН ТҮВШНИЙ ХЭМЖҮҮРҮҮД

SWIN + BioGPT	Үгийн түвшний хэмжүүрүүд							
	BLEU-1	***BLEU-2***	***BLEU-3***	***BLEU-4***	***METEOR***	***ROUGE-1***	***ROUGE-2***	***ROUGE-L***
DA-гүй	0.7226	0.6287	0.564	0.5154	0.4068	0.7461	**0.578**	0.7182
DA-тай	**0.7314**	**0.6339**	**0.5686**	**0.5185**	**0.4149**	**0.7477**	0.5748	**0.7212**

Хүснэгтэд 12-д Swin+BioGPT загварын data augmentation хийхээс өмнөх болон хийсний дараах эмнэлзүйн хэмжүүрийг харуулав.

ХҮСНЭГТ 12. ЭМНЭЛЗҮЙН ХЭМЖҮҮРҮҮД

SWIN+ BioGPT	Эмнэлзүйн хэмжүүр	
	CheXbert F1	***RG_ER F1***
DA-гүй	**0.8298**	**0.5687**
DA-тай	0.8254	0.5645

Data augmentation хийсний үр дүнд загварын гүйцэтгэл бага зэрэг сайжирсан байна. Тухайлбал:

- BLEU-1, BLEU-2, BLEU-3, BLEU-4 хэмжүүрүүд 0.7226, 0.6287, 0.5640, 0.5154 байсан бол augmentation ашигласан үед 0.7314, 0.6339, 0.5686, 0.5185 болж нэмэгдсэн байна.

Энэ нь тайлангийн үгийн түвшний төстэй байдал сайжирсныг харуулж байна.

- Харин эмнэлзүйн хэмжүүр мэдэгдэхүйц өөрчлөгдөөгүй ч IU+MIMIC нийлмэл өгөгдөл дээр үзүүлсэн өсөлтөө тогтвортойгоор авч үлдсэн. Үүний үр дүнд үгзүйн болон эмнэлзүйн хэмжүүрүүд хоёулаа сайжирснаар тайлан нь найруулгын хувьд чанартай бөгөөд оношилгооны хувьд үнэн зөв болж, загвар нь бодит эмнэлзүйн хэрэглээнд илүү найдвартай болсон.

E. Загварын үүсгэсэн тайлангийн жишээ

Хамгийн сайн гүйцэтгэлтэй хослол болох Swin+ BioGPT загвараар үүсгэсэн цээжний рентген зургийн оношилгооны тайлангийн жишээг харууллаа.

Оролтын зураг



Бодит тайлан

The cardiomediastinal silhouette is within normal limits for appearance. The thoracic aorta is tortuous. A calcified granuloma is identified in the right middle lobe. No focal areas of pulmonary consolidation. No pneumothorax. No pleural effusion. The thoracic spine appears intact. No acute displaced rib fractures. No acute intrathoracic abnormality

Үүсгэсэн тайлан

The heart cardiomediastinal silhouette is within normal limits for size. Subtle granuloma is identified in the lobe. No pneumothorax. No pleural effusion No focal areas of pulmonary consolidation. No acute intrathoracic abnormality.

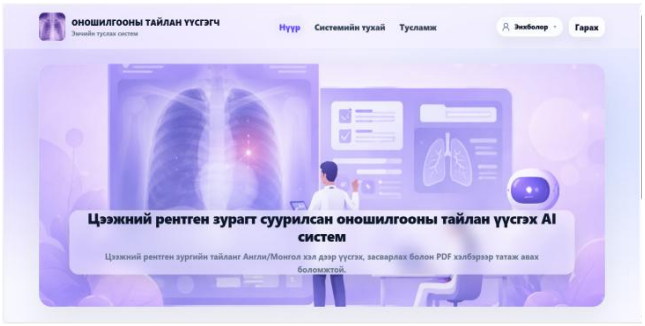
Орчуулгын загварын үр дүн

Зүрхний хэмжээ хэвийн хязгаарт байна. Уушгины дэлбэнд бүдэг гранулома илэрсэн. Уушгинд хий хуримтлагдаагүй. Гялтангийн хөндийд шингэн илрээгүй. Уушгины голомтот нягтаршил илрээгүй. Хурц цээжний хөндийн эмгэг өөрчлөлт илрээгүй.

Загварын үүсгэсэн тайлан нь бодит тайлангийн ихэнх гол эмнэлзүйн мэдээллийг зөв илэрхийлсэн. Гэхдээ зарим нарийвчилсан анатомийн тайлбаруудыг бүрэн илэрхийлж чадаагүй байна. Жишээ нь: Уушгины дэлбэнд бүдэг гранулома илэрснийг илэрхийлсэн боловч “баруун дэлбэнд” гэдгийг дурдаагүй.

F. Хэрэглэгчийн интерфейс

Систем нь хэрэглэгчдэд цээжний рентген зургийг оруулж, автоматаар оношилгооны тайлан үүсгэх, тайланг монгол хэл дээр харах, засах, pdf-ээр татах боломж бүхий веб интерфейстэй. Хэрэглэгч зураг оруулсны дараа backend сервер дээр гүн сургалтын загвар ажиллаж, үүсгэсэн тайланг дэлгэцэд харуулна. Зураг 6-т веб системийн нүүр хуудсыг, зурагт 7-д тайлан үүсгэсэн байгаа байдлыг харуулав.



Зураг 6. Нүүр хуудас



Зураг 7. Тайлан үүсгэсэн байгаа байдал

V. ДҮГНЭЛТ

Энэхүү судалгааны ажлын хүрээнд цээжний рентген зураг суурилсан оношилгооны тайлан үүсгэх AI загварыг боловсруулж, түүний гүйцэтгэлийг үгийн түвшний болон эмнэлзүйн хэмжүүрээр үнэллээ. Судалгаанд Swin Transformer, ViT, DEiT, BEiT энкодеруудыг DistilGPT болон BioGPT декодертой хослуулан туршсан.

Туршилтын үр дүнгээс харахад BioGPT декодер ашигласан загварууд DistilGPT декодертой загваруудаас илүү сайн гүйцэтгэл үзүүлсэн бөгөөд эмнэлзүйн хэмжүүрүүд дээр илүү өндөр үзүүлэлттэй байв. Энэ нь BioGPT загвар оношилгооны тайланг эмнэлзүйн хувьд үнэн зөв үүсгэж байгааг илтгэнэ. Мөн Swin+BioGPT хослол нь үгийн түвшний болон эмнэлзүйн хэмжүүрүүд дээр хамгийн сайн үр дүн үзүүлсэн.

Өгөгдлийн хэмжээг нэмэгдүүлэх туршилтаар IU Chest X-ray өгөгдөлд MIMIC-CXR өгөгдлийг нэмэхэд загварын үгийн түвшний хэмжүүр мэдэгдэхүйц өсөөгүй ч, эмнэлзүйн хэмжүүр тогтвортой өссөн. Энэ нь өгөгдлийн хэмжээг нэмж, илүү олон төрлийн өвчний тохиолдлуудыг агуулсан өгөгдөл дээр загварыг сургах нь тайлангийн чанарыг сайжруулж, эмнэлзүйн хувьд бодит, зөв тайланг үүсгэх боломжтойг харуулсан.

IU+MIMIC нийлмэл өгөгдөл дээр data augmentation хэрэгжүүлснээр үгийн түвшний хэмжүүрүүд бага зэрэг сайжирсан. Энэ нь data augmentation ашигласнаар, энкодер жижиг геометр эсвэл гэрэлтүүлгийн өөрчлөлтөд мэдрэмтгий бус болж, зургийн онцлогийг илүү тогтвортой төлөөлөх боломжтой болсныг харуулж байна. Үүний үр дүнд загварын үүсгэсэн тайлан бодит тайлантай давхцах n-gram (үг, хэлц) нэмэгдсэн буюу үгийн түвшний хэмжүүр өссөн. Ингэснээр үүсгэсэн тайлан нь бодит тайлантай найруулга болон үгзүйн хувьд илүү адил болсон.

Харин IU+MIMIC нийлмэл өгөгдөл дээр data augmentation хэрэгжүүлэхэд загварын эмнэлзүйн

хэмжүүр өсөөгүй боловч өмнөх туршилтуудад хүрсэн гүйцэтгэлийг тогтвортой хадгалсан. Үүний үр дүнд загварын үгүйн болон эмнэлзүйн хэмжүүрүүд суурь загвартай харьцуулахад сайжирч, найруулгын хувьд чанартай, эмнэлзүйн мэдээллийг зөв илэрхийлсэн тайлан үүсгэх чадвартай болсон. Энэ нь боловсруулсан системийг бодит практик хэрэглээнд ашиглах боломжийг нэмэгдүүлж байна.

Ном зүй

[1] Radiological Society of North America. "*AI chest X-ray classification*." RSNA News, (2022). https://www.rsna.org/news/2022/september/ai-chest-x-ray-classification

[2] World Health Organization. "Sustained Efforts to Reduce the Tuberculosis Burden in Mongolia." *WHO Mongolia News*, (2025). Sustained efforts to reduce the tuberculosis burden in Mongolia

[3] Munkhbaatar Dagvasumberel, Tugsjargal Purevsukh, and Oyundari Gonchigsuren. "Training of Radiology Residents in Mongolia." *Korean Journal of Radiology* 26, no. 2 (2025): 94–96. https://doi.org/10.3348/kjr.2024.1095

[4] Yuan, Tao, L. Rodney Long, Zhiyun Xue, Sameer Antani, George R. Thoma, and Xiaolei Huang. "Multimodal Recurrent Model with Attention for Automated Radiology Report Generation." in *Medical Image Computing and Computer Assisted Intervention – MICCAI* (2018), 457–466. Springer, (2018). https://doi.org/10.1007/978-3-030-00928-1

[5] Raminedi, Santhosh, S. Shridevi, and Daehan Won. "Multi-modal transformer architecture for medical image analysis and automated report generation." *Scientific Reports* 14 (2024): 19281 Multi-modal transformer architecture for medical image analysis and automated report generation | Scientific Reports

[6] Singh, Prateek, and Sudhakar Singh. "ChestX-Transcribe: a multimodal transformer for automated radiology report generation from chest X-rays." *Frontiers in Digital Health* 7 (2025). Frontiers | ChestX-Transcribe: a multimodal transformer for automated radiology report generation from chest x-rays

[7] Demner-Fushman, Dina, Marc D. Kohli, Marc B. Rosenman, Sonya E. Shooshan, Laritza Rodriguez, Sameer Antani, George R. Thoma, and Clement J. McDonald. "Preparing a collection of radiology examinations for distribution and retrieval." *Journal of the American Medical Informatics Association* 23, no. 2 (2016): 304–310. Preparing a collection of radiology examinations for distribution and retrieval | Journal of the American Medical Informatics Association | Oxford Academic

[8] Johnson, Alistair E. W., Tom J. Pollard, Seth J. Berkowitz, Nathaniel R. Greenbaum, Matthew P. Lungren, Chih-ying Deng, Roger G. Mark, and Steven Horng. "*MIMIC-CXR: A large publicly available database of labeled chest radiographs.*" arXiv preprint arXiv:1901.07042 (2019). [1901.07042] MIMIC-CXR-JPG, a large publicly available database of labeled chest radiographs

[9] Ian Goodfellow,Yoshua Bengio and Aaron Courville. *"Deep Learning", MIT Press, Goodfellow-et-al-2016 (2016), Deep Learning*

[10] Dosovitskiy, Alexey, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, et al. "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale." International Conference on Learning Representations (ICLR), 2021. [2010.11929] An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale

[11] Bao, Hangbo, Li Dong, and Furu Wei. "BEiT: BERT Pre-Training of Image Transformers." International Conference on Learning Representations (ICLR), 2022 [2106.08254] BEiT: BERT Pre-Training of Image Transformers

[12] Liu, Ze, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows." International Conference on Computer Vision (ICCV), 2021. 2103.14030

[13] Sanh, Victor, Lysandre Debut, Julien Chaumond, and Thomas Wolf. "DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter." NeurIPS Workshop on Energy Efficient Machine Learning, 2019. [1910.01108] DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter

[14] Luo, Renqian, Luyang Xu, Ruiyan Chang, Xiaoyang Liu, Xu Sun, and Jianfeng Gao."BioGPT: Generative Pre-trained Transformer for Biomedical Text Generation and Mining." Briefings in Bioinformatics 23, no. 6 (2022). 2210.10341

[15] Team Gemini, Rohan Anil, Sebastian Borgeaud, Yonghui Wu, Zhifeng Chen, Matthias Minderer, et al. "Gemini: A Family of Highly Capable Multimodal Models." arXiv preprint arXiv:2312.11805, 2024. [2312.11805] Gemini: A Family of Highly Capable Multimodal Models

[16] Shorten, Connor, and Taghi M. Khoshgoftaar. "A survey on image data augmentation for deep learning." Journal of Big Data 6, no. 1 (2019): 60. A survey on Image Data Augmentation for Deep Learning | Journal of Big Data | Springer Nature Link

[17] Papineni, Kishore, Salim Roukos, Todd Ward, and Wei-Jing Zhu. "BLEU: A Method for Automatic Evaluation of Machine Translation." Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), 2002: 311–318. bleu.dvi

[18] Banerjee, Satanjeev, and Alon Lavie. "METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments." Proceedings of the ACL Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization, 2005: 65–72. METEOR: An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments - ACL Anthology

[19] Lin, Chin-Yew. "ROUGE: A Package for Automatic Evaluation of Summaries." Proceedings of the ACL Workshop on Text Summarization Branches Out, 2004: 74–81. ROUGE: A Package for Automatic Evaluation of Summaries - ACL Anthology

[20] Smit, Akshay, Saahil Jain, Pranav Rajpurkar, Anuj Pareek, Andrew Y. Ng, and Matthew P. Lungren. "CheXbert: Combining Automatic Labelers and Expert Annotations for Accurate Radiology Report Labeling Using BERT." Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020. Combining Automatic Labelers and Expert Annotations for Accurate Radiology Report Labeling Using BERT - ACL Anthology

[21] Jain, Saahil, Ashwin Kalyan, Abhinav Rajpurkar, Pranav Rajpurkar, and Matthew P. Lungren. "RadGraph: Extracting Clinical Entities and Relations from Radiology Reports." Proceedings of the 2021.

Нярайн амьсгалын тоог дулааны зураглалаас тасралтгүй хэмжих машин сургалтад суурилсан арга

1st Алтансүхийн Жавхлан
Электроник, Холбооны
Инженерчлэлийн Тэнхим
МУИС-МТЭС
Улаанбаатар, Монгол
javkhlan.a@num.edu.mn

2nd Өнөрсайханы Батбаяр†
Электроник, Холбооны
Инженерчлэлийн Тэнхим
МУИС-МТЭС
Улаанбаатар, Монгол
batbayar.unursaikhan@num.edu.mn

3rd Бямбацогтын Гэрэлмаа
Электроник, Холбооны
Инженерчлэлийн Тэнхим
МУИС-МТЭС
Улаанбаатар, Монгол
gerelmaa.b@num.edu.mn

4th Дагданпүрэвийн Сумъяаханд
Электроник, Холбооны
Инженерчлэлийн Тэнхим
МУИС-МТЭС
Улаанбаатар, Монгол
sumiyakhand@num.edu.mn

5th Бямбаагийн Мөнхтулга
Электроник, Холбооны
Инженерчлэлийн Тэнхим
МУИС-МТЭС
Улаанбаатар, Монгол
munkhtulga@num.edu.mn

6th Чоймаагийн Лодойравсал
Электроник, Холбооны
Инженерчлэлийн Тэнхим
МУИС-МТЭС
Улаанбаатар, Монгол
lodoiravsal@num.edu.mn

***Хураангуй*—Нярайн эрчимт эмчилгээний тасаг (NICU)-т амьсгалын тоог (RR) тасралтгүй хянах нь эмнэлзүйн доройтлыг илрүүлэх болон нярайн эндэгдлийг бууруулахад нэн чухал ач холбогдолтой. Амин үзүүлэлт хэмжих уламжлалт аргууд нь нярайн эмзэг арьсыг цочроох, халдварын эрсдэл үүсгэх, мөн хөдөлгөөнийг хязгаарлах сул талтай байдаг. Энэхүү судалгаанд эрчимт эмчилгээний тасгийн нярайн амьсгалын тоог хүрэлцэхгүйгээр, тасралтгүй хэмжих дулааны зураглалд суурилсан машин сургалтын аргыг танилцуулж байна.**

Энэ систем нь нярайгаас 1 метр зайнаас 320×240 пикселээр, секундэд 30 фреймийн хурдтай авсан дулааны дүрсийг ашигласан бөгөөд хамар орчмыг сонирхлын бүс (ROI) болгон тодорхойлсон.

Сонирхлын бүсийг автоматаар тодорхойлохын тулд YOLO алгоритмд суурилсан машин сургалтын аргыг боловсруулан, ашигласан. Нярайн дулааны зургийг Twilight Shifted алгоритм ашиглан өнгөт зураг руу хувирган, хөдөлгөөнөөс болон камерын чичиргээнээс үүсэх шуугианыг Гауссын шүүлтүүрээр шүүж, шуугианы нөлөөг багасгасан. Машин сургалтаар илрүүлсэн сонирхлын бүсэд Channel and Spatial Reliability Tracker (CSRT) тракинг алгоритм ашиглав. Богино хугацааны амьсгалын динамикийг харахын тулд оролтын бичлэгийг 10 секундийн сегментүүдэд хуваах, амьсгалын дохиод оргил илрүүлэх алгоритм хэрэгжүүлэн амьсгалын тоог breath per minute (bpm) нэгжээр илэрхийлсэн.

Клиникийн баталгаажуулалтыг “Монгол Улсын Эх, Нярай, Эмэгтэйчүүдийн төв”-д 4–28 хоногтой 11 нярайд хийсэн. Нэг нярайд 60 хэмжилт хийснээр нийт 660 хэмжилтийн өгөгдөл цуглуулсан. Амьсгалын жишиг утгыг тогтоохдоо синхончлогдсон өнгөт бичлэгт EVM аргыг ашиглан хэвлийн хөдөлгөөнийг өсгөж, нүдэн баримжаагаар тооолсон.

Сонирхлын бүс илрүүлэх машин сургалтын загвар 0.995 recall болон 0.995 mAP үзүүлэлттэйгээр маш өндөр гүйцэтгэлтэй ажилласан. Энэ системийн хэмжсэн амьсгалын тоо нь жишиг утгатай өндөр хамаарал (r = 0.9, p < 0.001) үзүүлсэн бөгөөд 95%-ийн тохирлын хязгаар нь –5.62-аас 6.92 bpm хооронд байсан. Эдгээр үр дүн нь энэ системийг нярайн эрчимт эмчилгээний тасагт амьсгалын тоог тасралтгүй, хүрэлцэхгүй аргаар хянах боломжтой шийдэл болохыг харуулж байна.

Түлхүүр үг—Нярайн эрчимт эмчилгээний тасаг, Амьсгалын тоо хэмжих, Хүрэлцэхгүйгээр хэмжих, Дулааны зураглал, Нярайн амин үзүүлэлтүүд

I. ОРШИЛ

Дэлхий даяар жилд төрөх нярайн тоо нэмэгдэж, нярайн амин үзүүлэлтүүдийг тасралтгүй хянах найдвартай системийн хэрэгцээ улам нэмэгдэж байна.

“Монгол Улсын Эрүүл Мэндийн Хөгжлийн Үндэсний Төв”-ийн мэдээгээр 2023 онд Монгол улсад 793 нярайн эндэгдэл бүртгэгдэж, нярайн эндэгдлийн түвшин 1,000 амьд төрөлтөнд 12.3 байна [1].

Монгол Улсад нярай болон тав хүртэлх насны хүүхдүүдийн өвчлөлийн тэргүүлэх шалтгаан нь (өвчлөлийн 60 орчим хувь) амьсгалын замын өвчин байна. Үүнээс уушгины хатгалгаа 75.5%, цочмог бронхит 15.2%, цочмог ларингит/трахеит 2.0%-ийг тус тус эзэлж байна. Эдгээр өвчний үед амьсгалын тооны өөрчлөлт хамгийн түрүүнд илэрдэг тул нярайн амьсгалын тоог тасралтгүй хянах нь эмгэгийг эрт илрүүлэх, эмчилгээний үр дүнг сайжруулахад чухал ач холбогдолтой.

Нярайн дутуу төрөлт нь эрчимт эмчилгээний тасгийн ачаалалд ихээхэн нөлөөлдөг. 2020 онд дэлхий даяар ойролцоогоор 13.4 сая нярай (амьд төрөлтийн 9.9%) жирэмсний 37 долоо хоногоос өмнө төрсөн байна [2][3]. Sharma-ийн судалгаагаар эрчимт эмчилгээний тасагт хэвтэн эмчлүүлж буй дутуу төрсөн нярайн 86.45% нь амьсгалын тогтолцооны эмгэгтэй байна [4]. Дээрх баримтууд нь нярайн амьсгалын тоог тасралтгүй, хүрэлцэхгүй хянах шаардлага өндөр байгааг харуулж байна.

Хүний биеийн амин үзүүлэлтүүд болох амьсгалын тоо, зүрхний цохилт, биеийн температурыг хэмжих

уламжлалт аргууд нь хүрэлцэх мэдрэгч, электрод ашигладаг. Мэдрэгч болон электрод нь арьс цочроох, халдварын эрсдэл нэмэгдүүлэх, мөн нярайн хөдөлгөөнийг хязгаарлах сул талтай [5]. Дүрс боловсруулалтад суурилсан хүрэлцэхгүйгээр амин үзүүлэлт хэмжих арга нь нярайг тасралтгүй хянах боломжтой бөгөөд илүү аюулгүй, нярайн хөдөлгөөнд саад учруулахгүй давуу талтай. Түүнчлэн дулааны зураглал ашиглах үед орчны гэрэлтүүлгээс үл хамааран хэмжилт хийх боломж бүрддэг.

Энэхүү судалгаагаар нярайн амьсгалын тоог дулааны зураглалаас тасралтгүй хэмжих машин сургалтад суурилсан аргыг дэвшүүлж байна. Уг систем нь нярайн хамар, ам орчмын температурын динамик өөрчлөлтийг авч, нярайн амьсгалын тоог тасралтгүй, бодит хугацаанд хэмжих боломжийг бүрдүүлнэ. Хүрэлцэхгүйгээр хэмжилт хийх тул эрчимт эмчилгээний тасаг дахь нярайн тав тухыг алдагдуулахгүй, халдварын эрсдэлийг бууруулсан, илүү аюулгүй, найдвартай хяналтын орчныг бүрдүүлсэн.

А. Судлагдсан байдал

Нярайн амьсгалын тоог хянах нь эрүүл мэндийн байдлыг үнэлэх, эмнэлзүйн хүндрэлийг эрт илрүүлэхэд чухал ач холбогдолтой. Уламжлалт мэдрэгчид суурилсан аргууд өргөн хэрэглэгдэж байгаа ч дутуу төрсөн нярайн арьсыг цочроох, халдварын эрсдэл үүсгэх, хөдөлгөөн хязгаарлах зэрэг хүндрэл үүсгэдэг [5].

Эдгээр хязгаарлалтыг шийдэх зорилгоор судлаачид дүрс боловсруулалтад суурилсан хүрэлцэхгүйгээр хэмжилт хийх аргуудыг идэвхтэй судалж байна [6-8]. Ихэнх судалгаа хамар орчмын агаарын температурын өөрчлөлтийг илрүүлж, амьсгалын тоог тооцоолдог. Жишээлбэл, Maurya өнгөт болон дулааны дүрслэлийг хослуулсан аргыг санал болгож, субъектийн хөдөлгөөн их байсан ч найдвартай тооцоолол хийсэн [9-10]. Энэхүү арга нь эмзэг эсвэл идэвхтэй хөдөлгөөнтэй нярайд тохиромжтой. Гэвч сонирхлын бүсийг бэлэн машин сургалтын модел “Yolo5Face” ашиглан тодорхойлдог. Бэлэн моделиуд нь насанд хүрсэн хүний өгөгдөл дээр суралцсан учир нярайн биеийн бүтцээс (мөч богино, толгойн хэмжээ цээжтэй харьцуулахад том г.м) шалтгаалан сонирхлын бүсийг алдаж таамаглах магадлалтай.

Харин Pereira зөвхөн дулааны зураглалд суурилсан аргыг танилцуулж, хамар хэсгийн температурын өөрчлөлтийг хянах замаар дутуу нярайд дунджаар 3.42%-ийн алдаатай амьсгалын тоог тооцоолсон нь энэ аргыг клиник хэрэглээнд нэвтрүүлэх боломжтойг харуулсан [11]. Гэвч энэ систем нь сонирхлын бүсийг гараар тодорхойлдог бөгөөд тракинг алгоритм алдсан үед гараар дахин тохируулахаас нааш систем ажиллагаагүй болдог.

Хэдийгээр дээрх судалгааны үр дүн хүлээн зөвшөөрөхүйц түвшинд хүрсэн боловч одоогийн аргууд нь ихэвчлэн RGB болон дулааны дүрслэлийг хослуулан ашигладаг тул системийн бүтэц төвөгтэй болж, тоног төхөөрөмжийн шаардлага нэмэгддэг. Өнгөт зураглалаас субъектийн царайг таних боломжтой тул хувь хүний нууцлалтай холбоотой асуудал үүснэ. Энэ нь эмнэлзүйн орчинд онцгой анхаарах шаардлагатай асуудал юм. Мөн эдгээр судалгааг урьдчилан бэлтгэсэн, хяналттай лабораторийн орчинд туршсан байгаа тул бодит эрчимт эмчилгээний тасгийн нөхцөлд уг системийг шууд хэрэгжүүлэх боломжтой эсэх нь тодорхой бус, эргэлзээтэй хэвээр байна.

Бидний санал болгож буй арга нь зөвхөн дулааны дүрслэл ашигладаг бөгөөд энэ нь өвчтөний нууцлалыг хадгалахын зэрэгцээ техник хангамжийг энгийн болгодог. Хөдөлгөөнөөс болон камерын чичиргээнээс үүсэх шуугианыг бууруулах зорилгоор Гауссын шүүлтүүр ашиглаж, сонирхлын бүсэд CSRT тракинг алгоритм ашигласан. Бичлэгийг 10 секундийн интервал болгон хуваасан. Энэ нь амьсгалын тоог хэмжих стандарт арга биш бөгөөд дулааны дохиог боловсруулахад ашигласан дохио боловсруулах интервал юм. Өмнөх дүрс боловсруулалтад суурилсан нярайн амьсгалын тоог хэмжих судалгаануудад 5, 10, 20 секундийн интервалыг хэрэглэж ирсэн [12-13]. Иймд 10 секундийг хугацааны нарийвчлал болон амьсгалын тооны өөрчлөлтийг илрүүлэх боломжтой интервал гэж сонгосон.

Системийг бодит эрчимт эмчилгээний тасагт 11 нярайд клиник туршилтаар үнэлж, нярайд хүрэлцэхгүйгээр, тэдний хөдөлгөөнд саад учруулахгүйгээр амьсгалын тоог тооцоолсон.

II. АРГА ЗҮЙ

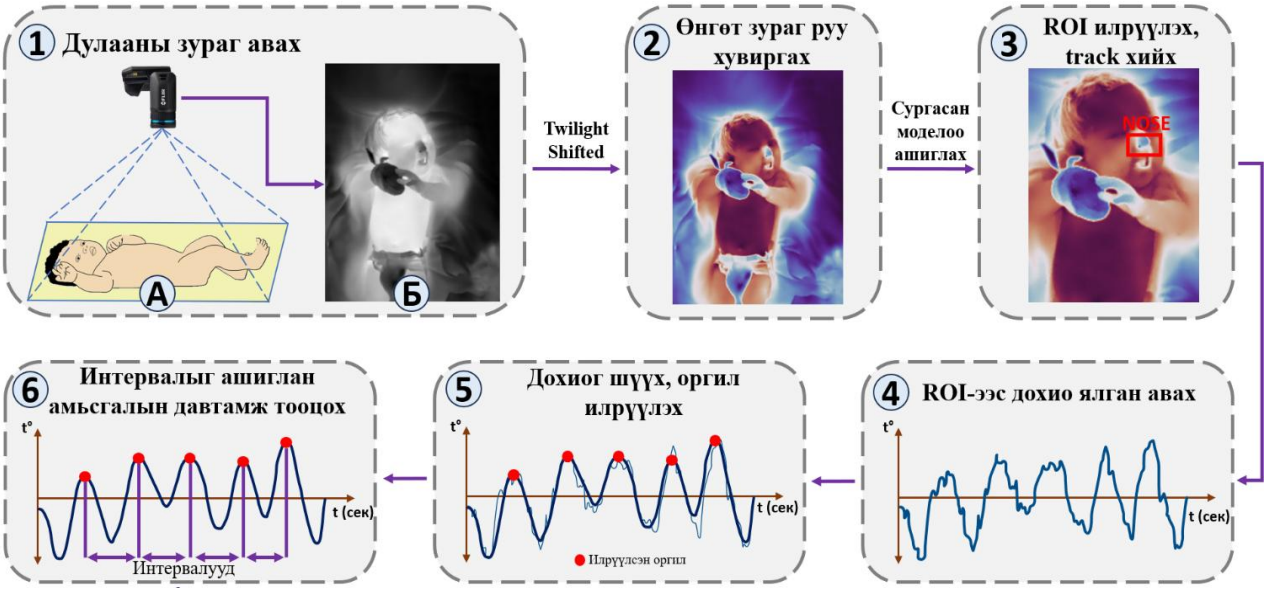
А. Ажиллах зарчим

Системийн ажиллах зарчмыг дулааны дүрс авах үеэс эхлэн амьсгалын тоог тооцоолох хүртэлх Зураг 1-д үзүүлэв. Дулааны дүрсийг 320×240 пиксел нягтралтай, секундэд 30 фреймийн хурдтайгаар авсан. Өгөгдөлд пайтон программчлалын хэл ашиглан дараах боловсруулалтыг хийв.

Сонирхлын бүс илрүүлэх болон тракинг алгоритмын гүйцэтгэлийг сайжруулах зорилгоор 32-битийн саарал өнгийн, дулааны зургийг twilight shifted алгоритмыг ашиглан өнгөт зураглал руу хувиргасан бөгөөд ингэснээр өнгөний болон ирмэгийн ялгарах байдал жигд нэмэгдсэн.

Бичлэг бүрийн хамгийн эхний фреймд машин сургалтын загвар ашиглан сонирхлын бүсийг автоматаар илрүүлсэн. Дараа нь фрейм тус бүрийн сонирхлын бүс дэх пикселүүдийн дундаж температурын утгыг тооцоолж, хугацааны дараалсан (time-series) дохио үүсгэсэн.

Нярайн хөдөлгөөн болон камерын чичиргээнээс үүсэх шуугианыг бууруулахын тулд Гауссын шүүлтүүр хэрэглэсэн. Шүүсэн дохионы оргилыг илрүүлж, оргилуудын хоорондох хугацааны интервалуудыг ашиглан амьсгалын тоо (bpm)-г тооцоолсон. Энэхүү судалгаанд амьсгалын тоог гар аргаар тоолж анхан шатны жишиг утгыг тогтоосон. ДЭМБ-ын IMCI удирдамжид хүүхдийн амьсгалын тоог нэг минутын турш тоолохыг зөвлөдөг [14]. Гэвч гар аргаар тоолох нь



Зураг 1. Системийн ажиллах зарчим. (1) Дулааны зураг авах; (А) Дулааны зургийг нярайн цээжтэй перпендикуляр (90°) байрлуулан авах; (Б) 32 битийн дулааны зураг; (2) Twilight Shifted алгоритм ашиглан өнгөт зураг руу хувиргах; (3) Сонирхлын бүс буюу ROI-г машин сургалтын аргаар автоматаар илрүүлж CSRT тракинг алгоритм ашиглан мөрдөх; (4) Сонирхлын бүс дэх пикселийн дундаж утгыг хугацааны дагуу тооцоолж амьсгалын дохио үүсгэх; (5) Дохиог Гауссын шүүлтүүр ашиглан шүүх, оргил илрүүлэх; (6) Оргил хоорондын интервалыг ашиглан амьсгалын тоог тооцоолох.

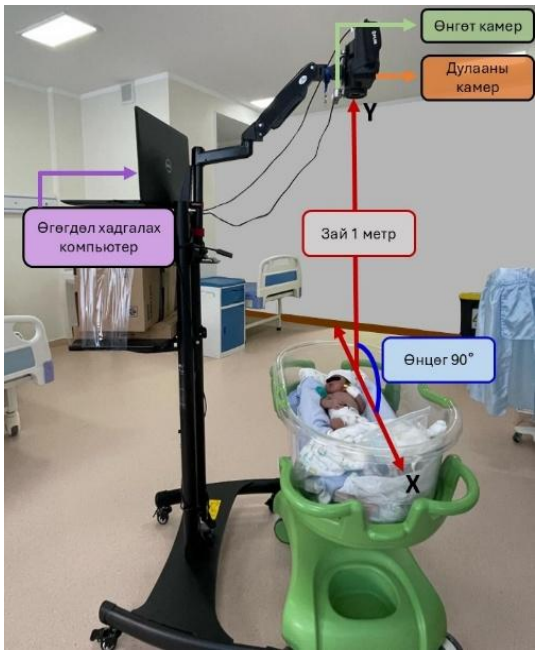
стандарт арга биш юм [12]. Ялангуяа нярайд амьсгалын хэлбэлзэл өндөр, хөдөлгөөн их, амьсгалын тооны өөрчлөлт илэрч болох тул ажиглагчийн алдаа болон ажиглагч хоорондын ялгаа үүсэх боломжтой [15-17]. Иймээс капнографи, амьсгалын индукцийн плетизмографи болон импеданс пневмографи ашиглаагүй нь энэхүү судалгааны чухал хязгаарлалт бөгөөд цаашдын судалгаанд дулааны бичлэгийг синхрон клиник жишиг төхөөрөмжтэй харьцуулах шаардлагатай [18-20].

Б. Хэмжилтийн тохиргоо

Системийн клиник туршилтыг Монгол Улсын Эх, Нярай, Эмэгтэйчүүдийн төв-д хийсэн. Судалгаанд нийт 11 нярай (4–28 хоногтой; 6 эмэгтэй, 5 эрэгтэй) оролцсон бөгөөд бүгд эмнэлэгт хэвтэн эмчлүүлж байв. Хэмжилтийн тохиргоог Зураг 2-т үзүүллээ.

Дулааны зураглалыг 1 метрийн зайд байрлуулсан FLIR T530 дулааны камераар 320×240 пиксел нягтралтай, секундэд 30 фреймийн хурдтайгаар авсан. FLIR T530 камерыг судалгааны түвшний дулааны өгөгдөл цуглуулах зорилгоор ашигласан боловч санал болгож буй арга нь зөвхөн энэ төхөөрөмжөөр хязгаарлагдахгүй. Амьсгалтай холбоотой температурын өөрчлөлт ойролцоогоор 1–3°C байсан тул FLIR Boson зэрэг (ойролцоогоор <60 mK NETD үзүүлэлттэй) илүү хямд дулааны мэдрэгч ашиглах боломжтой. [21-22] Цаашид хямд өртөгтэй төхөөрөмжүүдийн NETD, калибраци, радиометрийн гаралт, линзний харах өнцөг, секундэд бичих фреймийн тоо болон өртөг–гүйцэтгэлийн харьцуулалтыг тусгайлан үнэлэх шаардлагатай.

Богино хугацааны амьсгалын динамикийг судлах зорилгоор хэмжилт бүрийг 10 секундийн интервалд хуваасан. Нэг нярайд 10 минутын турш хэмжилт хийсэн бөгөөд нийт 660 удаа хэмжилт хийсэн. Амьсгалын тооны жишиг утгыг авахдаа синхрончлогдсон өнгөт бичлэгт Eulerian Video Magnification (EVM) ашиглан хэвлий хэсгийн хөдөлгөөнийг өсгөж, түүнээсээ нүдэн баримжаагаар тоолсон [23].

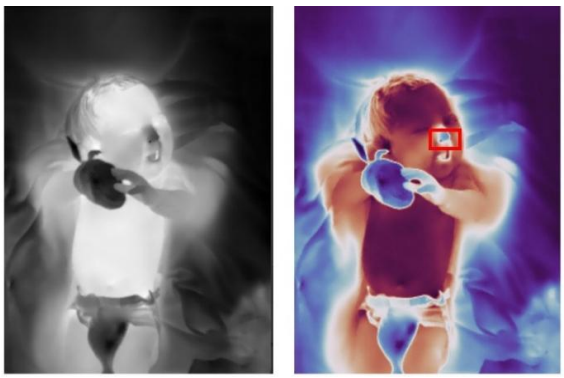


Зураг 2. Хэмжилтийн тохиргоо. Дулааны зургийг FLIR T530 камер ашиглан бичсэн (320x240 пиксел, секундэд 30 фрейм)

В. Дулааны зургийн хувиргалт

Дулааны дүрсийг 32-битийн саарал өнгийн форматтайгаар авсан (Зураг 3(А)). Дүрсний контрастийг сайжруулж, тракинг хийх нарийвчлалыг нэмэгдүүлэхийн тулд OpenCV-ийн Twilight Shifted алгоритмыг

ашигласан. Өнгөт зураг руу хувиргасан байдлыг Зураг 3(Б)-д үзүүлэв.



А Б

Зураг 3. Дулааны зургийг өнгөт зураг болгон хувиргах. (А) Дулааны камераас авсан 32 битийн саарал зураг; (Б) Twilight Shifted алгоритм ашиглан өнгөт зураглал руу хувиргаж, 25×15 пиксел хэмжээтэй тэгш өнцөгт ROI-г тодорхойлсон байдал.

Г. Сонирхлын бүс ба тракинг алгоритм

Үйлдвэрлэгчийн үзүүлэлтээр 42° линзтэй үед IFOV 2.41 mrad/pixel, 24° линзтэй үед 1.31 mrad/pixel, 14° линзтэй үед 0.75 mrad/pixel байна. Иймд 1 метрийн зайд нэг пикселд харгалзах бодит хэмжээ ойролцоогоор 2.41 мм, 1.31 мм, эсвэл 0.75 мм байх боломжтой. Энэ нь хамар/ам орчмын ROI-г 25 × 15 пикселээр хамруулах боломжтой. Мөн энэ хэмжээ нь хамрын агаарын урсгалын дулааны өөрчлөлтийг бүрэн хамруулахын зэрэгцээ ойролцоох бүсүүдийг аль болох багтаахгүй байхаар сонгосон.

Сонирхлын бүсийг илрүүлэх зорилгоор YOLO-д суурилсан объект илрүүлэх загварыг сургасан. Өгөгдлийн сан нь өнгөт зураг руу хувиргасан 1,100 ширхэг дулааны зургаас бүрдсэн бөгөөд нэг нярайд 100 зураг ноогдож байв. Өгөгдлийг санамсаргүй байдлаар 90:10 харьцаагаар сургалтын болон баталгаажуулалтын багц болгон хуваасан.

Сургалтыг YOLOv11s архитектур ашиглан Ultralytics YOLO pipeline-аар 300 эпох гүйцэтгэсэн. Google Colab орчинд NVIDIA Tesla T4 GPU ашигласан сургасан. Моделийн ажиллаж буй жишээг Зураг 4-т үзүүлэв.



Зураг 4. Машин сургалтаар нярайн хамар хэсгийг илрүүлж, сонирхлын бүсээр тодорхойлж буй байдал

Д. Дохио боловсруулалт

Камерын чичиргээ болон нярайн хөдөлгөөнөөс үүссэн шуугиантай дохиог Gaussian шүүлтүүрээр шүүсэн.

α = 5.2 гэсэн утгыг дохионы давтамжийг алдагдуулахгүйгээр өндөр давтамжийн шуугианыг хангалттай дарах оновчтой тохируулга гэж үзэн сонгосон. Өөрөөр хэлбэл, уг параметр нь амьсгалын 0.5–1.5 Hz (30–90 удаа/мин) мужид орших үндсэн хэлбэлзлийг хадгалж үлдэхийн зэрэгцээ шаардлагагүй шуугианыг дарж байгаа.

Гауссын шүүлтүүрийн стандарт хазайлт α-ийг дараах томьёогоор тооцоолсон:

$$\alpha = \frac{f_s}{2\pi f_c} \quad (1)$$

энд fs нь дээж авах давтамж (sampling frequency 30 Hz), fc нь таслах давтамж (cutoff frequency) болно.

Эрүүл нярайн амьсгалын тоо нь 30–60 bpm-ийн хооронд байдаг [24]. Харин 60 bpm-ээс дээш амьсгалын тоог тахипноэ гэж үзэх бөгөөд энэ нь амьсгалын хямралын шинж тэмдэг байж болзошгүй [25-26]. Иймд оргил илрүүлэх дээд хязгаарыг тооцоолохдоо хамгийн ихдээ 90 bpm гэж авч үзсэн.

Дараалсан оргил хоорондын хамгийн бага зайг дараах байдлаар тооцоолсон:

$$\text{Оргил хоорондын хамгийн бага зай} = \frac{fps \cdot 60}{BPM_{max}} \quad (2)$$

Энд fps нь камерын секундэд авах фреймийн тоо (30fps) BPM_{max} нь өмнө дурдсанчлан амьсгалын хамгийн их давтамж буюу 90 bpm байна.

Хуурамч илрүүлэлтийг багасгах зорилгоор оргил хоорондын хамгийн бага зайг 20 фрейм байхаар тогтоосон. Дохионы цэг бүрийг хөрш утгуудтай нь харьцуулж, зөвхөн хамгийн бага зайны шалгуурыг хангаж буй оройн утгуудыг хүчинтэй оргил гэж тодорхойлсон.

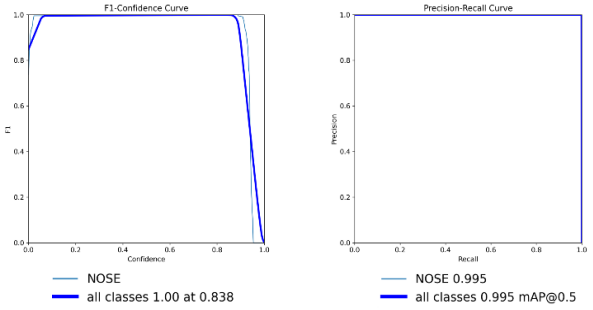
Оргил хоорондох интервалуудын дундаж утгыг ашиглан амьсгалын тоог дараах байдлаар тооцоолсон:

$$\text{Амьсгалын тоо} = \frac{60 \cdot fps}{\text{Дундаж интервал}} \quad (3)$$

III. ҮР ДҮН

Санал болгож буй системийг 11 нярайн дулааны бичлэг дээр баталгаажуулж, нийт 660 хэмжилт хийсэн. Тооцоолсон амьсгалын тоог синхрончлогдсон өнгөт бичлэгт Eulerian Video Magnification (EVM) ашиглан хэвлий хэсгийн хөдөлгөөнийг өсгөж, түүнээсээ нүдэн баримжаагаар тоолсон утгатай харьцуулсан.

YOLOv11s-д суурилсан хамрын сонирхлын бүс (ROI) илрүүлэх загвар нь нярайн дулааны дүрсэн дээр өндөр нарийвчлалтай ажилласан. Загварын үнэлгээний үр дүнд recall = 0.99, mAP@0.5 = 0.99 үзүүлэлтүүд гарсан. Мөн F1-score нь 0.838 босгон дээр 1.00-д хүрсэн нь нарийвчлал (precision) болон бүрэн илрүүлэлтийн (recall) хооронд оновчтой тэнцвэр хангагдсаныг харуулж байна.



А Б

Зураг 5. Хамрын илрүүлэлтийн Precision–Recall болон F1–confidence муруйнууд. (А) Precision–Recall муруй: хамрын ангиллын хувьд recall = 0.99, precision (mAP@0.5) = 0.99 үзүүлсэн; (Б) F1–confidence муруй: 0.84 босго утгад F1 оноо хамгийн ихдээ 1.00 хүрснийг харуулна.

Зураг 6(а)-д системээр тооцоолсон амьсгалын тоо болон жишиг утгуудын хоорондын хамаарлыг үзүүлсэн. Өгөгдлийн цэгүүд диагональ шугамын дагуу байгаа нь хэмжсэн болон жишиг утгууд хоорондоо өндөр хамааралтай байгааг харуулж байна.

Мөн Pearson корреляцийн коэффициент r = 0.90 ($p < 0.001$, n = 11) гарсан нь системийн үр дүн болон уламжлалт нүдээр ажиглан тоолох аргын хооронд хүчтэй тохирол байгааг дахин баталсан.

Монгол Улсын Эх, Нярай, Эмэгтэйчүүдийн төвд хийсэн клиник туршилт нь системийг бодит эрчимт эмчилгээний тасагт хэрэгжүүлэх боломжтойг харуулсан. Нярайгаас ойролцоогоор 1 метрийн зайд байрлуулсан дулааны камер нь биетэй огт хүрэлцэхгүй хэмжилт хийх боломж бүрдүүлж, мэдрэгчид суурилсан аргуудад нийтлэг тохиолддог арьсны цочрол, халдварын эрсдэл, хөдөлгөөний хязгаарлалт зэрэг асуудлыг бүрэн шийдсэн.

Хэдийгээр систем хүлээн зөвшөөрөхүйц үр дүн үзүүлсэн боловч дараах хязгаарлалтууд илэрсэн. Үүнд:

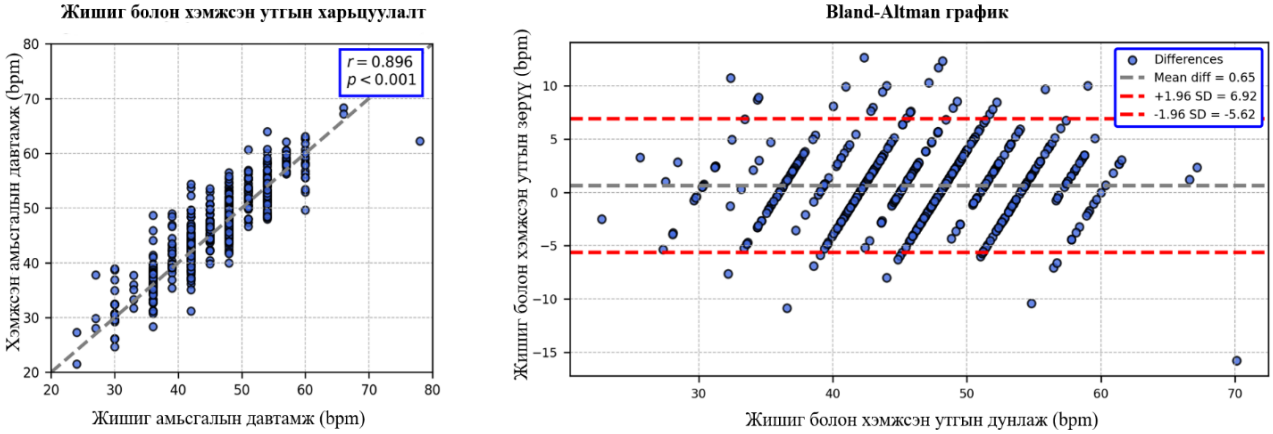
Хэмжилт хийсэн субъектийн тоо (n = 11) харьцангуй бага тул илүү өргөн хүрээний судалгаа шаардлагатай. Мөн жишиг утгыг нүдээр баримжаагаар тоолох нь бага зэрэг алдаатай байх магадлалтай.

IV. ДҮГНЭЛТ

Энэхүү судалгаагаар эрчимт эмчилгээний тасгийн нярайн амьсгалын тоог хүрэлцэхгүй, тасралтгүй хэмжих зөвхөн дулааны зураглалд суурилсан машин сургалтын аргыг аргыг танилцууллаа. Энэхүү арга нь хамрын бүсийг автоматаар илрүүлж, тракинг хийх замаар дулааны өөрчлөлтийг авч, хөдөлгөөнөөс үүсэх шуугианыг багасган амьсгалын дохиог амжилттай ялган авсан.

11 нярайд гүйцэтгэсэн хэмжилтийн үр дүнд системээр хэмжсэн амьсгалын тоо нь жишиг амьсгалын тоотой өндөр хамааралтай (r = 0.90, $p < 0.001$) байсан бөгөөд Bland–Altman шинжилгээгээр хүлээн зөвшөөрөгдөхүйц үр дүнтэй байгаа нь батлагдсан.

Хүрэлцэхгүйгээр хэмжилт хийх систем нь уламжлалт мэдрэгчид суурилсан системд тулгардаг хүндрэлүүдээс



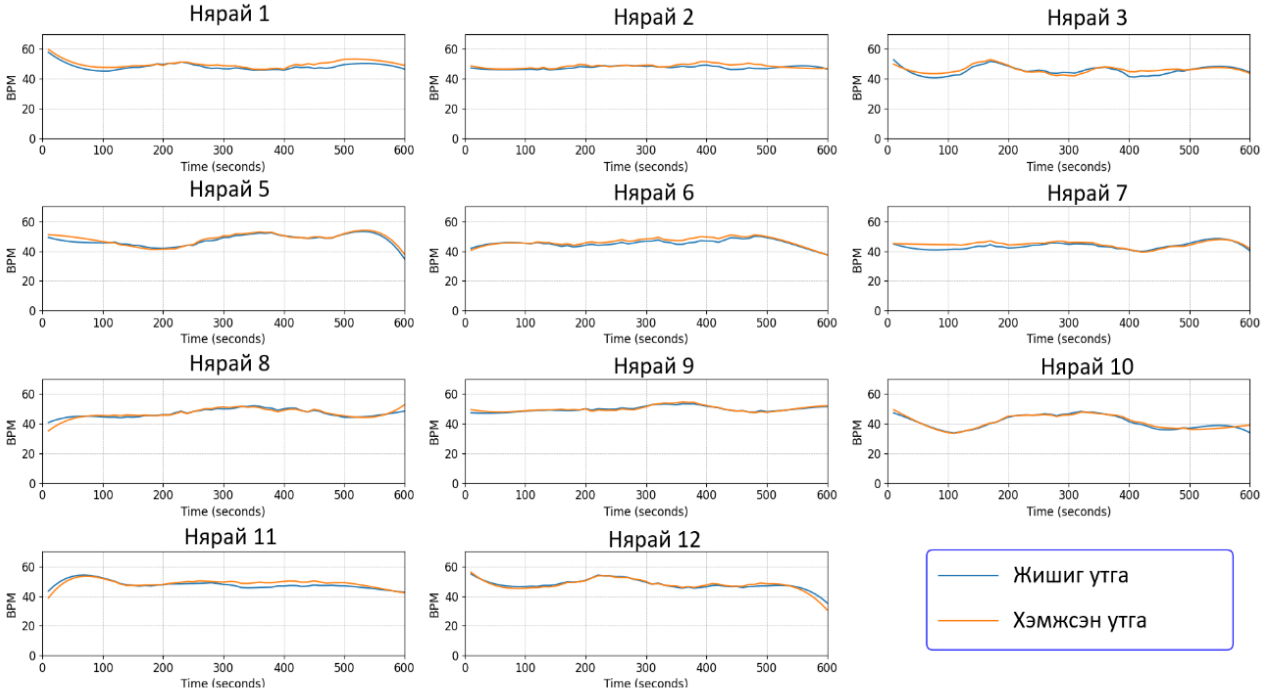
Зураг 6. Системийн баталгаажуулалт. (а) Системээр тооцоолсон амьсгалын тоо болон жишиг утгуудын харьцуулалтыг тархалтын (scatter) графикаар үзүүлсэн байдал; (b) Bland–Altman plot ашиглан хийсэн үнэлгээ.

Зураг 7-д нярай тус бүрийн амьсгалын тооны харьцуулалтыг үзүүлсэн бөгөөд хөх өнгө нь жишиг амьсгалын тоо, улбар шар өнгө нь системийн хэмжсэн амьсгалын тоог илэрхийлнэ. Хэдийгээр бага зэргийн зөрүү ажиглагдсан ч жишиг утга болон хэмжсэн утгын өсөлт бууралтууд синхрон байдлаар илэрч байна. Энэ нь абсолют утгад бага зөрүү гарч болох ч систем амьсгалын динамикийг үнэн зөв барьж чаддагийг харуулж байна.

зайлсхийх боломж олгож, өвчтөний аюулгүй байдлыг нэмэгдүүлж байна.

Судалгааны үр дүн нь манай системийг нярайн эрчимт эмчилгээний тасагт амьсгалын тоог тасралтгүй, бодит цагийн горимоор хянах ирээдүйтэй, хувилбар болохыг баталж байна.

Энэхүү судалгаа нь n=11 нярайг хамруулсан pilot feasibility study бөгөөд клиник баталгаажуулалт хийхэд хангалттай түүвэр биш. Үр дүн нь дулааны зураглалаас нярайн амьсгалтай холбоотой дохио ялган авах техникийн боломжийг харуулж байгаа боловч субъект хоорондын хэлбэлзлийг бүрэн дүгнэхэд хангалтгүй. Цаашид том түүвэртэй, зүрхний цохилт, амьсгалын тоо, биеийн температур зэрэг амин үзүүлэлтийг найдвартай хэмжих төхөөрөмж (электрокардиограмм эсвэл фото-плетизмографи, кафнограм) ашиглан жишиг утгыг синхрон авч судалгаа хийхийг зорьж байна.



Зураг 7. Нярайн амьсгалын тоог хэмжсэн утга болон жишиг утгын харьцуулалт (нярай тус бүрээр).

Номзүй

[1] Center for health development of Mongolia with WHO, “Health Indi cators 2023,” 2023.

[2] PMNCH (Partnership for Maternal, Newborn and Child Health).: “UN warns of preterm birth rates flatlining in every region,” 2023. doi: https://doi.org/10.1136/bmj.p1054

[3] World Health Organization (WHO), “Born too soon: decade of action on preterm birth,” 2023.

[4] Y. Sharma, O. K. Pathak, B. Poudel, A. Sharma, R. P. Sapkota, and K. Devkota, “Preterm Neonates Admitted in the Neonatal Intensive Care Unit at a Tertiary Care Centre: A Descriptive Cross-sectional Study,” 2023. doi: 10.31729/jnma.8126

[5] M. Broom, A. M. Dunk, and A.-L. E. Mohamed, “Predicting neonatal skin injury: The first step to reducing skin injuries in neonates,” 2019

[6] K. Abbas, K. Heimann, K. Jergus, T. Orlikowsky, and S. Leonhardt, “Measurement of respiratory rate using infrared thermography,” 2011. doi:10.1186/1475-925X-10-93

[7] C. Massaroni, A. Nicolò, M. Sacchetti and E. Schena, "Contactless Methods For Measuring Respiratory Rate: A Review," in IEEE Sensors Journal, vol. 21, no. 11, pp. 12821-12839, 1 June1, 2021, doi: 10.1109/JSEN.2020.3023486.

[8] Kim JD et al.: “Non-contactrespiration monitoring using impulse radioultrawideband radar in neonates,” 2019. doi: http://dx.doi.org/10.1098/rsos.190149

[9] L. Maurya, P. Mahapatra and D. Chawla, “Non-contact breathing moni toring by integrating RGB and thermal imaging via RGB-thermal image registration,” doi: https://doi.org/10.1016/j.bbe.2021.07.002. 202

[10] L. Maurya, R. Reyer Zwiggelaar, D. Chawla and P. Mahapatra, “Non contact respiratory rate monitoring using thermal and visible imaging: a pilot study on neonates,” 2023. doi: https://doi.org/10.1007/s10877-022 00945-8

[11] C. B. Pereira, K. Heimann, B. Venema, V. Blazek, M. Czaplik and S. Leonhardt, “Estimation of respiratory rate from thermal videos of

[12] J. Coleman et al., “Assessment of neonatal respiratory rate variability,” Archives of Disease in Childhood - Fetal and Neonatal Edition, 2022.

[13] N. Koolen et al., “Automated respiration detection from neonatal video data,” in Proc. International Conference on Health Informatics, 2015, pp. 164–169.

[14] World Health Organization, Integrated Management of Childhood Illness: Module 3 – Assess and Classify the Sick Child Age 2 Months up to 5 Years. Geneva, Switzerland: WHO, 2014.

[15] W. J. Daw, C. J. Kingshott, L. Elphick, and M. A. Elphick, “Poor inter-observer agreement in the measurement of respiratory rate in children: A prospective observational study,” BMJ Paediatrics Open, vol. 1, no. 1, 2017.

[16] N. Kallioinen, A. Hill, K. Horswill, H. E. Ward, and E. Watson, “Quantitative systematic review: Sources of inaccuracy in manually measured adult respiratory rate data,” Journal of Advanced Nursing, vol. 77, no. 1, pp. 98–124, 2021.

[17] S. Fleming et al., “Non-invasive measurement of respiratory rate in children using the photoplethysmogram,” in Proc. 30th Annual International IEEE EMBS Conference, 2008.

[18] B. B. Lahiri, S. Bagavathiappan, T. Jayakumar, and J. Philip, “Medical applications of infrared thermography: A review,” Infrared Physics & Technology, vol. 55, no. 4, pp. 221–235, 2012.

[19] Y. Zhao et al., “Non-contact infrared thermometers and thermal scanners for human body temperature monitoring: A systematic review,” Sensors, vol. 23, no. 17, 2023.

[20] I. Lorato et al., “Multi-camera infrared thermography for infant respiration monitoring,” Biomedical Optics Express, vol. 11, no. 9, pp. 4848–4861, 2020.

[21] Teledyne FLIR, “FLIR T530 Professional Thermal Camera: Technical Specifications,” Teledyne FLIR. Accessed: May 4, 2026.

[22] Teledyne FLIR, “FLIR Boson® Thermal Imaging Core Datasheet,” FLIR Systems, 2021. Accessed: May 4, 2026.

[23] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. T. Freeman,“ Eulerian Video Magnification for Revealing Subtle Changes in the World,” ACM Trans. Graph. 31(4), 65:1--65:8, 2012. doi: 10.1145/2185520.2185561

[24] D. DiMaio, “Vital signs in children” Medscape, 2023.

[25] WHO Young Infants Clinical Signs Study Group: “Clinical signs that predict severe illness in young infants aged 0–59 days,” 1999.

[26] Darmstadt G. L. et al.: Validation of simplified clinical algorithms for identifying young infants with severe illness in rural Bangladesh. (2011)



ҮЗҮҮЛЭН ИЛТГЭЛҮҮД



Хиймэл оюун болон AR технологид суурилсан анхны тусламжийн системийн судалгаа, хөгжүүлэлт

Үүрцайхын Мөнхтуяа
Мэдээллийн Технологийн тэнхим Мэдээлэл,
Холбооны Технологийн Сургууль, ШУТИС
Улаанбаатар, Монгол улс
B221930012@must.edu.mn

Нарантуягийн Ганхуяг
Мэдээллийн Технологийн тэнхим
Мэдээлэл, Холбооны Технологийн Сургууль, ШУТИС
Улаанбаатар, Монгол улс
xuygaa_mn@must.edu.mn

Хураангуй — Улаанбаатар хотын Түргэн Тусламжийн Төв 103 (НТТТ)-ийн сүүлийн 3 жилийн тайланд 103 дугаарт дуудлага өгөхөд, дуудлагыг хүлээн авснаас хойш дунджаар “28” минутын дараа тухайн байршилд очдог ба хотын хүн амд ойролцоогоор 170 орчим түргэн тусламжийн баг ногдож ажиллах шаардлагатай ч бодит байдалд 45-50 дуудлагын баг 5 бүсэд хуваагдан ажиллаж байгааг 103 төвийн эмч, ажилчид мэдээллэсэн. Энэ нь түргэн тусламжийн хүртээмжийг бууруулах, шаардлагатай анхны тусламжийг цаг алдалгүй авах боломжийг хязгаарлаж байна. Мөн ДЭМБ-ын мэдээлснээр дэлхийн хэмжээнд 2025 онд 3.1 сая хүн анхны тусламж үйлчилгээг цаг алдалгүй авч чадаагүйн улмаас амь насаа алдаж, ойролцоогоор 17.6 сая хүн насан туршийн хөдөлгөөн, хараа, сонсгол, сэтгэхүй, хэл ярианы бэрхшээлтэй болсон байна. Харин Монголд НТТТ-д жилд гэмтэл ослоос үүдэлтэйгээр эмч очихоос өмнө нас барж байгаа 1800 гаруй тохиолдол тогтмол бүртгэгдэж байна. Иймээс бид эмнэлгийн тусламж ирэх хүртэлх хугацаанд иргэдэд стандартын дагуу ойлгомжтой, найдвартай зааварчилгаа өгөх, тухайн нөхцөлд чиглүүлэх, дур мэдэн буруу алхам хийхээс сэргийлэх ухаалаг системийг хөгжүүлэхээр зорьсон болно. Систем нь ухаан алдах, их хэмжээний цус алдалт, амьсгал боогдолт, угаартах зэрэг нийт 26 төрлийн эмнэлзүйн тохиолдол дээр 252 бодит түргэн тусламжийн дуудлагын хүрээнд цуглуулсан өгөгдлөөр 4 үндсэн үе шаттайгаар хөгжүүлсэн. Өгөгдлийн багц нь амь тэнссэн, яаралтай, яаралтай бус гэсэн гурван хэлбэрээр ангилагдсан ба анхны болон амилуулах суурь тусламжийн эмнэлзүйн удирдамж, олон улсын өвчлөлийн ангиллаас бүрдэнэ. Систем нь нэгдүгээрт, хэрэглэгчээс аудио (Whisper ASR), дүрс (MediaPipe), текстэн хэлбэрээр нөхцөл байдлын мэдээллийг хүлээн авч боловсруулна. Хоёрдугаарт, VLM, RAG технологид суурилан анхны тусламжийн зааварчилгааг боловсруулна. Гуравдугаарт, боловсруулсан зааврыг дуу хоолой, текст, зураг болон AR технологиор дамжуулан хүргэнэ. Дөрөвдүгээрт, хэрэглэгч “CPR” зааварчилгааг алхамчилсан байдлаар зөв гүйцэтгэж буй эсэхийг бодит цаг хугацаанд хянах механизмуудыг агуулна. Хөгжүүлэлтийн үр дүнд нийт системийн нарийвчлал, гүйцэтгэлийг эмнэлзүйн тохиолдол, модуль бүрийн түвшинд тусгай метрикүүдээр хэмжин системийн ерөнхий гүйцэтгэл 96.5%, системийн найдвартай байдал 92% гарсан.

Түлхүүр үг— Хиймэл оюун, Өргөтгөсөн бодит байдал (Augmented Reality, AR), Whisper-ASR, MediaPipe, VLM-RAG, Анхны болон Амилуулах суурь тусламж

Компьютерын харааnd суурилсан объект алдагдлыг илрүүлэх судалгаа

Сүрэнжавын Азжаргал
Компьютерын ухааны тэнхим
ШУТИС, Мэдээлэл холбоо технологийн сургууль
Улаанбаатар хот
B232270052@must.edu.mn

Очирын Оюунжаргал
Компьютерын ухааны тэнхим
ШУТИС, Мэдээлэл холбоо технологийн сургууль
Улаанбаатар хот
B232270019@must.edu.mn

Бямбадоржийн Золзаяа
Мэдээллийн технологийн тэнхим
ШУТИС, Мэдээлэл холбоо технологийн сургууль
Улаанбаатар хот
b.zolzaya@must.edu.mn

Хураангуй — Сүүлийн жилүүдэд хулгайн гэмт хэргийн тоо нэмэгдэж байгаа ч сургууль, хувийн байгууллагуудын хяналтын систем нь ихэвчлэн камерын бичлэг хадгалах түвшинд хязгаарлагдаж, зөрчил гарсны дараа шалгах байдлаар ашиглагддаг сул талтай. Энэ нь цаг хугацаа алдах, санхүүгийн эрсдэлийг нэмэгдүүлэх шалтгаан болдог. Иймээс хяналт, аюулгүй байдлыг сайжруулах шаардлага тулгарч байна. Энэхүү судалгааны зорилго нь хиймэл оюунд суурилсан лабораторийн хяналтын систем боловсруулж, тоног төхөөрөмжийн зөвшөөрөлгүй алдагдлаас урьдчилан сэргийлэхэд оршино. Лабораторийн объектыг илрүүлэх загварын хувьд Background subtraction болон Объект илрүүлэлтийн аргуудыг харьцуулан үнэлэв. Туршилтын үр дүнгээс харахад background subtraction арга нь гэрлийн өөрчлөлтөд мэдрэмтгий бөгөөд объектын худал илрүүлэлт ихтэй байсан. Харин YOLOv8 загвар нь P(91%), Recall(90%), F1(90.5%)-ийн нарийвчлалтай ажиллаж, бодит цагийн нөхцөлд илүү үр дүнтэй илрүүлэлт хийж байв. Мөн илрүүлэлтийн гүйцэтгэлийг сайжруулах зорилгоор SAHI (Slicing Aided Hyper Inference) болон ESPCN (Efficient Sub-pixel Convolutional Neural Network) аргуудыг нэмэлтээр хэрэгжүүлсэн. Туршилтын үр дүнгээс харахад SAHI (96.45%) болон ESPCN (94.17%) аргуудыг ашигласан тохиолдолд суурь хувилбартай харьцуулахад илүү өндөр үзүүлэлттэй гарсан. Мөн лабораторийн тоног төхөөрөмжийн тоо зорилтот хэмжээнд хүрэхгүй тохиолдолд систем нь хариуцагч ажилтанд мессеж, имэйл, мобайл аппликэйшнээр дамжуулан автоматаар мэдэгдэл илгээдэг. Энэхүү судалгаа нь хяналт шаардлагатай ямар ч орчинд хямд зардлаар хяналт, аюулгүй байдлыг сайжруулж, яаралтай арга хэмжээ авах боломжийг бүрдүүлэх ач холбогдолтой юм.

Түлхүүр үг — Background subtraction, SAHI, ESPCN, Объект илрүүлэлт, Мэдэгдэл

Нүүрсний үнийн богино хугацааны таамаглалд decomposition-based hybrid болон ensemble загваруудын харьцуулалт

Д.Дорждэрэм
Эдийн засгийн тэнхим
МУИС, НУС
Улаанбаатар хот, Монгол улс
24M1NUM0395@stud.num.edu.mn

М.Золжаргал
Мэдээлэл, компьютерийн ухааны тэнхим
МУИС, МТЭС
Улаанбаатар хот, Монгол улс
zoljargal@num.edu.mn

Хураангуй – Энэхүү судалгааны зорилго нь Таншаны таамаглалын нарийвчлалыг сайжруулж буй ч эдгээр нүүрсний үнийн 2017–2025 оны хугацааны цуваанд (2017– аргуудыг ижил нөхцөлд системтэйгээр харьцуулсан 2024 сургалт, 2025 тест) тулгуурлан уламжлалт статистик, судалгаа хомс хэвээр байна. Гүн сургалтын дарааллын загвар болон decomposition-based Хэдийгээр нүүрсний үнийн хэлбэлзэл нь олон төрлийн hybrid, ensemble аргачлалуудын таамаглалын гүйцэтгэлийг гадаад хүчин зүйлсийн нөлөөнд оршдог боловч энэхүү харьцуулан үнэлэхэд оршино. судалгаа нь зөвхөн үнийн хугацааны цуваанд тулгуурласан Үнийн өгөгдлийг VMD болон CEEMDAN аргаар задлан загварчлалд төвлөрч, decomposition-based hybrid болон бүрэлдэхүүн хэсгүүдэд хуваж, тус бүрд нь LSTM, GRU, ensemble аргачлалуудын гүйцэтгэлийг харьцуулан үнэлэхэд CNN-LSTM, DeepTCN загваруудыг ашиглан hybrid чиглэж байна. Иймээс нүүрсний үнийн олон давтамжтай, загваруудыг байгуулж, эцсийн прогнозыг Stacking Ensemble шугаман бус динамикийг илүү нарийвчлалтай загварчлах аргаар нэгтгэн гаргав. зорилгоор эдгээр хандлагуудыг нэг хүрээнд авч үзсэн нь Загваруудын гүйцэтгэлийг MAE, RMSE, MAPE болон R^2 онолын болон практик ач холбогдолтой юм. үзүүлэлтээр үнэлэхэд decomposition-based hybrid болон ensemble хандлагууд нь нүүрсний үнийн олон давтамжтай, II. ИЖИЛ ТӨСТЭЙ СУДАЛГАА шугаман бус хэлбэлзлийг илүү нарийвчлалтай тусгаж, дан гүн сургалтын болон уламжлалт статистик загваруудаас[1] нь нүүрсний үнийн олон алхамт илүү өндөр гүйцэтгэл үзүүлсэн. таамаглалд LSTM болон DNN-ийн хосолсон загвар хэрхэн Ялангуяа Stacking Ensemble загвар нь хамгийн бага алдаа ажилладагийг туршсан. Судалгаанд олон улсын дулааны (RMSE=9.98, MAE=7.23, MAPE=1.80%) болон хамгийн нүүрсний (API2, Newcastle) 10–15 жилийн сарын үнийн өндөр тайлбарлах чадвар (R^2=0.83)-ыг үзүүлж, LSTM өгөгдлийг ашиглаж, LSTM–DNN загварыг MLP, SVM болон загвартай харьцуулахад RMSE-ийг ойролцоогоор 20%-иар уламжлалт ARIMA төрлийн загваруудтай харьцуулжээ. Үр бууруулсан байна. дүнгээр гүн сургалтын дарааллын загварууд (ялангуяа Судалгааны үр дүн нь нүүрсний үнийн таамаглалд LSTM) нь RMSE, MAE зэрэг гол хэмжүүрүүдээр статистик decomposition болон ensemble аргачлалуудыг хослуулсан загваруудаас илүү гүйцэтгэлтэй байж, нүүрсний үнийн загварууд илүү нарийвчлалтай, тогтвортой таамаглал гаргах уртын хамаарлыг илүү сайн барьж чаддагийг баталсан. Энэ боломжтойг харуулж байна. судалгаа нь нүүрсний үнийн шугаман бус, олон давтамжтай

Түлхүүр үг— нүүрсний үнэ, хугацааны цуваа, цуваанд deep learning үр ашигтай гэдгийг тодорхой машин сургалт, decomposition, LSTM, CNN-LSTM, харуулсан анхны судалгааны нэг юм.
CEEMDAN, Stacking Ensemble, үнийн таамаглал.[2]

[1] (Zakaria Alameer, 2020)
[2] (Kefei Zhang, 2022) нүүрсний үнийг Variational Mode

Эртний Рүни-Монгол бичээсийг орчин цагт “амилуулах” нь: Дүрсэд суурилсан систем

Соном-Очирын Өлзийбаяр[1],
Мэдээллийн технологийн тэнхим
ШУТИС, МХТСургууль
Монгол улс, Улаанбаатар хот
ulziibayar@must.edu.mn

Бямбадоржийн Золзаяа[2]
Мэдээллийн технологийн тэнхим
ШУТИС, МХТСургууль
Монгол улс, Улаанбаатар хот
b.zolzaya@must.edu.mn

Юндэнбатын Болдбаатар[3]
Түүх соёл, аялал жуулчлалын тэнхим
ШУТИС, НХСургууль
Монгол улс, Улаанбаатар хот
bold@must.edu.mn

Хураангуй—Төв Азийн түүх, соёл, хэл шинжлэлийн судалгаанд эртний рүни бичгийн дурсгалууд онцгой байр суурь эзэлдэг. Түрэг болон эртний монгол угсаатны соёл, төрийн байгуулал, нийгмийн харилцаа, хэлний хөгжлийн талаархи чухал мэдээллийг хадгалсан эдгээр хадны бичээс, дурсгалууд нь өнөөдөр түрэг судлал, монгол судлал, түүхсоёлын судалгааны үндсэн эх сурвалжийн нэг юм. Гэвч рүни бичгийн уншлага, тайлал нь ихэвчлэн мэргэжлийн цөөн судлаачдын гар ажиллагаанд тулгуурладаг бөгөөд олон дурсгалын бичээсүүд бүрэн судлагдаагүй, эсвэл судалгаанд ашиглахад хүндрэлтэй хэвээр байна. Энэхүү судалгааны зорилго нь эртний рүни-монгол бичгээр бичигдсэн хадны бичээс, дурсгалын зургийг оролт болгон авч, тэдгээрийг орчин үеийн кирилл монгол хэл рүү автоматаар хөрвүүлэх дижитал системийн загварыг боловсруулахад оршино. Судалгаанд дүрсийн боловсруулалт, OCR (Optical Character Recognition), дараалал-дараалал (seq2seq) нейрон загварууд болон хэлний том загвар (LLM)-д суурилсан контекстийн засвар, тайлбарын шаталсан архитектурыг ашиглахыг санал болгож байна.

Санал болгож буй систем нь гурван үндсэн шаттай. Нэгдүгээр шатанд хадны бичээс болон дурсгалын зургаас рүни үсгийг илрүүлж таних дүрсийн боловсруулалтын алгоритм болон OCR загварыг ашиглана. Хоёрдугаар шатанд танигдсан тэмдэгтүүдийг анхан шатны галиглалд хөрвүүлэх seq2seq загвар хэрэгжинэ. Гуравдугаар шатанд хэлний том загварт суурилсан контекстийн засварын механизм ашиглан тухайн бичээсийн цаг үеийн хэлзүйн онцлог, түүхэн утга агуулгыг харгалзан орчин үеийн монгол хэл рүү илүү оновчтой хөрвүүлэлт, тайлбар хийхийг зорьж байна. Мөн судалгаанд хэд хэдэн seq2seq архитектуруудыг туршин харьцуулж, рүни бичгийн галиглал болон орчуулгын гүйцэтгэлийг үнэлэх туршилтын орчныг төлөвлөсөн. Энэхүү судалгаа нь эртний бичгийн дурсгалуудыг дижитал хэлбэрт хадгалах, судалгаанд ашиглах боломжийг өргөжүүлэхээс гадна түрэг-монголын түүх, хэл шинжлэл, соёлын өвийг орчин үеийн хиймэл оюуны технологитой холбох шинэ боломжийг нээх ач холбогдолтой юм.

Түлхүүр үг— Түрэг-монгол судлал, Рүни бичээс, Оптик тэмдэгт таних (OCR), Дараалал-дараалал (Seq2Seq), Хэлний том загвар (LLM)

Comparative Analysis of Building Video Intercom Systems in New Residential Buildings in Mongolia

Bayarmunkh, Bulten
Department of Security Studies
Mongolian National Defense University
Mongolia, Ulaanbaatar
bultenbayarmunkh@gmail.com

Bayarkhuu, Purevsuren
Department of Security Studies
Mongolian National Defense University
Mongolia, Ulaanbaatar
p.bayarkhuu0305@gmail.com

Abstract—In recent years, the construction of new buildings in Mongolia has increased significantly, leading to the widespread adoption of various international video intercom systems. The selection of an appropriate intercom system plays a crucial role in ensuring building security and mitigating potential risks. This study examines the general classification of intercom systems and identifies the most commonly used systems in Mongolia based on a survey of construction companies. Furthermore, a comparative analysis of these systems is conducted to evaluate their technical characteristics and suitability. The findings provide insights into current usage trends and support recommendations for improving intercom system selection and implementation in Mongolian residential buildings.

Keywords— Video Intercom Systems, Building Security, Mongolia, Residential buildings, Comparative Analysis

CNN, MobileNetV2, CRNN+CTC архитектурын харьцуулалт

Эрдэнэбатын Баянбат
МКУТ,МТЭС
Монгол Улсын Их Сургууль
Монгол Улс, Улаанбаатар хот
buddybayanbat@gmail.com

Цогтбаатарын Энхтуяа
МКУТ,МТЭС
Монгол Улсын Их Сургууль
Монгол Улс, Улаанбаатар хот
enkhtuya.ts@num.edu.mn

Хураангуй—CAPTCHA нь вэб системийг автоматжуулсан халдлагаас хамгаалах өргөн хэрэглэгддэг механизм юм. Гүн сургалтын аргуудын хөгжил нь текстэн CAPTCHA-ийн найдвартай байдалд ноцтой сорилт болж байна. Энэхүү ажилд бид өөрсдийн боловсруулсан CNN загварын зэрэгцээ ImageNet дээр урьдчилан сургасан MobileNetV2-ийн хоёр хувилбар, мөн CRNN+CTC дарааллын загвар гэсэн нийт таван сургалтын схемийг ижил өгөгдөл, ижил тооцоолох орчинд (CPU) сургаж харьцуулсан. Үндсэн үр дүнд: (i) 772 ширхэг жинхэнэ Kaggle зургаар сургасан CNN бүтэн CAPTCHAийн түвшинд 46.58% нарийвчлалд хүрэв; (ii) 5,000 зохиомол CAPTCHA нэмж сургахад жинхэнэ тестэн дээрх нарийвчлал 46.58%-аас 29.81% болж буурсан хэдий ч гажуудалтай тестэн дээрх тэмдэгтийн нарийвчлал 3.24%-аас 28.44% болж 8.8 дахин нэмэгдсэн; (iii) MobileNetV2-ийн хоёр хувилбар болон CRNN+CTC хоёулаа бараг суралцаж чадаагүй. Эдгээр үр дүн нь текстэн CAPTCHA-ийн эмзэг байдал нь нарийн архитектураас үл хамааран, хэн ч хийж болох энгийн CNN сургалтад нээлттэй байгааг харуулж байна.

Түлхүүр үг—CAPTCHA, конволюцын мэдрэлийн сүлжээ, MobileNetV2, CRNN+CTC, гүн сургалт, аюулгүй байдал, зохиомол өгөгдөл, шилжүүлэх сургалт.

Нөөц багатай нөхцөлд монгол-англи холимог бичвэрийн мэдрэмжийн шинжилгээний аргачлал

Ариунболдын Цолмонбаяр
Мэдээллийн технологи
ШУТИС, МХТС
Монгол улс, Улаанбаатар
tsolmonbayar122@gmail.com

Бямбадоржийн Золзаяа
Мэдээллийн технологи
ШУТИС, МХТС
Монгол улс, Улаанбаатар
b.zolzaya@must.edu.mn

Хураангуй— Монгол хэл нь эх хэлний боловсруулалт (Natural Language Processing, NLP)-ын судалгаанд нөөц багатай хэлд тооцогддог бөгөөд том хэмжээний шошготой өгөгдөл, сургалтын корпус хязгаарлагдмал байдаг. Ялангуяа сошиал медиа орчинд түгээмэл ажиглагддаг монгол-англи холимог (code-mixed) бичвэр нь хэлний бүтэц болон бичгийн хэв маягийн ялгаанаас шалтгаалан мэдрэмжийн шинжилгээнд хүндрэл үүсгэдэг. Энэхүү судалгаанд 3000 мөр шошготой холимог бичвэрийг ашиглан mBERT, XLM-RoBERTa, ALBERT загваруудыг эхнээс нь сургасан бөгөөд 35–37%-ийн нарийвчлал үзүүлсэн.

Харин урьдчилан сурсан загваруудыг нарийн тохируулга (fine-tuning) хийснээр гүйцэтгэл 74–84%-д хүрч, XLM-RoBERTa загвар 84%-ийн нарийвчлалтайгаар хамгийн сайн үр дүнг үзүүлсэн. Цаашид сургалтын өгөгдлийн хэмжээг нэмэгдүүлэх зорилгоор өгөгдлийн нэмэгдүүлэлтэд суурилсан синтетик өгөгдөл үүсгэх (data augmentation) аргаар өгөгдлийн хэмжээг 5000 мөрт хүргэхэд XLM-RoBERTa загвар 85%-д хүрсэн. Мөн хагас хяналттай сургалтад суурилсан автомат шошгололт (semi-supervised learning) ашиглан өгөгдлийн хэмжээг 6050 мөрт хүргэснээр XLM-RoBERTa загвар 86%-ийн нарийвчлалтай болж сайжирсан. Эдгээр аргуудыг хослуулан ашигласнаар загварын гүйцэтгэл 87%-д хүрч, нарийн тохируулга хийсэн суурь үр дүнтэй харьцуулахад 3%-иар сайжирсан. Энэхүү судалгаа нь нөөц багатай нөхцөлд өгөгдлийн чанартай нэмэгдүүлэлт нь мэдрэмжийн шинжилгээний гүйцэтгэлийг сайжруулах боломжтойг харуулж байна.

Түлхүүр үг— mBERT, XLM-RoBERTa, ALBERT.

CNN архитектуруудын олон удаагийн туршилтад суурилсан нарийвчилсан статистик харьцуулалт

Сундуйн Энхбаяр
Статистик, хэрэглээний математикийн тэнхим
Мэдээллийн технологи, электроникийн сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
s.enkhbayar@gmail.com

Нугэжгарын Пүрэвцогт
Статистик, хэрэглээний математикийн тэнхим
Мэдээллийн технологи, электроникийн сургууль
Монгол Улсын Их Сургууль
Улаанбаатар, Монгол
purevtsogt@num.edu.mn

Хураангуй—Суурь мэдээлэл: Гүн сургалт (deep learning) эмнэлгийн зургийн шинжилгээнд шинэ дэвшлийг авч ирсэн боловч CNN (convolutional neural network) архитектуруудын давуу болон сул талуудын талаарх мэдээлэл дутмаг байна. Цөөн тооны судалгаанд ганц цэгийн нарийвчлалаар харьцуулах, мөн тодорхой бус байдлын хэмжүүрүүд, effect size, multiple comparison-ийн зохих засваргүйгээр дүгнэлт гаргадаг нь CNN архитектуруудыг шинжлэх ухааны үндэслэлтэйгээр харьцуулан дүгнэхэд төвөгтэй байдлыг үүсгэж байна.

Зорилго: Энэхүү судалгаа нь цээжний рентген зургаас уушгины хатгалгаа илрүүлэхэд ашиглаж буй гурван түгээмэл CNN архитектурыг (ResNet-50, DenseNet-121, EfficientNet-B0) харьцуулах иж бүрэн статистик арга зүйг тодорхойлж, үнэлэх явдал юм.

Арга зүй: 5-fold stratified cross-validation-ийг 5 өөр санамсаргүй seed-ээр давтаж, архитектур бүрд 25 бие даасан гүйцэтгэлийн хэмжилт бий болгосон. Paired t-test, Wilcoxon signed-rank test, DeLong test, McNemar тест зэрэг параметрик болон параметрик бус шинжилгээнүүдийг Cohen’s d, Cliff’s delta effect size хэмжүүрүүд, Bonferroni/Holm-Bonferroni multiple comparison correction-тэй хамт хэрэглэсэн.

Үр дүн: Авч үзсэн гурван загвар бүгд 0.996-аас дээш AUC-ROC буюу тухайн архитектурын чанар маш сайн гэсэн дүгнэлтэд хүрсэн. Үүнээс DenseNet-121 хамгийн өндөр дундаж AUC-ROC (0.9977 ± 0.0006) үр дүнг өгсөн бол ResNet-50 (0.9975 ± 0.0008) нь статистикийн хувьд ялгаагүй (p =0.259, Cohen’s d =−0.32). Хоёулаа EfficientNet-B0 (0.9966 ± 0.0010)-ооc статистик ач холбогдолтой илүү сайн (p < 0.01, Cohen’s d > 1.0). McNemar тест нь DenseNet-121 нь хүнд тохиолдлуудыг ангилахдаа EfficientNet-B0-ооc давуу байгааг батлагдсан. DenseNet-121 EfficientNet-B0-ийн буруу ангилсан 31 тохиолдлыг зөв ангилсан (OR = 4.43).

Дүгнэлт: DenseNet-121-ийг хамгийн бага хэлбэлзэлтэй хамгийн өндөр дундаж гүйцэтгэл, дээд зэргийн параметрийн үр ашиг (ResNet-50-аас 3 дахин бага параметр), хүнд тохиолдлуудыг ангилах давуу талтай тул эмнэлгийн хэрэглээнд түлхүү ашиглах боломжтой гэсэн дүгнэлтэд хүрч байна.

Түлхүүр үг—Конволюцийн мэдрэлийн сүлжээ, Уушгины хатгалгаа илрүүлэх, Статистик харьцуулалт, DeLong тест, McNemar тест, Гүн сургалт, Цээжний рентген, Cross-validation, Effect size

The Technical Debt of Rapid Digitalization: A Quantitative Analysis of Software Maturity in Mongolia

Youngmi Kim
Computer Science department
Huree College of ICT ymkim@huree.edu.mn

Abstract—This study quantitatively evaluates the software maturity of five key projects in Mongolia's finance, infrastructure, and NLP sectors using SonarQube-based static analysis. By examining software quality metrics such as Lines of Code (LOC), Cyclomatic Complexity (CC), and the Maintainability Index (MI), the research identifies a significant divergence from global benchmarks. While the projects maintain structural simplicity (Avg. CC 3.2), they exhibit a critical "verification gap" characterized by 0.0% test coverage and a high density of security hotspots, leading to an overall Security Grade of E . These findings demonstrate a prevalent pattern of "Emerging-market Technical Debt," where long-term reliability and security are compromised for immediate development speed. The study concludes by proposing a strategic roadmap, including mandatory SQA frameworks, to mitigate these risks and enhance the national IT ecosystem.

Keywords—Software Quality Metrics, Static Analysis, SonarQube, Technical Debt, Security Hotspots, Test Coverage , Maintainability Index (MI) , Software Maturity.

Рансомвэр төрлийн халдлагаас урьдчилан сэргийлэх илрүүлэх программ хангамж хөгжүүлэх нь

Ганзоригийн ГАНБАЯР
Кибер аюулгүй байдлын тэнхмийн оюутан
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар хот, Монгол улс
ganbayar.ganzorig@gmail.com

Содномдамбын СУГАРСҮРЭН
Кибер аюулгүй байдлын тэнхмийн оюутан
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар хот, Монгол улс
ssugarsuren023@gmail.com

Зоригтбаатарын МӨНХ-ЭРДЭНЭ
Кибер аюулгүй байдлын тэнхмийн оюутан
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар хот, Монгол улс
z.monherdene22@gmail.com

Баярмагнайн УНДРАЛ
Кибер аюулгүй байдлын тэнхмийн багш
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар хот, Монгол улс
undral.b@must.edu.mn

Цэдэнбалын ТЭМҮҮЛЭН
Кибер аюулгүй байдлын тэнхмийн оюутан
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар хот, Монгол улс
tsedenbaltemuulen2@gmail.com

Бат-Эрдэнийн МӨНХБАЯР
Кибер аюулгүй байдлын тэнхмийн эрхлэгч, доктор (Ph.D), дэд профессор
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар хот, Монгол улс
munkhbayar.b@must.edu.mn

Хураангуй—Рансомвер халдлага нь сүүлийн жилүүдэд кибер аюулгүй байдлын салбарын хамгийн ноцтой заналхийллийн нэг болж, дэлхий даяар байгууллага, хувь хэрэглэгчдэд асар их хохирол учруулсаар байна. Одоогийн хамгаалалтын системүүд нь хуурамч эерэг дохиоллын өндөр түвшин, рансомвер процессыг нарийн тодорхойлох чадвар сул, системийн гүйцэтгэлд үзүүлэх нөлөө их зэрэг сул талуудтай байдаг. Энэхүү судалгааны ажил нь рансомвер халдлагыг эрт үе шатанд илрүүлэх, урьдчилан сэргийлэх зорилгоор Windows үйлдлийн системд зориулсан WinShield нэртэй программ хангамжийн шийдлийг боловсруулсан. Судалгааны аргазүйн хүрээнд Kernel-Mode Minifilter драйверт суурилсан файлын системийн бодит цагийн хяналт, олон түвшний шинжилгээний механизм,
Trap file, Registry, Entropy, VSC, API call зэрэг зургаан өөр эх үүсвэрт суурилсан илрүүлэлт, XGBoost машин сургалтын алгоритм ашигласан хөнгөн жинтэй процесс ангилагчийг хослуулсан цогц системийг хэрэгжүүлсэн. Мөн илрүүлсэн халдлагад процесс блоклох, файлын үйлдлийг хаах, сүлжээнээс тусгаарлах, лог хөтлөх, файл сэргээх зэрэг хариу үйлдлийн таван механизмыг нэгтгэсэн. Туршилтын үр дүнгээр WinShield систем нь 97.8% нарийвчлалтай рансомвер илрүүлэлт, 2.3% хуурамч эерэг түвшин, 0.8 секундийн дундаж хоцролт, CPU ачаалал дунджаар 4.2%иар нэмэгдэх, 12 төрлийн zero-day рансомвер илрүүлэх чадвартай болох нь тогтоогдсон. Эдгээр үр дүн нь санал болгож буй hybrid илрүүлэлтийн арга нь рансомвер халдлагыг илрүүлэх нарийвчлалыг нэмэгдүүлэх, хуурамч эерэг дохиоллыг бууруулах, системийн гүйцэтгэлд үзүүлэх нөлөөг багасгах боломжтойг харуулж байна. Судалгааны шинэлэг тал нь олон түвшний илрүүлэлтийн механизм болон машин сургалтад суурилсан процессын ангилалтыг хослуулсан hybrid загвар боловсруулсан явдал юм.

Түлхүүр үг— Рансомвер, кибер аюулгүй байдал, KernelMinifilter драйвер, машин сургалт

Нарны хавтангийн гадаргуу цэвэрлэх алсын удирдлагатай робот системийн судалгаа, хөгжүүлэлт

Лхагвасүрэн Анужин
Электроникийн тэнхим
ШУТИС-МХТС
Монгол улс, Улаанбаатар
anujin.l@mobicom.mn

Энхбаатар Сумъяабазар
Электроникийн тэнхим
ШУТИС-МХТС
Монгол улс, Улаанбаатар
sumya0026@gmail.com

Баатар Зоригтбаатар
Электроникийн тэнхим
ШУТИС-МХТС
Монгол улс, Улаанбаатар
bzorigt@must.edu.mn

Найманнаран Чулуунбанди
Холбооны Инженерчлэлийн тэнхим
ШУТИС-МХТС
Монгол улс, Улаанбаатар
chuka@must.edu.mn

Хураангуй— Энэхүү судалгааны ажлаар алслагдсан бүс нутагт байрлах нарны эрчим хүчинд суурилсан системийн найдвартай ажиллагааг нэмэгдүүлэх зорилгоор нарны хавтангийн гадаргуугийн бохирдол, сүүдэрлэлт, халхлалтын нөлөөг үнэлж, түүнд суурилсан ухаалаг цэвэрлэгээний робот системийг боловсруулж хэрэгжүүлэв. Судалгаанд механик хийц, математик загварчлал, алгоритм боловсруулалт, алсын хяналт-удирдлагын IoT архитектур, туршилтын мониторингийн аргуудыг нэгтгэн ашигласан.

Систем нь ремен дамжуургат шүүрэн механик бүтэц, Pololu Dual VNH5019 мотор драйвер, 12V 296 rpm редукторт хөдөлгүүр, микроконтроллерт суурилсан удирдлага, сервер-талын өгөгдлийн сан, веб dashboard бүхий алсын хяналтын платформоос бүрдэнэ. Төхөөрөмж нь нарны хавтангийн хүчдэл, гүйдлийг бодит цагт хэмжиж серверт дамжуулах, хуваарьт, алсын командын болон хүчдэлийн бууралтад суурилсан ухаалаг (emergency) горимоор цэвэрлэгээ хийх боломжтой.

Мөн системд эрчим хүчний хэрэглээг бууруулах зорилгоор цэвэрлэгээ хийгдээгүй үед sleep горимыг хэрэгжүүлж, шаардлагагүй ажиллагааг хязгаарласан. Туршилтын үр дүнд хэсэгчилсэн сүүдэрлэлт нь хүчдэлийг аажмаар бууруулдаг бол гадаргуугийн бүрэн халхлалт нь огцом уналт үүсгэдэг болохыг тогтоосон. Урт хугацааны мониторингийн өгөгдөлд бохирдлын нөлөөгөөр гаралтын хүчдэлийн түвшин буурч, цэвэрлэгээний дараа сэргэсэн хандлага ажиглагдсан.

Иймээс гаралтын хүчдэлийн өөрчлөлтөд суурилсан бохирдлын индекс тодорхойлон, түүнийг цэвэрлэгээ эхлүүлэх шалгуур болгон ашиглах боломжтой бөгөөд боловсруулсан систем нь алслагдсан нарны эрчим хүчний байгууламжийн засвар үйлчилгээ, ашиглалтын үр ашгийг нэмэгдүүлэх инженерийн шийдэл болохыг харуулж байна.

Түлхүүр үг— нарны хавтан, цэвэрлэгч робот, алсын хяналт, математик загварчлал, эрчим хүчний баланс, үүрэн холбооны суурь станц

Монгол улсад дронд суурилсан хүргэлтийн системийг хэрэгжүүлэх боломжийн судалгаа

Б.Амаржаргал
Мэдээллийн технологийн салбар
ШУТИС, Мэдээлэл, холбооны технологийн сургууль
Монгол улс, Улаанбаатар хот
J.DS25E005@must.edu.mn

С.Өлзийбаяр
Мэдээллийн технологийн салбар
ШУТИС, Мэдээлэл, холбооны технологийн сургууль
Монгол улс, Улаанбаатар хот
ulziibayar@must.edu.mn

Хураангуй — Энэхүү судалгааны зорилго нь Монгол Улсын нөхцөлд дронд суурилсан хүргэлтийн системийг хэрэгжүүлэх техникийн болон зохион байгуулалтын боломжийг тодорхойлох, мөн уур амьсгал, газарзүй, хотын орчны онцлогт нийцсэн зам төлөвлөлтийн аргуудыг үнэлэхэд оршино. Судалгаанд OpenStreetMap өгөгдөлд суурилан 2 м нарийвчлалтай торон орчин үүсгэж, барилга байгууламжийг саад болгон загварчилсан. Зам төлөвлөлтөд зай, салхи болон барилгаас аюулгүй зайг нэгтгэсэн жинлэсэн өртгийн функц ашиглан Dijkstra, A*, Theta*, эрчим хүчийг тооцсон A* болон PSO зэрэг аргуудыг харьцуулан туршсан. Үр дүнгээс Theta* алгоритм нь хамгийн шулуун, эргэлт багатай маршрут үүсгэж, нислэгийн геометрийн хувьд оновчтой шийдэл болохыг харуулсан. Харин эрчим хүчийг тооцсон A* алгоритм нь салхи болон аюулгүй зайг харгалзан үзсэн илүү тогтвортой, бодит нөхцөлд тохиромжтой маршрут гаргасан. PSO арга нь тооцооллын хугацаа их, үр дүн тогтворгүй тул бодит хэрэглээнд тохиромжгүй болох нь тогтоогдсон. Мөн төлөвлөсөн маршрутуудыг PX4 SITL болон Gazebo симуляцийн орчинд гүйцэтгэж, төлөвлөсөн зам ба бодит нислэгийн гүйцэтгэлийг харьцуулан үнэлэхэд алгоритмын үр дүнг бодит нислэгийн орчинд хэрэгжүүлэх боломжтойг баталсан. Судалгааны үр дүн нь Монгол Улсад дронд суурилсан хүргэлтийн системийг үе шаттайгаар нэвтрүүлэх техникийн үндэс болон зам төлөвлөлтийн оновчтой шийдлийг тодорхойлж байна.

Түлхүүр үг — UAV, Drone Delivery, Path Planning, Energy-aware Routing, Simulation

Хувийн санхүүгийн удирдлагыг оновчтой болгох AI-д суурилсан ухаалаг туслах мобайл аппликейшн боловсруулах нь

Ц. Энхтуяа
Мэдээлэл, Компьютерын Ухааны Тэнхим
МТЭС, МУИС
enkhtuya.ts@num.edu.mn
Г. Содномдолгор
Санхүүгийн тэнхим
Бизнесийн сургууль, МУИС
dolgorganaa0125@gmail.com

Б. Бат-Эрдэнэ
Мэдээлэл, Компьютерын Ухааны Тэнхим
МТЭС, МУИС
baterdenebatbayr321@gmail.com
*С. Ариунаа**
Санхүүгийн Тэнхим
Бизнесийн сургууль, МУИС
s.ariunaa@num.edu.mn

Abstract—Монгол Улсад хэрэглээний зээлийн хэмжээ огцом өсөж, 2024 оны эцсийн байдлаар банкны систем дэх хэрэглээний зээлийн үлдэгдэл 11.4 их наяд төгрөгт хүрч, өмнөх оноос 35.5 хувиар нэмэгдсэн байна. Орлогын өсөлтөөс давсан зардлын өсөлт, тэтгэврийн насны иргэдийн зээлийн өндөр хамаарал, мөн өр, орлогын харьцааны зохицуулалт 2024 онд 55 хувь, 2025 онд 50 хувь, 2026 оны эхнээс 45 хувь болж дараалан чангатгагдсан нь иргэдийн хувийн санхүүгийн удирдлагыг сайжруулах шаардлага нэмэгдэж байгааг харуулж байна.
Энэхүү судалгааны ажлын зорилго нь Монголын нөхцөлд тохирсон, хиймэл оюунд суурилсан хувийн санхүүгийн ухаалаг туслах мобайл аппликейшний системийн загварыг боловсруулахад оршино. Санал болгож буй систем нь (i) орлого, зарлагын өгөгдөлд тулгуурласан шинжилгээ болон LSTM загвараар богино хугацааны таамаглал хийх, (ii) зарлагыг “танах боломжтой” болон “зайлшгүй” гэж ангилан what-if симуляци гүйцэтгэх, (iii) хэрэглэгчийн зан төлөвт суурилсан K-Means кластерчлал болон Reinforcement Learning аргачлалаар хувь хүнд тохирсон зөвлөмж боловсруулах, (iv) зээлийн бүтцийг оновчлох алгоритмын шийдлийг багтаана.
Судалгааны хүрээнд санал болгож буй модулиудыг үнэлэх арга зүйг боловсруулж, LSTM таамаглалыг MAE, RMSE, MAPE үзүүлэлтээр, зарлагын ангиллыг Accuracy, Precision, Recall, F1-score үзүүлэлтээр, K-Means кластерчлалыг Silhouette score болон Davies-Bouldin indexээр, Reinforcement Learning зөвлөмжийн модулийг cumulative reward, convergence болон saving rate үзүүлэлтээр үнэлэхээр тодорхойлов. Мөн санал болгож буй AI аргачлалуудыг moving average, naive forecast, rule-based зөвлөмж болон Avalanche өр төлөлтийн стратеги зэрэг baseline аргуудтай харьцуулах үнэлгээний хүрээг санал болгосон.
Ингэснээр энэхүү судалгаа нь зөвхөн хувийн санхүүгийн бүртгэлийн аппликейшн бус, хэрэглэгчийн зан төлөвт дасан зохицож, санхүүгийн шийдвэр гаргалтыг дэмжих AI-д суурилсан ухаалаг туслах системийн архитектур, үнэлгээний арга зүйг Монголын нөхцөлд нийцүүлэн боловсруулснаараа ач холбогдолтой юм.

Index Terms—хувийн санхүү, хиймэл оюунд суурилсан зөвлөмж, зан төлөвийн шинжилгээ, зардлын хяналт, зээлийн удирдлага, what-if шинжилгээ, машин сургалт

Чичиргээнд суурилсан автомашины хөдөлгүүрийн эвдрэл илрүүлэх: domain adaptation аргаар загварыг шинэ орчинд дасан зохицуулах

Зоригтбаатар Сарантуяа
Мэдээллийн технологи
ШУТИС, МХТС
Монгол, Улаанбаатар
B231890012@must.edu

Дарханбаатар Наранбат
Мэдээллийн технологи
ШУТИС, МХТС
Монгол, Улаанбаатар
231890043@must.edu.mn

Соном-Очир Өлзийбаяр
Мэдээллийн технологи
ШУТИС, МХТС
Монгол, Улаанбаатар
ulziibayar@must.edu.mn

Хураангуй: Автомашины хөдөлгүүрийн эвдрэл нь ажлын бүтээмж, аюулгүй байдал болон үйлчилгээний зардалд шууд нөлөөлдөг тул түүнд хийн оношилгоо чухал үүрэг гүйцэтгэдэг. Энэ судалгаанд гурван тэнхлэгт акселерометрээр хэмжигдсэн чичиргээний дохиог ашиглан хөдөлгүүрийн хэвийн ба эвдрэлтэй төлвүүдийг ялгах машин сургалтын загварыг боловсруулж, олон RPM ба орчны нөхцөл (температур, чийгшил)–ийн нөлөөг судаллаа. Өгөгдлийг шүүлт, нормчлол, шинж тэмдгийн олборлолт (RMS, FFT-д суурилсан шинжүүд) болон raw 1D дохиог шууд ашиглах аргаар боловсруулсан. Random Forest, SVM, XGBoost болон 1D-CNN зэрэг аргуудыг харьцуулсан бөгөөд олон RPM-ийг нэгтгэн ашигласнаар Random Forest-ийн гүйцэтгэл 89%–аас 95% хүртэл сайжирсан; орчны нөхцлийг шинжүүлсэн загварууд 96–99% нарийвчлалтай ажилласан. Гэвч өөр төрлийн автомашины (target domain) өгөгдөл дээр эхний CNN-ийн гүйцэтгэл буурсан тул domain adaptation аргыг (CORAL) few-shot fine-tuning-тэй хослуулан хэрэглэв. Target domain дээр 20 sample-аар fine-tune хийсний үр дүнд CNN загварын нарийвчлал 42%–аас 90.1% болж нэмэгдэв. Судалгааны үр дүн нь RPM ба орчны нөхцөл нь чичиргээний оношилгоонд чухал нөлөөтэйг, мөн CORAL болон few-shot стратегиудыг хослуулан domain shift-ийг бууруулж загварыг шинэ орчинд амжилттай дасан зохицуулж болдог болохыг харуулсан ба автомашины хөдөлгүүрийн эвдрэлийг чичиргээний сигналын шинжилгээ болон хиймэл оюун ухааны аргаар үр дүнтэй илрүүлэх боломжтойг харуулсан.

Түлхүүр үг: машин сургалт, чичиргээний анализ, автомашины оношилгоо, CNN, domain adaptation

Үүлэн орчин дахь аюулгүй байдлын удирдлагыг автоматжуулах нь

Буяндэлгэрийн ЭРДЭНЭБАЯР
Кибер аюулгүй байдлын тэнхимийн оюутан
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар, Монгол улс
B231870052@must.edu.mn

Төмөрбаатарын СЭРГЭЛЭН
Кибер аюулгүй байдлын тэнхимийн оюутан
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар, Монгол улс
B231870007@must.edu.mn

Дорждэрэмийн БИЛГҮҮН
Кибер аюулгүй байдлын тэнхимийн оюутан
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар, Монгол улс
B231870007@must.edu.mu

Баярмагнай УНДРАЛ
Кибер аюулгүй байдлын тэнхим
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар, Монгол улс
дэд профессор undral.b@must.edu.mn

Бат-Эрдэнийн МӨНХБАЯР
Кибер аюулгүй байдлын тэнхим
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар, Монгол улс
тэнхимийн эрхлэгч, доктор (Ph.D),

Хураангуй— Үүлэн тооцооллын хэрэглээ эрчимтэй өсөхийн хэрээр буруу тохиргоо, хэт өргөн эрхийн бодлого, журналжуулалтын сул хэрэгжилт, шифрлэлтийн алдаа, стандартын нийцлийн алдагдал зэрэг эрсдэл байгууллагын мэдээллийн аюулгүй байдалд томоохон нөлөө үзүүлж байна. Арилжааны Cloud Security Posture Management (CSPM) шийдлүүд өргөн хэрэглэгддэг боловч өртөг өндөр, хаалттай архитектуртай, байгууллагын онцлог шаардлагад бүрэн нийцүүлэхэд хязгаарлагдмал байх нь бий.

Энэхүү судалгаанд үйл явдалд суурилсан (eventdriven) зарчимд тулгуурласан, мета-өгөгдөлд суурилсан үүлэн аюулгүй байдлын хяналтын архитектур санал болгож, IAM, хадгалалт, сүлжээний ангиллаар дүрмийн шалгалт хийх, эрсдэлийг үнэлэх, тайлагнах урсгалыг нэгтгэсэн. Судалгаанд CloudTrail, EventBridge, SNS зэрэг үйлчилгээн дээр суурилсан идэвхжүүлэгч (триггер) механизм, Python/Boto3 ашигласан нөөцийн метаөгөгдөл цуглуулалт, CIS Benchmark-д уялдуулсан 51 дүрэм бүхий шалгалтын санг ашиглав.

Туршилтаар AWS-д 2.4 секунд, Azure-д 3.1 секунд, GCP-д 1.8 секунд дундаж хугацаанд үнэлгээ гүйцэтгэж болохыг тогтоосон бөгөөд энэ нь бараг бодит цаг (near real-time) хяналт хэрэгжүүлэх боломжтойг харуулсан. Үр дүнгээс харахад санал болгож буй аргачлал нь гараар хийдэг аудит болон тогтмол хугацаат сканнераар харьцуулахад илрүүлэлтийн хурдыг сайжруулж, буруу тохиргоог эрт илрүүлэх, эрсдэлийг ангилах, залруулгын зөвлөмж өгөхөд үр дүнтэй байна.

Түлхүүр үг—Үүлэн аюулгүй байдал, буруу тохиргоо, үйл явдалд суурилсан хяналт, олонүүлэн орчин, нийцэл

Гүн сургалтын аргаар компьютер томограф зургаас гавал тархины гэмтлийг илрүүлэх нь

Доржгочоо САРАНГЭРЭЛ
Мэдээллийн Технологийн тэнхим
ШУТИС, МХТС
Монгол улс, Улаанбаатар
saran@must.edu.mn

Батжаргал НАМУУН
Мэдээллийн Технологийн тэнхим
ШУТИС, МХТС
Монгол улс, Улаанбаатар
b.namuun2006@gmail.com

Үүрцайх МӨНХТУЯА
Мэдээллийн Технологийн тэнхим
ШУТИС, МХТС
Монгол улс, Улаанбаатар
munkhtuyauurtsaikh@gmail.com

Сарсенбай НУРСУЛУ
Мэдээллийн Технологийн тэнхим
ШУТИС, МХТС
Монгол улс, Улаанбаатар
nnnxd47@gmail.com

Хураангуй — Хиймэл оюуны технологи нь эрүүл мэндийн салбарт, ялангуяа дүрс оношилгоонд хурдацтай хөгжиж, эмнэлгийн мэдээллийг илүү нарийвчлалтай, хурдан боловсруулах боломжийг бүрдүүлж байна. Гавал тархины гэмтэл нь дэлхий даяар нас баралт болон насан туршийн хөгжлийн бэрхшээлийн тэргүүлэх шалтгаануудын нэг юм. Монгол Улсад зам тээврийн осол, өндрөөс унах, үйлдвэрлэлийн болон ахуйн осол зэрэг шалтгаанаас үүдэлтэй толгой, тархины гэмтэл нь гэмтлийн шалтгаант нас баралтын 56%-ийг бүрдүүлдэг. Түүнчлэн Монгол орны хүйтэн, эрс тэс уур амьсгал, өвлийн улиралд халтиргаа гулгаа ихсэх, зам тээврийн ослын давтамж нэмэгдэх зэрэг хүчин зүйлс нь тархины гэмтлийн эрсдэлийг нэмэгдүүлэхэд нөлөөлдөг. Сүүлийн жилүүдэд хиймэл оюун ухаан, ялангуяа гүн сургалтын аргууд нь эмнэлгийн дүрс боловсруулалтын салбарт өргөн хэрэглэгдэх болж, СТ зурагнаас хүний нүдээр шууд илрүүлэхэд хүндрэлтэй нарийн дүрслэлийн хэв шинжүүдийг автоматаар илрүүлэх боломжийг бий болгож байна. Үүнд жижиг хэмжээний цус хуралт, тархины хавангийн эхний үеийн өөрчлөлт, тархины эдийн нягтралын нарийн ялгаа зэрэг багтана. Энэхүү судалгаанд өөрсдийн цуглуулсан компьютер томографын өгөгдлийг ашиглан гавал тархины гэмтлийг автоматаар илрүүлэх болон өвчтөний эмнэлзүйн үр дүнг урьдчилан таамаглах зорилгоор конволюшин нейрон сүлжээнд суурилсан гүн сургалтын аргыг ашиглаж туршсан. Мөн туршилтын үр дүнг эмч нарын клиник үнэлгээтэй харьцуулсан.

Түлхүүр үг — Хиймэл оюун, Гавал тархины гэмтэл, Компьютер томограф (СТ), Гүн сургалт, Хиймэл оюуны тусламжтай оношилгоо

Хиймэл оюунд суурилсан имэйлийн фишинг ангиллын систем хөгжүүлэх

Ган-Эрдэнийн СУГАРХҮҮ
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн оюутан
ШУТИС, МХТС
Улаанбаатар, Монгол улс
sugaraa4960@gmail.com

Батнасангийн ТӨГӨЛДӨР
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн оюутан
ШУТИС, МХТС
Улаанбаатар, Монгол улс
B231870015@must.edu.mn

Батболдын ЖАРГАЛ
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн оюутан
ШУТИС, МХТС
Улаанбаатар, Монгол улс
B231870028@must.edu.mn

Баярмагнайн УНДРАЛ
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн багш
ШУТИС, МХТС
Улаанбаатар, Монгол улс
undral.b@must.edu.mn

Бат-Эрдэнийн МӨНХБАЯР
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн эрхлэгч, доктор (Ph.D.), дэд профессор
ШУТИС, МХТС
Улаанбаатар, Монгол улс
munkhbayar.b@must.edu.mn

Хураангуй— Өнөөгийн цахим орчинд фишинг халдлага нь кибер аюулгүй байдлын хамгийн өргөн тархсан, хохирол ихтэй аюулуудын нэг болоод байна. APWG-ийн мэдээгээр 2019–2025 онд фишинг халдлагын тоо 400 гаруй хувиар өсөж, санхүүгийн хохирол 9.8 тэрбум ам.долларт хүрсэн байна. Уламжлалт дүрэмд суурилсан болон хар жагсаалтын аргууд нь орчин үеийн нарийн төвөгтэй, хувьсан өөрчлөгдөж буй фишинг халдлагыг бүрэн илрүүлэх чадваргүй болоод байна. Энэхүү судалгааны ажлын зорилго нь CNN (Convolutional Neural Network) болон BiLSTM (Bidirectional Long Short-Term Memory) загваруудыг Attention механизмтай хослуулан, имэйлийг автоматаар өндөр нарийвчлалтайгаар фишинг эсвэл аюулгүй гэж ангилах системийг боловсруулах юм. Аргазүйн хувьд бид 2023–2025 оны орчин үеийн халдлагуудыг багтаасан MEAJOR, CIC-Trap4Phish болон Монгол хэл дээрх зохиомол имэйлүүдээс бүрдсэн нийт 188,377 өгөгдлийн багцыг ашиглан CNN загвараар хэв маяг илрүүлж, BiLSTM загвараар өгүүлбэрийн агуулга, контекстийг боловсруулан, Attention механизмаар фишинг шинжтэй үгсийг онцолдог нэгдсэн архитектурыг бүтээсэн. Мөн SPF, DKIM, DMARC баталгаажуулалтыг нэгтгэн ашигласан. Бодит орчны туршилтыг Монгол хэл дээрх 1,000 имэйл дээр явуулахад систем нь 85.4%-ийн нийт нарийвчлал (Accuracy) үзүүлсэн бөгөөд хамгийн чухал нь бодит фишинг халдлагуудыг 100%-ийн магадлалтайгаар (Recall=100.0%) нэг ч алдаагүй илрүүлсэн. Хэдийгээр хууль ёсны маркетинг болон төлбөрийн мэдэгдлийг андуурч ангилах (FP) сорилт байгаа ч, систем нь фишинг халдлагаас хэрэглэгчийг бүрэн хамгаалах өндөр практик ач холбогдолтой болох нь тогтоогдсон.

Түлхүүр үг— Нейрон сүлжээ, Кибер аюулгүй байдал, Гүн сургалт, Баталгаажуулалт, Өгөгдөл боловсруулалт

Бэхжүүлэх сургалтад суурилсан халдлага илрүүлэх систем хөгжүүлэх

Лхагва-Очирын ЧИНГАРАЛ
Кибер аюулгүй байдлын судалгааны баг,
Кибер аюулгүй байдлын тэнхимийн оюутан,
ШУТИС, МХТС
chingaral.lkha@gmail.com

Тамирын ЭНХБАЯР
Кибер аюулгүй байдлын судалгааны баг,
Кибер аюулгүй байдлын тэнхимийн оюутан,
ШУТИС, МХТС
Tamirenkhbayar3@gmail.com

Сумъяагийн ХАЙДАВ
Кибер аюулгүй байдлын судалгааны баг,
Кибер аюулгүй байдлын тэнхимийн оюутан,
ШУТИС, МХТС
khaidavsumiya@gmail.com

Баярмагнайн УНДРАЛ
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн багш
ШУТИС, МХТС
Улаанбаатар, Монгол улс
undral.b@must.edu.mn

Бат-Эрдэнийн МӨНХБАЯР
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн эрхлэгч, доктор (Ph.D.), дэд профессор
ШУТИС, МХТС
Улаанбаатар, Монгол улс
munkhbayar.b@must.edu.mn

Хураангуй— Мэдээллийн технологийн салбарт кибер аюулгүй байдлыг хангах асуудал чухлаар тавигдсаар байна. Сүүлийн жилүүдэд халдлагын хэлбэр улам нарийн төвөгтэй, зохион байгуулалттай болж байгаа нь уламжлалт дүрэмд суурилсан илрүүлэлтийн системүүдэд томоохон сорилт болж байна. Иймд энэхүү судалгааны ажлаар бэхжүүлэх сургалт буюу Reinforcement Learning алгоритмд суурилсан динамик халдлага илрүүлэх системийг санал болгож байна. Уг систем нь сүлжээний хувьсан өөрчлөгдөх урсгалд дасан зохицож, орчинтойгоо харилцан үйлчлэх замаар бие даан суралцах агент ашигладгаараа онцлогтой. Туршилтын үр дүнд Q -Learning болон Deep Q-Network алгоритмууд нь хуурамч дохиоллын түвшнийг бууруулж, шинэ төрлийн халдлагыг илрүүлэх нарийвчлалыг мэдэгдэхүйц нэмэгдүүлсэн юм. Системийн үр ашгийг олон улсад хүлээн зөвшөөрөгдсөн NSL-KDD, CICIDS2017 болон UNSW-NB15 өгөгдлийн багцууд дээр туршиж баталгаажууллаа.

Түлхүүр үг—Кибер аюулгүй байдал, халдлага илрүүлэх систем, бэхжүүлэх сургалт, zero - day халдлага, хиймэл оюун ухаан.

Нам өндрийн агаарын хөдөлгөөний удирдлагын системийн өгөгдлийн сангийн загварчлал

Гансүхийн Дүүриймаа, Болдын Наранчимэг†*
Мэдээлэл, компьютерын ухааны тэнхим
МТЭС, МУИС
*Email: *21m1num0367@stud.num.edu.mn, †naranchimeg@num.edu.mn*

Хураангуй—Нисгэгчгүй нисэх хэрэгсэл нь энгийн хэрэглээний төхөөрөмжөөс давж, бизнесийн олон салбарт (үйлчилгээ, логистик, аюулгүй байдал гэх мэт) өргөн ашиглагдах болсноор түүнийг ашиглагчид, иргэд болон нисэхийн байгууллагуудын хоорондын харилцааг зохицуулах хэрэгцээ улам бүр нэмэгдэж байна. Олон улсад дроны хэрэглээг зохицуулах эрх зүйн зохицуулалтууд олон үе шаттайгаар туршигдаж, үүний үндсэн дээр агаарын хөдөлгөөний зохицуулалтын системүүд (UTM) үүсэж хөгжиж байна.

Энэхүү судалгааны ажилд Монгол улсын нам өндрийн нисгэгчгүй нисэх төхөөрөмжийн (UAV) агаарын хөдөлгөөний удирдлагын системд (UTM) зориулсан газарзүйн орон зайн өгөгдлийн сангийн загварыг боловсруулан, прототайп хэлбэрээр хэрэгжүүлж, туршин үнэллээ. Тодруулбал, International Civil Aviation Organization (ICAO)-ийн UTM мэдээллийн загварт тулгуурлан Монгол Улсын онцлог нөхцөл (алслагдсан бүс нутаг болон хүн амын нягтрал өндөртэй хот суурин)—д нийцсэн өгөгдлийн сангийн схемийг боловсруулсан. Үүний зэрэгцээ PostGIS-ийн орон зайн индексийг ашиглан нислэгийн зөрчил илрүүлэх болон geofence шалгалтын асуулгуудыг оновчлон гүйцэтгэсэн.

Прототайпыг PostgreSQL/PostGIS өгөгдлийн сан, Python/FastAPI бэкэнд болон React/Leaflet фронтэнд ашиглан хөгжүүлэв. Улаанбаатар хот болон орон нутгийн бүсэд зориулсан дөрвөн туршилтын хувилбарыг ашиглан spatial query-н хурд, орон зайн индексийн үр ашиг болон нэгэн зэрэг олон UAV-ийн ачааллын үеийн гүйцэтгэлийг үнэллээ. Энэхүү судалгаа нь Монгол улсын Иргэний Нисэхийн Ерөнхий Газар (МСАА)-д UTM системийн өгөгдлийн дэд бүтцийг бий болгох бодлогын шийдвэр гаргахад болон ирээдүйд бүрэн хэмжээний UTM систем нэвтрүүлэхэд шинжлэх ухааны үндэстэй хувь нэмэр болно хэмээн найдаж байна.

Түлхүүр үг—UTM систем, нисгэгчгүй нисэх төхөөрөмж, UAV, UAS, агаарын зайн удирдлага, нислэгийн зөвшөөрөл, прототайп

Нисгэгчгүй нисэх төхөөрөмжийн оптик холбооны судалгаа

Дуламжавын Долгорсүрэн
Communication Engineering Department
SICT, MUST
dolgorsuren@must.edu.mn

Зундуйн Буянхишиг
Communication Engineering Department
SICT, MUST
zbuya@must.edu.mn

Шагдарын Ганболд
Communication Engineering Department
SICT, MUST
sh_ganbold@must.edu.mn

Хураангуй— Харилцаа холбооны төхөөрөмжүүд нь radio frequency (RF) буюу радио давтамж (РД) ашиглан одоогоор төрөл бүрийн бүхий л утасгүй сүлжээнд өргөн холбогдож байна. Одоо байгаа утасгүй технологиуд 10ГГцээс доош давтамжтай цахилгаан соронзон спектрийг өргөнөөр ашигладаг. Цаашилбал, энэ муж (10ГГц-ээс доош) нь нарийн спектрийн зурвас нь спектр ашиглалтын хууль, ойр байрлах РД хандалтын цэгүүдийн хооронд маш их харилцан нөлөөлөл үүсгэж хязгаарлагдах болно. Free space optic (FSO) communication буюу чөлөөт орон зайн оптик холбоо (ЧО^{2}Х) нь дараа үеийн өндөр хурд, өргөн зурвасын утасгүй холбоо бөгөөд газрын гадарга дээрх төхөөрөмж хооронд дэлхийн агаар мандлаар хэт улаан туяаны (ХУТ) болон үзэгдэх гэрлийн дохиог дамжуулдаг. ЧО^{2}Х нь түүнчлэн алсын зайн радио одон орон, нисгэгчгүй нисэх төхөөрөмж (ННТ), цэргийн, гамшгаас хамгаалах, хэрэглэгчийн хандалтын болон утасгүй үүрэн холбооны сүлжээ зэрэг олон чиглэлээр хэрэглэгдэж байна. Энэхүү судалгаа нь зөвшөөрөлгүй зурвасд ННТ-ийг команд, удирдлагын холбоогоор хангахын өөр давтамжийн төрөл буюу оптик зурвас дээр ажилладаг холбооны технологийн судалгааны хэрэгцээг тодорхойлно. Мөн хөгжиж буй ННТийн холбооны системийн найдвартай байдлыг сайжруулах зөвлөмж өгөхөд оршино.

Түлхүүр үг— чөлөөт орон зайн оптик холбоо ЧО^{2}Х, хэт улаан туяа ХУТ, байршил тогтоох систем БТС

Өсвөр насныхны сэтгэцийн эрүүл мэнд ба асран хамгаалагчийн хүн ам зүйн байдлын хоорондын хамаарлын загварчлал

Лувсан-Иш Ажнай
Физик, Мэдээлэл зүй
АШУҮИС, Био-Анагаах
Монгол, Улаанбаатар
ajnai@mnums.edu.mn

Цагаан Баатархүү
Мэдээллийн технологи
ШУТИС, МХТС
Монгол, Улаанбаатар
baatarkhuu@must.edu.mn

*Цэнд Жавзмаа**
Физик, Мэдээлэл зүй
АШУҮИС, Био-Анагаах
Монгол, Улаанбаатар
Javzmaa.ts@mnums.edu.mn

Лувсанцэрэн Пүрэвдолгор
Физик, Мэдээлэл зүй
АШУҮИС, Био-Анагаах
Монгол, Улаанбаатар
purevdolgor@mnums.edu.mn

Ахыт Тилеубай
Мэдээллийн технологи
МҮИС, Мэдээлэл Технологи
Дизайны Сургууль
Монгол улс, Улаанбаатар
Akhit@mnums.edu.mn

Оюунбилэг Бат-Энх
Мэдээллийн технологи
ШУТИС, Мэдээлэл Холбоо
Технологийн Сургууль
Монгол, Улаанбаатар
batenkh@must.edu.mn

Болдбаатар Одгэрэл
Физик, Мэдээлэл зүй
АШУҮИС, Био-Анагаах
Монгол, Улаанбаатар
odgerel.b@mnums.edu.mn

Энхжаргал Энх-Үрэл
Физик, Мэдээлэл зүй
АШУҮИС, Био-Анагаах
Монгол, Улаанбаатар
enkh-urel@mnums.edu.mn

Чилхаасүрэн Баасандорж
Физик, Мэдээлэл зүй
АШУҮИС, Био-Анагаах
Монгол, Улаанбаатар
baasandorj@mnums.edu.mn

Хураангуй — Дэлхийн хэмжээнд өсвөр насныхны сэтгэцийн эрүүл мэнд нь тулгамдсан асуудал хэвээр байгаа бөгөөд ДЭМБ-ын мэдээлснээр 10–19 насны өсвөр үеийнхний ойролцоогоор 7 хүн тутмын 1 нь сэтгэцийн эрүүл мэндийн эмгэгтэй байна. Түүнчлэн сэтгэл түгшилт, сэтгэл гутрал, зан үйлийн эмгэгүүд нь өсвөр насныхны өвчлөл, хөдөлмөрийн чадвар алдалтын гол шалтгаануудын нэг бөгөөд сэтгэцийн эрүүл мэндийн эмгэгүүдийн гуравны нэг нь 14 наснаас өмнө, тал хувь нь 18 наснаас өмнө эхэлдэг болохыг онцолсон байна. Иймээс эрт илрүүлэг, ялангуяа сургууль, гэр бүлд суурилсан скрининг үнэлгээ чухал ач холбогдолтой гэж үздэг. Олон улсын судалгаанд SDQ асуумж нь өсвөр насныхны сэтгэл хөдлөл, зан үйл, анхаарал төвлөрөл, үе тэнгийн харилцааны асуудлыг илрүүлэх найдвартай шалгах хэрэгсэл болно 2021 онд ШУТСАН-гийн докторын дараах инновацийн төслийн хүрээнд СЭМҮТ-тэй хамтран Говь-Алтай аймгийн 6–12 дугаар ангийн 3764 сурагч, 3151 асран хамгаалагч, багш нарыг хамруулсан нийт 10,600 гаруй оролцогчоос SDQ асуумж ашиглан өгөгдөл цуглуулсан. Судалгаагаар сурагч, асран хамгаалагч болон ангийн багшийн сэтгэл хөдлөл, зан үйлийн үнэлгээний өгөгдлийн санг бүрдүүлж, сурагч болон асран хамгаалагчийн хүн ам зүйн үзүүлэлтүүдэд тулгуурлан машин сургалтын аргаар өсвөр үеийнхний сэтгэцийн эрүүл мэндийн асуудлыг илрүүлэх загваруудыг боловсруулан тооцоолж, дүрслэн шинжилсэн.

Түлхүүр үг— data mining, adolescents, C5.0

“Машин сургалт ашиглан макро эдийн засгийн эрсдэлийн индекс боловсруулах: Монгол улсын инфляцийг урьдчилан таамаглах судалгаа”

Ц.Солонгоо
Компьютерын ухаан
Шинэ Монгол технологийн их сургууль
Монгол, Улаанбаатар
solongo.ts@nmit.edu.mn

Э.Батцэцэг
Мэдээллийн технологи
ШУТИС-ийн МХТС
Монгол, Улаанбаатар
e.battsetseg@must.edu.mn

Хураангуй- Сүүлийн жилүүдэд Монгол Улсад инфляцийн түвшин өндөр хэлбэлзэлтэй байж, макро эдийн засгийн тогтвортой байдал болон бодлого боловсруулах үйл явцад томоохон сорилт үүсгэж байна.

Монгол Улсын инфляц нь импортын бараа бүтээгдэхүүн, хүнсний нийлүүлэлт, шатахууны үнийн хэлбэлзэл, мөн цаг уурын онцлог нөхцөл зэрэг олон хүчин зүйлээс хамаарах нийлмэл шинжтэй тул бодит өгөгдөлд суурилан инфляцийн динамикийг тайлбарлах, урьдчилан таамаглах аргачлал боловсруулах шаардлага улам бүр нэмэгдэж байна.

Энэхүү судалгаанд Монгол Улсын инфляцийн динамикийг тайлбарлах, урьдчилан таамаглах зорилгоор машин сургалт болон эконометрик аргуудыг хослуулан макро эдийн засгийн эрсдэлийн индекс (MERI)-ийг боловсруулсан. Судалгаанд 2006– 2025 оны Хэрэглээний Үнийн Индексийн (CPI) дэд бүлгүүд, цаг уурын үзүүлэлтүүд, улирлын эрэлтийг илэрхийлэх баяр ёслолын dummy хувьсагчдыг ашигласан. Эрсдэлийн үзүүлэлтүүдийг z-score болон volatility аргаар тооцоолж, олон эх сурвалжийн өгөгдлийг нэгтгэн тодорхойлсон.

Random Forest болон Elastic Net шинжилгээний үр дүнгээр инфляцид хамгийн их нөлөө үзүүлсэн хүчин зүйлс нь хүнсний үнэ, орон сууц ба эрчим хүч, тээврийн зардал, MERI индекс болон хувцасны үнэ болох нь тогтоогдсон. Эконометрик шинжилгээний үр дүнд MERI индекс нь инфляцийн өөрчлөлтийг 1–3 сарын өмнө урьдчилан тайлбарлах чадвартай (corr ≈ 0.64) болохыг баталсан. Цаашилбал, Инфляцийн цаг хугацааны хамаарлыг загварчлах зорилгоор ARIMAX загварыг ашигласан бөгөөд ARIMAX загвар нь өндөр таамаглалын нарийвчлал харуулсан ($R^2 \approx 0.92$). Иймээс машин сургалт болон цаг хугацааны цуваны загварыг хослуулах арга нь инфляцийн таамаглалын чанарыг мэдэгдэхүйц дээшлүүлж байгааг нотолж байна.

Түлхүүр үг - Инфляци, CPI, Макро эдийн засгийн эрсдэлийн индекс, Машин сургалт, z-score, шугаман регресс, Random Forest, ARIMAX

Машин сургалтын аргад суурилсан робот сургалтын төхөөрөмжийн хөгжүүлэлт

Хүрэлбаатар Цэвээнжав
Электроникийн тэнхим
Шинжлэх Ухаан Технологийн Их Сургууль
Улаанбаатар, Монгол
khurelbaatarts77@gmail.com

Ууганчимэг Ганчимэг
Электроникийн тэнхим
Шинжлэх Ухаан Технологийн Их Сургууль
Улаанбаатар, Монгол
uuganchimeg.g@must.edu.mn

Мөнхбаяр Адъяа
Холбооны инженерчлэлийн тэнхим
Шинжлэх Ухаан Технологийн Их Сургууль
Улаанбаатар, Монгол
A_munkhbayar@must.edu.mn

Хураангуй — Энэхүү судалгааны ажил нь Монголын уламжлалт шагайн наадгайг орчин үеийн робот болон дүрс таних технологитой хослуулсан сургалтын системийг хөгжүүлэхэд чиглэсэн. Судалгаанд шагайн дүрсийг таних хоёр өөр арга болох классик дүрс боловсруулалтын гинжин кодын арга болон гүн сургалтад суурилсан YOLO алгоритмыг ашиглан харьцуулсан. Дүрс боловсруулалтын шатанд хар цагаан болгон хувиргах, шуугиан дарах, ирмэг илрүүлэлт, багасгах зэрэг аргуудыг ашиглан гинжин код үүсгэсэн бол машин сургалтын шатанд YOLO загварыг ашиглан объект илрүүлэх болон ангилах асуудлыг шийдвэрлэсэн. Туршилтын үр дүнгээр гинжин кодын арга нь хяналттай орчинд өндөр нарийвчлалтай боловч давхацсан объект болон бодит орчинд сул, харин YOLO алгоритм нь илүү уян хатан, бодит цагийн гүйцэтгэлтэй болохыг тогтоов.

Түлхүүр үг— Гинжин код, Ирмэг илрүүлэлт, YOLO, Дүрс боловсруулалт

Хөдөлмөрийн аюулгүй байдлын VR сургалтын платформ

Б.Намуун
Мэдээлэл, компьютерын ухааны тэнхимийн докторант
МУИС-ийн Мэдээллийн технологи электроникийн сургууль
Монгол, Улаанбаатар хот
24D1NUM0040@stud.num.edu.mn

Х.Оюундолгор
Мэдээлэл, компьютерын ухааны тэнхимийн профессор
МУИС-ийн Мэдээллийн технологи электроникийн сургууль
Монгол, Улаанбаатар хот
okhorloo@num.edu.mn

Хураангуй— Энэхүү судалгаагаар уул уурхай, хүнд үйлдвэрийн салбарын хөдөлмөрийн аюулгүй байдал (ХАБ)-ын сургалтын үр өгөөж, хүртээмжийг нэмэгдүүлэх зорилго бүхий виртуал реалити (VR) технологид суурилсан интерактив сургалтын платформын архитектур, технологийн шийдлийг танилцуулна. Санал болгож буй систем нь гурван хэмжээст (3D) компьютер график болон 360° панорам дүрслэлийг хослуулсан виртуал орчинд "суралцах, дадлагажих, үнэлүүлэх" гэсэн гурван үе шаттай арга зүйг хэрэгжүүлснээрээ онцлог юм. Түүнчлэн суралцагчийн танин мэдэхүй болон практик гүйцэтгэлийн өгөгдлийг үүлэн тооцооллын дэд бүтцэд (AWS) бодит хугацаанд цуглуулж, багш нар веб хяналтын системээр дамжуулан үнэлэх боломжтой өгөгдөлд суурилсан VR-LMS (Learning Management System) загварыг боловсруулсан. Уг платформыг "Оюу толгой" ХХК болон "Эрдэнэт үйлдвэр" ТӨҮГ-ын үйлдвэрлэлийн бодит орчинд туршин нэвтрүүлэхэд ажилтнуудын сургалтын идэвх оролцоо, ур чадварын гүйцэтгэл мэдэгдэхүйц сайжирсан эерэг үр дүн гарсан. Энэхүү судалгааны ажил нь өндөр эрсдэлтэй орчны аюулгүй ажиллагааны сургалтад дэвшилтэт VR технологийг ашиглах, байгууллагын дотоод сургалтын үйл ажиллагааг дижиталчлах технологийн боломжит шийдэл болохыг харуулж байна.

Түлхүүр үг—интерактив сургалт, сургалтын удирдлагын систем (LMS), виртуал реалити, ХАБ-ын виртуал сургалт

Монгол гэрийн 3D модель үүсгэх арга

Өлзий-Ундралын Оргил
Машин оюуны лаборатори
Мэдээлэл, компьютерын ухааны тэнхим
МТЭС, МУИС
25d1num0110@stud.num.edu.mn

Ганболдын Амарсанаа
Машин оюуны лаборатори
Мэдээлэл, компьютерын ухааны тэнхим
МТЭС, МУИС
amarsanaag@num.edu.mn

Хураангуй—Энэхүү судалгааны ажилд нүүдэлчдийн уламжлалт сууц болох Монгол гэрийн 3D моделийг компьютер графикийн процедур загварчлалын аргаар бодит цаг хугацаанд үүсгэх алгоритмыг санал болгож байна. Уламжлалт, гар аргаар 3D модель үүсгэх, түүний бүтэц, хэмжээг өөрчлөхөд цаг их шаарддаг бол процедур арга нь математик параметрүүдийг өөрчлөх замаар олон хувилбарыг автоматаар хялбар үүсгэх боломжийг олгодог. Ялангуяа нүүдэлчдийн материаллаг соёлын өвийг дижитал хэлбэрт шилжүүлэн хадгалах, түүний архитектурын нарийн хийцийг технологийн тусламжтайгаар ирээдүй хойч үедээ өвлүүлэн үлдээхэд энэхүү автоматжуулсан арга нь стратегийн ач холбогдолтой юм.

Бидний боловсруулсан алгоритм нь гэрийн гадна бүтцийн геометрийн хамаарлыг тригонометрийн функцүүд ашиглан тодорхойлж, 3D моделийг динамикаар үүсгэдэг. Судалгааны хүрээнд бэлэн зураг ашиглахгүйгээр текстурыг алгоритмаар пикселийн түвшинд үүсгэж, UV буулгалтыг өөрийн модельтой автомат байдлаар холбосон. Энэхүү арга нь Монгол гэрийн 3D моделийг маш богино хугацаанд, санах ойн бага нөөцөөр үүсгэх боломжийг олгох бөгөөд дижитал музей, тоглоом хөгжүүлэлт болон соёлын өвийн виртуал сэргээн босголтын салбарт бодит хувь нэмэр оруулна гэж дүгнэж байна.

Түлхүүр үг—Процедур загварчлал, Монгол гэр, 3D сэргээн босголт, Дижитал өв

Монгол адууны тамганы нийлмэл дүрсийг задлан танихад зориулсан бүтцэд суурилсан дүрмийн арга

Пэрэнлэйлхүндэв Гантуяа, Батсуурь Сувдаа, Ганболд Уганбаяр
Монгол Улсын Их Сургууль
Мэдээллийн технологи, электроникийн сургууль
Мэдээлэл, компьютерийн ухааны тэнхим
Улаанбаатар, Монгол Улс
gantuya.p@num.edu.mn suvdaa@num.edu.mn uuganbayar.g@num.edu.mn

Хураангуй— Монгол адууны тамганы дүрсний санд олон шугам нэг контур болж нэгдсэн нийлмэл дүрсүүд түгээмэл тохиолддог бөгөөд энэ нь автомат танихад үндсэн хүндрэл болдог. Энэхүү өгүүлэлд нийлмэл тамгыг шууд нэг бүхэл объект гэж танихын оронд бүтцээр нь задлан уншихад зориулсан дүрмийн аргыг санал болгов. Судалгаанд 1401 зурагтай өгөгдлийн санг ашиглаж, дүрсийг бүрэлдэхүүн дүрсний тоогоор нь 1–6 бүлэгт ангилсан. Санал болгож буй арга нь урьдчилсан боловсруулалт, холбоост хэсэг илрүүлэлт, skeleton-д суурилсан задлал, дараа нь байрлал, хэмжээ, хэлбэрийн дүрмээр үндсэн дүрс, чимэг дүрс, хүрээ, дэвсгэрийг ялгах шатлалтай. Мөн бидний өмнөх судалгаанд тодорхойлсон уламжлалт тамга унших дарааллыг Rule 1–Rule 5 хэлбэрээр формалчилж, нийлмэл дүрсийг тайлбарлах алгоритмд тусгав. Туршилтаас харахад нэг болон хоёр бүрэлдэхүүнтэй тамганд энэхүү арга үр дүнтэй ажиллах боломжтой боловч олон бүрэлдэхүүнтэй, нэг контурт бүрэн нэгдсэн тамганд ambiguity ихсэж, segmentation-ээс илүү бүтцийн тайлбар шаардлагатай байна.

Түлхүүр үг— Монгол адууны тамга, нийлмэл дүрс, задлан таних, дүрмийн арга, бүтцийн задлал

Машин сургалтад суурилсан тосны шинжилгээ ба зогсолтын өгөгдлөөр тоног төхөөрөмжийн эвдрэлийг таамаглах нь

Г.Ганчимэг
Компьютерын ухааны тэнхим
ШУТИС, МХТС
Улаанбаатар, Монгол улс
ganaa@must.edu.mn

Ж.Болортунгалаг
Компьютерын ухааны тэнхим
ШУТИС, МХТС
Улаанбаатар, Монгол улс
bolorjargakh@gmail.com

Хураангуй—Энэхүү судалгаанд үйлдвэрийн тоног төхөөрөмжийн тосны шинжилгээний болон ажиллагааны зогсолтын өгөгдлийг нэгтгэн ашиглаж, эвдрэл үүсэх процесс болон урьдчилан таамаглах боломжийг судлав. Судалгаанд 2017–2025 оны хооронд цуглуулсан 25250 тосны шинжилгээний сорьц болон 2025–2026 оны хооронд бүртгэгдсэн 12 тоног төхөөрөмжийн зогсолтын мэдээллийг ашигласан. Тосны үзүүлэлтүүд болох төмөр (Fe), PQ индекс, цахиур (Si), усны агууламж нь эвдрэлийн түвшнээс хамааран статистикийн ач холбогдол бүхий ялгаатай (Kruskal–

Wallis, $p < 0.001$) болохыг тогтоов. Random Forest загвар нь эвдрэл илрүүлэхэд өндөр гүйцэтгэл (Accuracy = 0.9136, ROC-AUC = 0.9644) үзүүлсэн. Зогсолтын шинжилгээгээр хөдөлгүүрийн систем нийт 1701 цагийн зогсолтыг үүсгэж, UL016 (382 цаг), UL017 (293.9 цаг) зэрэг тоног төхөөрөмжүүд өндөр эрсдэлтэй болохыг тогтоов. Судалгааны үр дүнд бохирдол → элэгдэл → эвдрэл → зогсолт гэсэн дараалсан механизм тогтоогдсон. Энэхүү судалгаа нь тосны шинжилгээ болон зогсолтын өгөгдлийг нэгтгэх нь эвдрэлийг урьдчилан таамаглах гүйцэтгэлийг мэдэгдэхүйц сайжруулдгийг харуулж байна.

Түлхүүр үг—Predictive maintenance, oil analysis, machine learning, downtime analysis, reliability

Эрүүл мэндийн мэдээлэл зүйн хөтөлбөрийн төгсөгчдийн сэтгэл ханамжийн шинжилгээ ба машин сургалтын загварууд

Цагаан Баатархүү
Мэдээллийн технологи
ШУТИС, Мэдээлэл Холбоо
Технологийн Сургууль Монгол,
Улаанбаатар хот
baatarkhuu@must.edu.mn

Чилхаасүрэн Баасандорж*
Физик мэдээлэлзүйн тэнхим
АШУҮИС, БАСургууль,
Монгол улс, Улаанбаатар хот
baasandorj@mnums.edu.mn

Лувсанцэрэн Пүрэвдолгор
Физик мэдээлэлзүйн тэнхим
АШУҮИС, БАСургууль,
Монгол Улс
purevdolgor@mnums.edu.mn

Цэнд Жавзмаа,
Физик мэдээлэлзүйн тэнхим
АШУҮИС, БАСургууль,
Монгол улс, Улаанбаатар хот
javzmaa.ts@mnums.edu.mn

Болдбаатар Одгэрэл
Ерөнхий эрдмийн тэнхим
АШУҮИС, ССургууль,
Монгол улс, Улаанбаатар хот
odgerel.b@mnums.edu.mn

Лувсан Ажнай
Физик мэдээлэлзүйн тэнхим
АШУҮИС, БАСургууль,
Монгол улс, Улаанбаатар хот
ajnai@mnums.edu.mn

Тилуебай Ахыт
Физик мэдээлэлзүйн тэнхим
АШУҮИС, БАСургууль,
Монгол улс, Улаанбаатар хот
akhit@mnums.edu.mn

Уугандаваа Нямдаваа
Физик мэдээлэлзүйн тэнхим
АШУҮИС, БАСургууль,
Монгол улс, Улаанбаатар хот
nyamdavaa.u@mnums.edu.mn

Сандаг Оюунцэцэг
Харилцааны ур чадварын тэнхим
АШУҮИС, БАСургууль,
Монгол улс, Улаанбаатар хот
oyuntsetseg.sa@mnums.edu.mn

Эрдэнээ Алтанхундага
Физик мэдээлэлзүйн тэнхим
АШУҮИС, БАСургууль,
Монгол улс, Улаанбаатар хот
its22d005@st.mnums.edu.mn

Даваа Ивээл
Физик мэдээлэлзүйн тэнхим
АШУҮИС, БАСургууль,
Монгол улс, Улаанбаатар хот
its22d012@st.mnums.edu.mn

Хураангуй—Энэхүү судалгаагаар төгсөгчдийн сэтгэл ханамжийг олон талт 18 үзүүлэлтээр үнэлж, түүний бүтцийг тодорхойлох, мөн машин сургалтын ангиллын загварууд гүйцэтгэлийг харьцуулан судлах зорилго тавив. Судалгаанд Likert шкалд суурилсан асуумжаар 49 ажиглалтын өгөгдлийг цуглуулж, статистик болон машин сургалтын аргуудыг хослуулан боловсруулсан.

Хэмжилтийн найдвартай байдлыг Cronbach’s alpha коэффициентоор үнэлэхэд өндөр ($\alpha = 0.92$) гарсан нь үзүүлэлтүүд дотоод уялдаа сайтайг баталсан. Principal Component Analysis (PCA)-ийн үр дүнд эхний гурван үндсэн компонент нь нийт хэлбэлзлийн 64.51%-ийг тайлбарласан бөгөөд эхний компонент нь сургалтын чанар, багшийн ур чадвар, хичээлийн уялдаа зэрэг үзүүлэлтүүдийг нэгтгэсэн ерөнхий сэтгэл ханамжийн хүчин зүйл болох нь тогтоогдсон.

Машин сургалтын загваруудаас Naive Bayes болон Random Forest аргуудыг ашиглан ангилал хийж, 5-н давталтат хөндлөн баталгаажуулалтын аргаар үнэлэхэд Naive Bayes загвар илүү тогтвортой гүйцэтгэл (Accuracy = 0.9196, Kappa = 0.8568) үзүүлсэн бол Random Forest загвар тухайн өгөгдөл дээр төгс ангилалт (Accuracy = 100%) хийсэн боловч ерөнхий нөхцөлд гүйцэтгэлийн тогтвортой байдал харьцангуй бага байж болзошгүйг харуулсан.

Судалгааны үр дүнгээс харахад төгсөгчдийн 91.9% нь дунд болон өндөр түвшний сэтгэл ханамжтай байгаа нь сургалтын үйл ажиллагаа ерөнхийдөө эерэг үнэлэгдэж байгааг илтгэж байна. Иймээс төгсөгчдийн сэтгэл ханамж нь олон хүчин зүйлээс хамаарах цогц бүтэцтэй бөгөөд статистик болон машин сургалтын аргуудыг хослуулан ашиглах нь боловсролын судалгаанд үр дүнтэй болохыг энэхүү судалгаа харууллаа.

Түлхүүр үг—Төгсөгчдийн сэтгэл ханамж, Principal Component Analysis, машин сургалт, Naive Bayes, Random Forest загвар

Фишинг халдлагыг илрүүлэхаргуудын харьцуулсан судалгаа

Л.Энхцэцэг
УХИС-ийн докторант
Цэргийн төв эмнэлэг
Монгол, Улаанбаатар
enkhee10000@gmail.com

Ж.Амартайван
УБХИС-ийн магистрант
Их засаг их сургуулийн Үндэсний инженер, технологийн сургуулийн Мэдээллийн технологийн тэнхим
Монгол, Улаанбаатар
Amaraa1016@gmail.com

Л.Оюунцэцэг
Мэдээллийн технологийн тэнхим
Үндэсний инженер, технологийн сургууль
Монгол, Улаанбаатар
ORCID: 0009-0000-0500-915X

Abstract: In recent years, phishing attacks have become increasingly sophisticated, expanding beyond conventional email-based delivery to encompass a range of channels, including websites, social networks, mobile messaging services, and search engines. Phishing attacks that leverage search engines are particularly concerning due to their heightened capacity to exploit user trust. On the one hand, users commonly assume that the top-ranked results returned by search engines are reliable; adversaries capitalize on this behavioral tendency by positioning counterfeit websites at the top of search rankings or by using deceptive advertisements to mislead users. On the other hand, traditional phishing detection approaches - such as blacklisting, URL-based analysis, and signature-based techniques - are often insufficient to identify emerging attack variants or previously unobserved (zero-day) attacks. In this regard, search engine–based detection methods offer greater flexibility and near–real-time analytical capability by contrasting the actual content of a website with corresponding search results and by assessing the consistency of information as reflected in its frequency, geographic/structural placement, and textual coherence.

Key words: Phishing attack, web sites, Anti-Phishing Tools (detection methods)

Монгол хэлний үсгийн авиа танилтад машин ба гүн сургалтын загваруудын гүйцэтгэлийг харьцуулсан нь

Н. Чигцалмаа
Мэдээлэл, Компютерийн Ухааны Тэнхим
МТЭС, МУИС
Улаанбаатар, Монгол
23b1num0245@stud.num.edu.mn

Б. Солонгоо
Мэдээлэл, Компютерийн Ухааны Тэнхим
МТЭС, МУИС
Улаанбаатар, Монгол
23b1num1034@stud.num.edu.mn

Б. Нарансолонго
Мэдээлэл, Компютерийн Ухааны Тэнхим
МТЭС, МУИС
Улаанбаатар, Монгол
23b1num2631@stud.num.edu.mn

*Ц. Энхтуяа**
Мэдээлэл, Компютерийн Ухааны Тэнхим
МТЭС, МУИС
Улаанбаатар, Монгол
enkhtuya.ts@num.edu.mn

Удиртгал—Энэхүү өгүүлэлд Монгол хэлний үсгийн авиа танилтын гүйцэтгэлийг хэд хэдэн ялгаатай загвар ашиглан харьцуулсан судалгааг танилцуулж байна. Судалгаанд машин сургалтын SVM болон гүн сургалтын LSTM, CNN, RNN загварууд, тэдгээрийн хосолсон хувилбаруудыг ашиглан туршилт хийсэн. Загваруудын гүйцэтгэлийг нарийвчлал (accuracy), precision, recall болон F1-score үзүүлэлтүүдээр үнэлэв. Судалгааны гол зорилго нь бодит орчноос цуглуулсан бага хэмжээний өгөгдөл (хүүхдийн яриа) дээр Монгол хэлний авиаг таних хамгийн үр дүнтэй загварыг тодорхойлоход оршино. Туршилтын үр дүн нь Монгол хэлний авиа танилтын системийг сайжруулах, цаашдын судалгааны ажилд суурь өгөгдөл болох ач холбогдолтой.

Түлхүүр үг—Speech Recognition, Mongolian Language, Model Comparison, Machine Learning, SVM, Deep Learning, CNN, RNN, LSTM, MFCC.

“AQUA LEAF - AI” Хиймэл оюунд суурилсан тасалгааны ургамал усалгааны системийн судалгаа, хөгжүүлэлт

Баатархүүгийн Дүүрэнбаяр
Мэдээллийн Технологийн тэнхим
ШУТИС, Мэдээлэл, Холбооны Технологийн Сургууль
Монгол улс, Улаанбаатар
B231930041@must.edu.mn

Нямаагийн Даваасүрэн
Кибер аюулгүй байдлын тэнхим
ШУТИС, Мэдээлэл, Холбооны Технологийн Сургууль
Монгол улс, Улаанбаатар
davaasuren.n@must.edu.mn

Хураангуй: Манай систем нь хэрэглэгчийн оруулсан дүрс болон контекст өгөгдөлд тулгуурлан компьютер хараа, машин сургалт, ансамбль сургалтын аргуудыг ашиглан ургамлын төрөл зүйлийг таних, төлөв ба эрсдэлийг үнэлэх, хөрсний чийгийн төлөв тодорхойлох, усалгааны зөвлөмж боловсруулах боломжтой. Судалгаанд 54,000 гаруй ургамлын зураг, 32,000 гаруй хөрсний зураг ашиглагдсан. Ургамлын төрөл зүйл танихад Convolutional Neural Network (CNN) ангиллын загвар, ургамлын төлөв ба эрсдэлийн үнэлгээнд Vision Transformer (ViT), хөрсний чийг тодорхойлоход CNN–Random Forest хосолсон регресс, усалгааны зөвлөмж боловсруулахад Gradient Boosting Decision Tree алгоритмыг ашигласан. Туршилтын үр дүнгээс системийн ангиллын модуль 95%-иас дээш нарийвчлалтай ажиллаж, усалгааны регрессийн загвар R^2=0.96 үзүүлэлтэд хүрсэн. Систем нь нэмэлт мэдрэгч шаарддаггүй, хэрэглэгчийн энгийн өгөгдөлд тулгуурлан бодит хэрэглээнд нэвтрүүлэх боломжтой, өртөг багатай шийдэл юм. Энэхүү систем нь нэмэлт мэдрэгч төхөөрөмж шаардахгүй бөгөөд хэрэглэгчдэд хялбар, хүртээмжтэй шийдэл юм. Систем нь усны зохистой хэрэглээг дэмжих, ургамлын амьдрах чадварыг нэмэгдүүлэхэд ач холбогдолтой.

Түлхүүр үг - Хиймэл оюун, машин сургалт, компьютерын хараа, дүрс танилт, хөрсний чийгийн үнэлгээний алгоритм, усалгааны модуль

Байгууллагын ажилтны зан төлөвт суурилсан хиймэл оюун болон технологийн шийдлийг турших

Нугманы ИСА
Кибер аюулгүй байдлын тэнхимийн оюутан
аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар, Монгол улс
isanugman8889@gmail.com

Баярмагнайн УНДРАЛ
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн багш
ШУТИС, МХТС
Улаанбаатар, Монгол улс
undral.b@must.edu.mn

Жагаслагийн ХАЖМУХАН
Кибер аюулгүй байдлын тэнхимийн оюутан
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар, Монгол улс
khajmukhankhajku@gmail.com

Батжаргалын ХҮРЭЛБААТАР
Кибер аюулгүй байдлын тэнхимийн оюутан
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар, Монгол улс
khurelbaatarb20@gmail.com

Бат-Эрдэнийн МӨНХБАЯР
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн эрхлэгч, доктор (Ph.D.), дэд профессор
ШУТИС, МХТС
Улаанбаатар, Монгол улс
munkhbayar.b@must.edu.mn

Хураангуй— Энэхүү судалгааны ажил нь байгууллагын кибер аюулгүй байдалд хүний зан төлөв хэрхэн нөлөөлдгийг тодорхойлж, Isolation Forest машин сургалтын алгоритмд суурилсан дасан зохицох хамгаалалтын системийг бодит нөхцөлд туршсан. Судалгааны хүрээнд GoPhish хэрэгслийг AWS EC2 серверт суурилуулж нийт 29 ажилтанд хоёр үе шатт фишинг симуляцийн туршилт явуулсан. Туршилтын үр дүнд ажилтнуудын 69% нь фишинг мэйлийг нээсэн, 31% нь фишинг холбоос дарсан, 17% нь нууц үгийг маягтанд оруулснаар мэдээлэл алдсан болох нь тогтоогдов. Хэрэглэгчийн нэвтрэлтийн лог өгөгдлийг нэвтрэлтийн цаг, IP хаяг, хөтчийн төрөл, давтамж, сессийн хугацаа гэсэн таван хувьсагчаар Isolation Forest загварт оруулж боловсруулахад 91.3%-ийн нарийвчлал, 88.7%-ийн мэдрэмжит байдалтайгаар хэвийн бус нэвтрэлт илрүүлсэн бөгөөд загварын шийдвэрийн дундаж хурд 45 миллисекунд байлаа. Санал болгож буй Dynamic Policy Enforcement механизм нь эрсдэлийн оноонд тулгуурлан хандалт зөвшөөрөх, хоёр хүчин зүйлийн баталгаажуулалт (2FA) шаардах, хандалт хаах гэсэн гурван түвшний хариу үйлдлийг бодит цагт автоматаар хэрэгжүүлдэг бөгөөд загварыг сар бүр шинэ өгөгдлөөр дахин сургах замаар нарийвчлалыг байнга сайжруулна. Судалгаа нь кибер аюулгүй байдлын ХүнПроцесс-Технологийн гурван тулгуурт зарчмыг Isolation Forest алгоритм хиймэл оюунтай хослуулж, ажилтны бүтээмжид саад болохгүйгээр мэдээллийн аюулгүй байдлыг хангах боломжтойг харуулсан.

Түлхүүр үг— фишинг, Isolation Forest, хэрэглэгчийн зан төлөв, кибер аюулгүй байдал, дасан зохицох хамгаалалт

Speech Emotion Recognition via Attentive Pooling on Whisper Encoder Representations

Magvan-Erdene G.
SICT, MUST
Ulaanbaatar, Mongolia magvan-erdene@must.edu.mn

Zoljargal M.
SICT, MUST
Ulaanbaatar, Mongolia
zzoloo2727@gmail.com

Orgil J.
SICT, MUST
Ulaanbaatar, Mongolia
orgil@must.edu.mn

Abstract—Speech emotion recognition (SER) has broad applications in human-computer interaction, mental health monitoring, and multimedia analysis. While large-scale automatic speech recognition (ASR) models such as Whisper have demonstrated remarkable acoustic modeling capabilities, their potential as universal feature extractors for paralinguistic tasks remains underexplored. In this work, we conduct an empirical study of a lightweight adapter architecture that repurposes the frozen Whisper encoder for SER by appending an Attentive Statistics Pooling layer followed by a nonlinear projection head and a classifier. The pooling mechanism computes attention-weighted mean and standard deviation vectors, producing a fixed-dimensional utterance embedding that captures both spectral content and temporal dynamics without modifying the encoder weights. Experiments on IEMOCAP demonstrate competitive performance against fine-tuned baselines using only a small trainable head. Notably, a Whisper encoder further adapted for Mongolian automatic speech recognition generalizes to Mongolian emotion recognition without any emotion-labeled Mongolian data, providing preliminary evidence that emotional prosody is encoded in a largely language-invariant subspace of the Whisper encoder.

Index Terms—speech emotion recognition, Whisper, attentive statistics pooling, transfer learning, paralinguistic features, low-resource, Mongolian

LBPH болон Haar Cascade алгоритмд суурилсан Монгол хүний царай таних систем

Цагаан Баатархүү
Мэдээллийн технологи
ШУТИС, МХТС
Монгол, Улаанбаатар
baatarkhuu@must.edu.mn

Цэнд Жавзмаа
Физик, Мэдээлэл зүй
АШУИС, Био-Анагаахын сургууль,
Монгол, Улаанбаатар
javzmaa.ts@mnums.edu.mn

Оюунбилэг Бат-Энх
Мэдээллийн технологи
ШУТИС, МХТС
Монгол, Улаанбаатар
batenkh@must.edu.mn

Оюунбилэг Баттөгс
Үндэсний хөгжим
СУИС, Хөгжмийн Урлагийн Сургууль
Монгол, Улаанбаатар
battugs495@gmail.com

Хураангуй— Мэдээллийн эрин үед хиймэл оюун ухаанд суурилсан технологиуд эрчимтэй хөгжиж, нийгмийн олон салбарт өргөн хэрэглэгдэх болсон. Ялангуяа компьютер хараа болон биометрийн технологиуд нь аюулгүй байдал, нэвтрэх хяналт, иргэний бүртгэл зэрэг салбарт чухал үүрэг гүйцэтгэж байна. Монгол Улсын урт хугацааны хөгжлийн бодлого болох Алсын хараа 2050 -д дэвшүүлсэн зорилтуудтай уялдуулан үндэсний онцлогт тохирсон царай таних системийг хөгжүүлэх шаардлага улам бүр нэмэгдэж байна. Энэхүү судалгааны ажлаар Монгол хүний нүүрний өгөгдөлд суурилсан царай таних системийг боловсруулж, хэрэгжүүлсэн. Судалгаанд нийт 1000- хүний нүүрний өгөгдлийг цуглуулж, өгөгдлийн санд хадгалж, сургалт болон туршилтад ашигласан. Системийг C# програмчлалын хэл дээр боловсруулсан ба Emgu CV (OpenCV) сан ашиглан, нүүрийг илрүүлэхэд Haar cascade алгоритм, нүүрийг танихад LBPH (Local Binary Pattern Histogram) алгоритм-ыг ашигласан. Эдгээр алгоритмууд нь жижиг хэмжээний өгөгдөлд хурдан, нарийвчлалтай таних боломжийг олгодог. Туршилтын үр дүнгээс харахад уг систем нь 93.5 хувийн царай танилтын нарийвчлалтай ажилласан бөгөөд өгөгдлийн хэмжээ, гэрэлтүүлэг, зургийн чанар зэргээс танилтын хувь өөрчлөгдөж байсан.

Энэхүү судалгаа нь Монгол хүний өгөгдөлд тулгуурласан царай таних систем хөгжүүлэх боломжийг харуулсан бөгөөд цаашид өгөгдлийн санг өргөжүүлэх, өөр бусад гүн сургалтын аргуудыг ашиглах замаар системийн танилтын хувийг сайжруулах боломжтойг харуулж байна.

Хувь хүний мэдээлэл хамгаалах тухай Монгол улсын хуулийн “7.1 зүйлд Төрийн байгууллагаас бусад хуулийн этгээд, хуулийн этгээдийн эрхгүй байгууллага, хүн дараах тохиолдолд мэдээлэл цуглуулж, боловсруулж, ашиглах.” Мөн “ 7.1.3. хүнийг тодорхойлох боломжгүй болгож түүх, эрдэм шинжилгээ, урлаг, утга зохиолын бүтээл туурвих, нээлттэй өгөгдөл, статистикийн мэдээлэл бэлтгэх.” –д ашиглаж болох тухай заасан байдаг [9].

Түлхүүр үг— Царай танилт, Haar cascade, LBPH, OpenCV, Emgu CV, C#

Машин орчуулга ашиглан Монгол хэлний өгүүлбэрийн утгыг жиших загвар сургах нь (PAWS, SNLI)

Түвшинбаатар Мөнх-Орших
МУИС-МТЭС, МКУТ
Улаанбаатар, Монгол Улс
munkhorshikht@seas.num.edu.mn

Чагнаа Алтангэрэл
МУИС-МТЭС, МКУТ
Улаанбаатар, Монгол Улс
altangerel@seas.num.edu.mnZ

Хураангуй — Энэхүү судалгаанд нэг ижил архитектуртай (XLM-R ба нэг бинар ангилах толгойтой) загварын хувьд өгөгдөж буй сургалтын мэдээллийн хэлбэр нь хэл дамнасан нөхцөлд шийдвэрийн хилд хэрхэн нөлөөлж байгааг судлав. Үүний тулд (i) даалгаварт шууд нийцсэн, утга ижил өгүүлбэрийн хос дээр суурилсан чиглүүлэгдсэн мэдээлэл агуулсан PAWSonly горим болон (ii) өгүүлбэрийн хосын утгын нэмэлт сургалтын мэдээлэл агуулсан PAWS+SNLI горимыг харьцуулсан. Үр дүнгээс PAWS-only загвар нь PAWS жишиг сорил дээр өндөр үзүүлэлттэй байсан боловч Монгол хэлний хяналттай үнэлгээнд (MN10000-Synth) утга ялгаатай ангиллыг бараг ялгахгүй болж, ихэнх хосыг утга ижил гэж хэт таамаглах хэв шинж ажиглагдав. Харин SNLI-оос үүсгэсэн нэмэлт сургалтын мэдээлэл ашиглахад энэ хэлбийлт буурч, Монгол хэлний үнэлгээн дэх алдааны тархалт илүү тэнцвэртэй болсон. Гэсэн хэдий ч эдгээр үр дүн нь шалтгаант хатуу баталгаа биш бөгөөд хэл дамнасан ерөнхийшүүлэх чадварт сургалтын мэдээллийн хэлбэр нөлөөлж болохыг харуулсан ажиглалтын шинжтэй дүгнэлт болно.

Түлхүүр үг— paraphrase identification, cross-lingual transfer, decision boundary, calibration, XLM-R

Дөрвөн сэнстэй дэлгэгддэг дроны загвар болон хөгжүүлэлт

Э. Бат-Оргил
МУИС-МТЭС, Электроник, Холбооны Инженерчлэлийн тэнхим
Улаанбаатар, Монгол
batorgil.e1224@gmail.com

Б. Мөнхтулга
МУИС-МТЭС, Электроник, Холбооны Инженерчлэлийн тэнхим
Улаанбаатар, Монгол
munkhtulga@num.edu.mn

Хураангуй — Энэхүү судалгааны ажилд Кансат тэмцээний буцагч төрлийн даалгаварт зориулан бие даан боловсруулсан, GPS-д суурилсан автомат газардалтын системтэй дөрвөн сэнстэй (дөрвөн сэнстэй нисэгч) нисэгчийн загвар болон хэрэгжүүлэлтийг танилцуулж байна. Уг нисэгчийн гол давуу тал нь цойлуурын нарийн цилиндрт багтахаар эвхэгдэж, агаарт гарсны дараа автоматаар дэлгэгддэг өвөрмөц механик бүтцтэй байхын зэрэгцээ GPS-ийн байршлыг ашиглан алсын удирдлагагүйгээр бие даан заасан цэгт очиж газардах чадвартай байдаг. Системийн төв удирдлагад HAKRC F722V2 нислэгийн хянагч, байршил тогтооход UBLOX NEO-M8N GPS модуль, нислэгийн горим шилжилтэд Arduino Nano микроконтроллёрыг ашигласан бөгөөд INAV программ хангамжийг гол программ хангамж болгон сонгов. Механик бүтцийн бүх эд анги, дэлгэгдэх механизм зэргийг PLA материалаар 3D хэвлэх аргаар бие даан үйлдвэрлэв. Нисэгчийн нийт жин 370– 400 г бөгөөд батарей 12 минут ажиллах чадалтайг тооцоолсон. Туршилтын үр дүнд дэлгэгдэх механизм 34%-ийн моторын хүчтэйгээр тогтвортой ажиллаж, GPS байршлын нарийвчлал 6.3 м байгааг тогтоов.

Түлхүүр үг — Кансат, GPS, автомат газардалт, дөрвөн сэнстэй нисэгч, INAV, дэлгэгдэх механизм, Arduino Nano, HAKRC F722V2

Өөрийгөө хөгжүүлэх ухаалаг систем: Хиймэл оюуны шинжилгээ ба мэргэжилтний зөвлөгөөний хосолсон загвар

Доржгочоо Сарангэрэл
Мэдээллийн технологийн тэнхим
ШУТИС МХТС
Монгол улс, Улаанбаатар
saran@must.edu.mn

Батхүү Өсөх-Ирээдүй
Мэдээллийн технологийн тэнхим
ШУТИС МХТС
Монгол улс, Улаанбаатар
usukhireedui34@gmail.com

Хураангуй—Энэхүү өгүүлэлд хувь хүний өөрийгөө хөгжүүлэх үйл явцыг дэмжих зорилгоор хиймэл оюунд суурилсан, мэргэжилтний зөвлөгөөтэй хосолсон ухаалаг вэб систем хөгжүүлэхийг санал болгосон. Уг систем нь хэрэглэгчийн өдөр тутмын дадал, зорилго, ахиц дэвшлийн өгөгдлийг цуглуулж, хиймэл оюун ашиглан шинжилгээ хийж, хувь хүнд тохирсон зөвлөмж боловсруулах боломжийг бүрдүүлсэн. Мөн системд хэрэглэгч, мэргэжилтэн, админ гэсэн гурван түвшинд нэвтрүүлж, мэргэжилтний зөвлөгөөг хиймэл оюуны шинжилгээтэй уялдуулсан шийдэл хэрэгжүүлсэн. Үр дүнгийн үнэлгээг хэрэглэгчийн идэвхжил, зорилгын биелэлт, зөвлөмжийн үр ашиг зэрэг үзүүлэлтүүдээр хэмжиж, уламжлалт системүүдтэй харьцуулахад илүү үр дүнтэй болохыг тогтоов.

Түлхүүр үг— өөрийгөө хөгжүүлэх систем, хиймэл оюун, хувь хүнд тохирсон зөвлөмж, дадал мөрдөлт

Халдлага илрүүлэх системд ашиглах санг харьцуулан судлах нь

Н.Угтахбаяр, П.Баясгалан, Д.Бямбажав, Б.Өсөхбаяр*
Мэдээллийн Технологи, Электроникийн Сургууль, Монгол Улсын Их Сургууль Электроник, Холбооны Инженерчлэлийн Тэнхим bayasgalan.p@num.edu.mn

Хураангуй—Технологийн салбарын хурдацтай хөгжлөөс үүдэлтэй кибер халдлагын төрөл, тоо, хэмжээ болон төвөгтэй байдал (complexity) нэмэгдэж байгаатай холбоотойгоор сүлжээний халдлага илрүүлэх системийн хэрэгцээ, ач холбогдол улам бүр өсөж байна. Сигнатурт суурилсан халдлага илрүүлэх систем шинэ төрлийн “zero day” халдлагыг илрүүлэх боломж улам бүр хязгаарлагдмал болсон тул машин сургалт, хиймэл оюунд суурилсан халдлага илрүүлэх технологи, аргачлал сүүлийн жилүүдэд өргөн хэрэглэгдэх болсон. Бид энэхүү судалгааны ажлаар өөрсдийн гаргаж авсан сан болон бусад түгээмэл хэрэглэгддэг нийтэд нээлттэй санг ашиглан машин сургалтын алгоритмуудын гүйцэтгэлийг бодит сүлжээний орчинд харьцуулан үнэлэх зорилго тавин ажиллалаа. Бид судалгааны ажилдаа энэ чиглэлийн судлаачдын судалгааны ажилдаа түгээмэл ашиглаж буй, тухайн судалгааны ажлуудын хувьд сайн үр дүн үзүүлсэн Random Forest алгоритмыг ашиглан халдлага илрүүлэх системийн моделийг сургаж UNSW-NB15 өгөгдлийг ашиглан үр дүнгээ туршсан. Бидний судалгааны үр дүнд халдлага илрүүлэх системийн гүйцэтгэл нь зөвхөн алгоритмаас бус, ашиглаж буй сангаас ихээнэх хамаардаг болохыг туршилтаар харууллаа. Туршилтийн үр дүнд хоёр өөр сангаар сурсан моделууд халдлагыг 9-12 хувийн зөрүүтэй танилт хийж байгаа нь ажиглагдаж байна.

Түлхүүр үг— Кибер аюулгүй байдал, сүлжээний халдлага илрүүлэх систем, машин сургалт, хиймэл оюун

YOLO-д суурилсан хог хаягдлын автомат ангиллын аргын судалгаа

Дашхорлоогийн Хулан
Мэдээлэл, компьютерын ухааны тэнхим
МУИС, МТЭС, МКУТ
Монгол улс Улаанбаатар
email: hulanvn@gmail.com

Ганболдын Уганбаяр
Мэдээлэл, компьютерын ухааны тэнхим
МУИС, МТЭС, МКУТ
Монгол улс Улаанбаатар
email: uuganbayar.g@num.edu.mn

Хураангуй—Хог хаягдлыг зөв ангилан ялгах нь байгаль орчны бохирдлыг бууруулах, дахин боловсруулах үйл явцын үр ашгийг нэмэгдүүлэхэд чухал ач холбогдолтой ойлголт. Гэвч бодит нөхцөлд ихэвчлэн гар ажиллагаанд тулгуурласан байдаг тул алдаа ихтэй, тогтворгүй байдаг. Иймээс энэхүү судалгаанд компьютерын хараа болон гүн сургалтын аргад суурилсан YOLO төрлийн объект таних загварыг ашиглан хог хаягдлын төрлийг автоматаар ангилах боломжийг судлав. Судалгаанд өөрийн бүрдүүлсэн болон нээлттэй эхийн өгөгдлийг ашиглан өгөгдлийн багцыг үе шаттайгаар (500 → 1300 → 5000 зураг) өсгөж, загвар сургалт болон туршилтыг хийж гүйцэтгэсэн. Туршилтын явцад хэд хэдэн YOLO архитектурын загваруудыг харьцуулан үнэлж, нарийвчлал, илрүүлэлтийн тогтвортой байдал болон таамаглах (inference) хурдыг хамтад нь авч үзэн судалсан. Туршилтын үр дүнд сонгогдсон загвар нь mAP50 ≈ 0.87 үзүүлэлттэйгээр хог хаягдлыг ангилах нарийвчлалтай байсан. Эдгээр үр дүн нь YOLO-д суурилсан автомат ангиллын аргыг тогтвортой хог хаягдлын менежментэд ашиглах боломжтойг илтгэж байгаа бөгөөд цаашид бодит төхөөрөмж болон ухаалаг хогийн савны загвар системд нэгтгэн хэрэгжүүлэх боломжийг нээж байна.

Түлхүүр үг—Компьютерын хараа, объект таних, YOLO, хог хаягдлын ангилалт, гүн сургалт

IoT төхөөрөмжүүдийн сүлжээний дүрэмд суурилсан халдлага илрүүлэх систем боловсруулах

Пүрэвдоржийн БУЯНЦОГ
Кибер аюулгүй байдлын тэнхимийн оюутан
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар, Монгол улс
B221870534@must.edu.mn

Баярмагнайн УНДРАЛ
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн багш
ШУТИС, МХТС
Улаанбаатар, Монгол улс
undral.b@must.edu.mn

Алтансүхийн ТЭМҮҮЛЭН
Кибер аюулгүй байдлын тэнхимийн оюутан
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар, Монгол улс
B221870537@must.edu.mn

Бат-Эрдэнийн МӨНХБАЯР
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн эрхлэгч,
доктор (Ph.D.), дэд профессор
ШУТИС, МХТС
Улаанбаатар, Монгол улс
munkhbayar.b@must.edu.mn

Хураангуй—Эд юмсын интернет (IoT) сүлжээнд чиглэсэн тагнан судлах халдлагыг илрүүлэх нь нөөц хязгаарлагдмал орчинд тулгамдсан асуудал хэвээр байна. Уламжлалт дүрэмд суурилсан системүүд босго утгуудыг гараар тохируулдаг бол машин сургалтын загварууд тооцооллын хувьд IoT гарцанд байрлуулахад хэт өртөгтэй байдаг. Энэхүү ажлаар эдгээр хоёр хандлагын дунд байрлах шийдэл болгон зөвхөн пакетийн толгой хэсгийн мэдээлэлд тулгуурлан 14 тайлбарлах боломжтой дүрмийн хүрээнд скан болон түүнтэй төстэй зан төлөв бүхий 5 төрлийн халдлагыг илрүүлдэг, хөнгөн жинтэй, дүрэмд суурилсан IDS-ийг боловсруулав. Ажиллах үеийн машин сургалтын дүгнэлт шаардахгүй энэ систем нь босго утгуудаа машин сургалтаас зээлсэн торон хайлт болон давхарласан хөндлөн баталгаажуулалтаар өгөгдөлд тулгуурлан оновчилдог бөгөөд дүрмүүд нь бүрэн тайлбарлах боломжтой, нягтлан шалгаж болдог давуу талтай. Gotham 2025 өгөгдлийн багц дээр давхарласан 5 хэсэгт хөндлөн баталгаажуулалтын үр дүнд дундаж нарийвчлал 85.3%, дундаж сэргээн олох 93.3%, дундаж F1 оноо 88.8% үзүүлсэн нь санал болгож буй аргачлал IoT орчинд хэрэглэхэд үр дүнтэй болохыг харуулж байна.

Түлхүүр үг—IoT, халдлага илрүүлэх систем, дүрэмд суурилсан, босго оновчлол, тагнан судлах, гарцан аюулгүй байдал, хөндлөн баталгаажуулалт

Зургаас автомашины өнгөний ангилал

1st Хайдав Дармаасүрэн
Монгол Улсын Их Сургууль
Мэдээллийн Технологи,
Электроникийн Сургууль
Мэдээлэл, Компьютерийн Ухааны Тэнхим
Монгол, Улаанбаатар
kh.darmaasuren@gmail.com

2nd Ганболд Уганбаяр
Монгол Улсын Их Сургууль
Мэдээллийн Технологи,
Электроникийн Сургууль
Мэдээлэл, Компьютерийн Ухааны Тэнхим
Монгол, Улаанбаатар
uuganbayar.g@num.edu.mn

Хураангуй—Сүүлийн жилүүдэд Компьютер хараа, Гүн сургалт, Машин сургалтын арга техник, технологи хөгжихийн хэрээр зураг болон бичлэг дэх мэдээллийг таних, ангилах асуудал нь хэрэглээнд өргөн нэвтэрч байна. Үүнээс тээврийн хэрэгслийг автоматаар таних технологи нь замын хөдөлгөөний хяналт, хот болон хөдөө орон нутгийн камерын систем зэрэг олон салбарт чухал үүрэг оролцоотой байна. Гэвч автомашины загвар таних асуудал нь зөвхөн объект илрүүлэхээс илүү төвөгтэй бөгөөд өөр өөр үйлдвэрлэгчийн автомашинууд хоорондоо төстэй харагдах тохиолдол элбэг бөгөөд нэг үйлдвэрлэгчийн өөр өөр загварууд ч ойролцоо харагдах тохиолдол цөөнгүй байдаг. Тиймээс бүтэн зургаас тээврийн хэрэгслийг ангилах үед гэрэл сүүдэр, хажуугийн өөр биет зэргээс хамааран танилтын чанар буурах эрсдэлтэй байдаг. Энэхүү судалгаанд зураг дээрээс эхлээд YOLO загварын авто машин илрүүлэлтийг ашиглан тээврийн хэрэгслийг илрүүлж, дараа нь илэрсэн хэсэг дээр Resnet-50 backbone дээр суурилан тээврийн хэрэгслийн төрлийг ангилна. Судалгааны явцад авто машин илрүүлэх болон ангиллын өгөгдлийн багцыг хослуулан бэлтгэж, илрүүлэлтийн хязгаарлагч хүрээний координат дээр үндэслэн автомашины ангилах өгөгдлийн багцыг үүсгэн сургалтад ашигласан. Туршилтын үр дүнгээс харахад автомашиныг бүтэн зургаас тээврийн хэрэгслийн төрлийг тодорхойлохоос илүүтэйгээр эхлээд зөв илрүүлж, дараа нь түүн дээрээ төрлийг ангилах нь илүү оновчтой болох нь ажиглагдсан.

Түлхүүр үг—Make and model classification, Computer Vision, Deep Learning, YOLO, ResNet-50

Захын тооцооллыг ашиглан зүйлсийн интернэтийн өгөгдөл дамжуулалтын ачааллыг

Баяраагийн ЭРДЭНЭПҮРЭВ
Кибер аюулгүй байдлын тэнхимийн оюутан, Кибер аюулгүй байдлын судалгааны баг
ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль
Монгол Улс, Улаанбаатар хот
B231860015@must.edu.mn

Отгонбаярын БАЯРМАА
Кибер аюулгүй байдлын тэнхимийн оюутан, Кибер аюулгүй байдлын судалгааны баг
ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль
Монгол Улс, Улаанбаатар хот
B231860017@must.edu.mn

Золбаярын ТЭМҮҮЛЭН
Кибер аюулгүй байдлын тэнхимийн оюутан, Кибер аюулгүй байдлын судалгааны баг
ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль
Монгол Улс, Улаанбаатар хот
temvvlenzolbayar7@gmail.com

Баярмагнайн УНДРАЛ
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн багш
ШУТИС, МХТС
Улаанбаатар, Монгол улс
undral.b@must.edu.mn

Бат-Эрдэнийн МӨНХБАЯР
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн эрхлэгч, доктор (Ph.D.), дэд профессор
ШУТИС, МХТС
Улаанбаатар, Монгол улс
munkhbayar.b@must.edu.mn

Нороврэнчиний САРАНГОО
Кибер аюулгүй байдлын тэнхимийн оюутан, Кибер аюулгүй байдлын судалгааны баг
ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль
Монгол Улс, Улаанбаатар хот
B231860002@must.edu.mn

Хураангуй—Зүйлсийн интернэтийн системд үүсэх өгөгдөл дамжуулалтын ачааллыг захын тооцооллын аргаар бууруулах боломжийг судалж, туршилтаар үнэлэв. Зүйлсийн интернэт төхөөрөмж, захын тооцоолол, үүлэн тооцоолол гэсэн гурван шатлалт архитектур бүхий хэсгээс бүрдэнэ. Уг механизмд өгөгдлийг захын тооцоолол дээр урьдчилан шүүх, цэвэрлэх, нэгтгэх, шахах боловсруулалтуудыг хийж, зөвхөн шаардлагатай мэдээллийг үүлэн тооцоолол руу дамжуулах аргыг ашигласан. Туршилтыг GNS3 болон виртуал орчинд шууд дамжуулалт болон захын тооцооллыг харьцуулан хэрэгжүүлсэн. Захын тооцоолол ашигласан үед үүлэн тооцоолол руу дамжуулах өгөгдлийн хэмжээ 3221 пакетаас 404 пакет болж буурч, ойролцоогоор 87%-иар багассан. Дамжуулалтын хоцролт (RTT) 21.1638 ms-ээс 21.0433 ms болтол буурч 0.57%-ийн сайжрал үзүүлсэн бөгөөд өгөгдлийн дундаж дамжуулалтын хурд 6–9 kb/s-ээс 1–3 kb/s болж, зурвасын өргөний хэрэглээ 73%-иар буурсан.

Түлхүүр үг—Захын тооцоолол, Зүйлсийн интернэт, Үүлэн тооцоолол, хоцролт, сүлжээний ачаалал.

Бодит цагийн халдлагын дүрслэлтэй үүлэн орчинд суурилсан honeypot систем хөгжүүлэх нь

Пүрэвдоржийн ДАВААДАМБЭРЭЛ
Кибер аюулгүй байдлын тэнхимийн оюутан, Кибер аюулгүй байдлын судалгааны баг
ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль
Монгол Улс, Улаанбаатар хот
B221870524@must.edu.mn

Батсайханы ДАШНЯМ
Кибер аюулгүй байдлын тэнхимийн оюутан, Кибер аюулгүй байдлын судалгааны баг
ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль
Монгол Улс, Улаанбаатар хот
B221870532@must.edu.mn

Энхболдын БАТЗОРИГ
Кибер аюулгүй байдлын тэнхимийн оюутан, Кибер аюулгүй байдлын судалгааны баг
ШУТИС, Мэдээлэл Холбооны Технологийн Сургууль
Монгол Улс, Улаанбаатар хот
B221870504@must.edu.mn

Баярмагнайн УНДРАЛ
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн багш
ШУТИС, МХТС
Улаанбаатар, Монгол улс
undral.b@must.edu.mn

Бат-Эрдэнийн МӨНХБАЯР
Кибер аюулгүй байдлын судалгааны ба
Кибер аюулгүй байдлын тэнхимийн эрхлэгч,
доктор (Ph.D.), дэд профессор
ШУТИС, МХТС
Улаанбаатар, Монгол улс
munkhbayar.b@must.edu.mn

Abstract —Энэхүү судалгааны ажилд үүлэн орчинд (AWS EC2) суурилсан honeypot систем байршуулан бодит кибер халдлагуудыг татаж мэдээлэл цуглуулах, улмаар уг өгөгдлийг бодит цаг хугацаанд дүрслэн хянах хяналтын самбарын загвар боловсруулсан. Corwie болон Dionaea honeypot системүүдийг ашиглан SSH/Telnet болон олон протоколын халдлагын лог бүртгэж, Logstash/Kafka-аар дамжуулан Elasticsearch-д хадлагалсан.

Max Mind GeolP2 өгөгөдлийн сан ашиглан халдагчийн газарзүйн байрлалыг газрыг зурагт дүрслэх боломжтой болгож, Grafana платформ дээр нэгдсэн хяналтын самбар болгон хагжүүлэхээр төлөвлөсөн.

Туршилтын 7 хоногийн хугацаанд нийт 1,500 гаруй халдлага бүртгэгдэж, SSN brute-force 56%, FTP сканнинг 20%, Telnet нэвтрэх оролдлого 13%-ийг тус тус эзэлсэн.

Судалгааны үр дүнд нь байгууллагуудад кибер халдлагыг ыолит цагт илрүүлэх, хариу арга хэмжээг хурдасгах боломжийг олгоно.

Index Terms—Honeypot, кибер аюулгүй байдал, ELK, Stack, AWS EC2, Grafana, GeoIP, бодит цагийн хяанлт.

Лимбэний дуугаралтын цахилгаан дохионы судалгаа

Цагаан Баатархүү
Мэдээллийн технологи ШУТИС, МХТС
Монгол, Улаанбаатар
baatarkhuu@must.edu.mn

Цэнд Жавзмаа
Физик, Мэдээлэл зүй
АШУИС, , Био-Анагаахын сургууль,
Монгол, Улаанбаатар
javzmaa.ts@mnums.edu.mn

Оюунбилэг Баттөгс
Үндэсний хөгжим
СУИС, Хөгжмийн Урлагийн Сургууль
Монгол, Улаанбаатар
battugs495@gmail.com

Оюунбилэг Бат-Энх
Мэдээллийн технологи
ШУТИС, МХТС
Монгол, Улаанбаатар
batenkh@must.edu.mn

Хураангуй—Энэхүү судалгааны ажлын зорилго нь лимбэний дуугаралтыг цахилгаан дохио хэлбэрт шилжүүлэн боловсруулж, түүний үндсэн шинж чанаруудыг тодорхойлон, дохио боловсруулалтын аргуудыг ашиглан шинжилгээ хийхэд оршино. Судалгаа нь “Үндэсний их найрал хөгжмийн бүтэц бүрэлдэхүүн, дуугаралтын хүчний харьцааг боловсронгуй болгох нь” сэдэвт төслийн хүрээнд хийгдсэн бөгөөд уг төслөөс гарсан үр дүнгээс лимбэ хөгжмийн дуугаралтын цахилгаан дохионы үзүүлэлтүүдийг сонгон авч ашигласан болно.

Судалгаанд Соёл урлагийн их сургуулийн II–IV дамжааны нийт 28 оюутан өөрсдийн хөгжмийн зэмсгээр оролцож, дуугаралтын бичлэг хийлгэсэн. Мөн Үндэсний урлагийн их театрын эвэр бүрээч болон цордон бүрээч нар хамтран оролцсон. Энэхүү судалгааны хүрээнд лимбэний дуугаралтыг гоцлол болон хоршил гэсэн хоёр нөхцөлд авч үзэн, цахилгаан дохионы шинжилгээний үр дүнгүүдийг харуулсан. Судалгааны явцад хулсан болон хуванцар хоёр төрлийн лимбэ ашигласан ба лимбэ тус бүрийн хувьд бага наймцын си бемолиос ÷II наймцын ля хүртэлх нийт 14 эгшгийн судалгааг хийсэн.

Судалгааны үр дүн нь цаашид хөгжмийн автомат таних систем, дууны синтез, мөн соёл урлагийн салбарт дуугаралтын чанарын үнэлгээ хийхэд өргөнөөр ашиглагдах боломжтой. Түүнчлэн лимбэ хөгжмийн дуугаралтын онцлогийг цахилгаан дохионы түвшинд судалснаар хөгжмийн зэмсгийн дуугаралтыг илүү нарийвчлалтай загварчлах, ангилах алгоритмуудыг хөгжүүлэхэд чухал хувь нэмэр оруулна.

Түлхүүр үг—дууны дохио, цахилгаан дохио, лимбэний дуугаралт, үндсэн өнгө, үндсэн давтамж, дагалдах өнгө, давтамж

Нисэх буудлын хөөрч, буух зурваст гаднын биетийг илрүүлэх mmWave ба X-band радарын системийн судалгаа

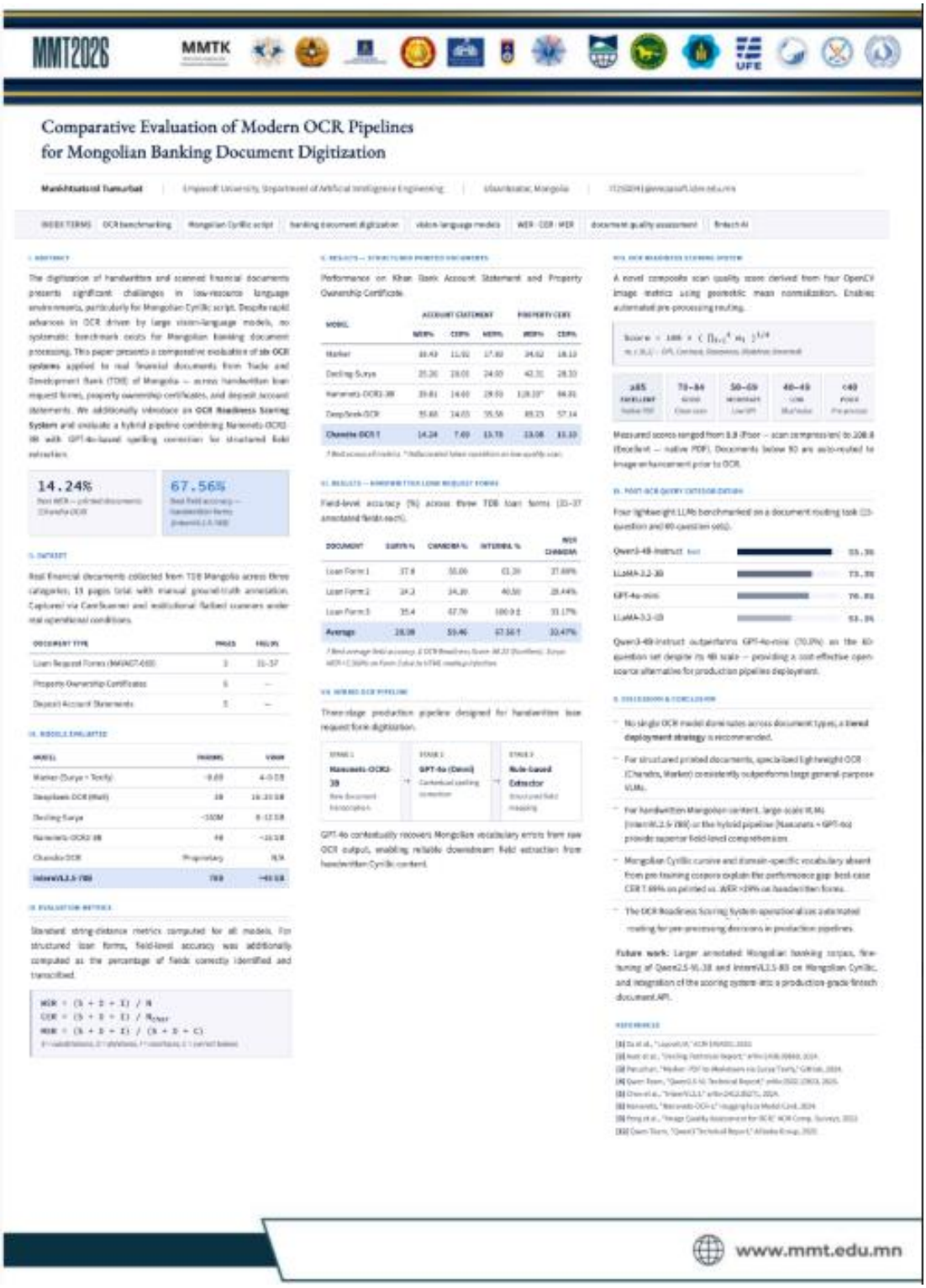
Эрхэмбаатарын Нарантуяа
Холбооны инженерчлэлийн тэнхим
Мэдээлэл Холбооны Технологийн Сургууль
Шинжлэх Ухаан Технологийн Их Сургууль
narantuya@must.edu.mn

2nd Баянмөнхийн Сэлэнгэсүрэн
Холбооны инженерчлэлийн тэнхим
Мэдээлэл Холбооны Технологийн Сургууль
Шинжлэх Ухаан Технологийн Их Сургууль
selenge7171@gmail.com

Хураангуй—Нисэх буудлын хөөрч буух зурвас (ХБЗ) дээрх гаднын биетүүд (Foreign Object Debris, FOD) нь нислэгийн аюулгүй байдалд ноцтой эрсдэл учруулдаг. Энэ судалгаанд FOD илрүүлэх радарын системүүд, тэдгээрийг Чингис Хаан олон улсын нисэх буудал (ОУНБ)-д нэвтрүүлэх боломжийг судлан, mmWave болон X-band-ыг хослуулсан FOD радарыг шийдэл болгож байна. FMCW ба MIMO технологиудыг нэгтгэснээр объектын байршил, хурд тодорхойлох, олон объект ялгах чадварыг нэмэгдүүлдэг. Симуляцийн үр дүн нь хосолсон (hybrid) систем FOD-оос үүсэх эрсдлийг бууруулж, ХБЗ-ын аюулгүй байдлыг сайжруулах боломжтойг харуулж байна.

Түлхүүр үг—FOD, радар систем, mmWave радар, X-band радар, Нейман-Пирсоны загвар

Comparative Evaluation of Modern OCR Pipelines for Mongolian Banking Document Digitization



Радарын хөндлөн огтлолын коэффициентийг бууруурууулах мета-гадаргуугийн загварчлал

Дашбаатар АНУДАРЬ[1], Баттулга УНДАРМАА[1], Дашгянцан САНЖСҮРЭН[1],
Баярсайхан ПҮРЭВЦЭРЭН[1], Пүрэвдорж УУГАНБАЯР[1a]

[1]Холбооны салбар, Мэдээлэл, Холбооны Технологийн Сургууль
Шинжлэх Ухаан Технологийн Их Сургууль
Улаанбаатар, Монгол улс
anudaridashbaatar@gmail.com[1], undarmaa.battulga12@gmail.com, sanjsurend@gmail.com ,
purevtseren_b@must.edu.mn[1], p.uuganbayar@must.edu.mn[1a]

Хураангуй — Энэхүү өгүүлэлд C болон X зурваст ажиллах мета-гадаргуугийн долгион шингээгч болон долгион сарниулагч нэгж элементийн загварыг танилцууллаа. Симуляцийн үр дүнгээс харахад шингээгч бүтэц 6 GHz орчимд $\boldsymbol{S_{11}} = -\mathbf{27}$dB утгатай гарч, өндөр шингээлтийн шинж чанартай болохыг харуулсан. Мөн шингээгч бүтэц C зурваст ойролцоогоор 175°, сарниулагч бүтэц X зурваст ойролцоогоор 66° фазын шилжилттэй гарсан нь долгионы шингээлт болон сарниулалтын горимыг удирдах боломжтойг илтгэж байна. Иймээс санал болгож буй мета-гадаргуугийн нэгж элементүүд нь долгионы тархалтын чиглэл болон шингээлтийн шинж чанарыг хянах боломжтой.

Түлхүүр үг— Direction of Arrival (DOA) тооцоолол, нэгж элемент, радарын хөндлөн огтлолын коэффициен

Монгол үндэсний хээ угалзын зарим дүрсийг математик загварчлалаар вектор график болгон зурах

П.Гантуяа
Мэдээлэл, Компьютерын Ухааны Тэнхим
Монгол Улсын Их Сургууль, Мэдээллийн Технологи, Электроникийн Сургууль
Монгол улс, Улаанбаатар хот

Г. Багаболд*
Мэдээлэл, Компьютерын Ухааны Тэнхим
Монгол Улсын Их Сургууль, Мэдээллийн Технологи, Электроникийн Сургууль
Монгол улс, Улаанбаатар хот
bagabold.g@num.edu.mn

А.Мөнхзул
Мэдээлэл, Компьютерын Ухааны Тэнхим
Монгол Улсын Их Сургууль, Мэдээллийн Технологи, Электроникийн Сургууль
Монгол улс, Улаанбаатар хот
munkhzul.a@num.edu.mn

Хураангуй— Монголчуудын уламжлалт соёлд хээ угалз нь маш чухал байр суурь эзэлж, олон зууны турш эдлэл хэрэгсэл, хувцас, чимэглэлийн зүйлс дээр өргөн хэрэглэгдсээр ирсэн. Хээ угалз нь зөвхөн гоёмсог дүрслэл бус, утга агуулга, бэлгэдэл, ёс заншлыг илэрхийлдэг гүн гүнзгий утга бүхий соёлын илэрхийлэл юм. Монголчуудын өдөр тутмын амьдралд хээ угалзыг дизайны салбарууд тэр дундаа график дизайн, эд зүйлсийг гоёж, чимэглэх зэрэг хэрэглээний урлагт ихээхэн хэрэглэсээр ирсэн бөгөөд түүгээр дамжуулан соёлын элемент, утга бэлгэдэл зэргийг онцгойлон авч, тухайн цаг үеийн оюун санаа, үнэт зүйлсийг илтгэдэг онцлогтой байжээ.

Энэ судалгаагаар бид Монголчуудын хээ угалзны зарим дүрсийг математик загварчлал ашиглан чанартай вектор дүрс бүтээх арга эрэлхийлсэн юм.

Мөн энэхүү ажлаар Монголын дүрс тэмдгийн тэр дундаа хээ угалзын дүрсийн элементүүдийг орчин үеийн дизайнд ашиглахад үр дүнтэй, чанартай, алдаагүй зөв бүтээгдсэн, хэмжээ бага эзлэх, цааш үргэлжүүлэн зурж дүрслэх боломжтой загвар гарган авахыг зорьж, аналитик геометр аргаар туршин судаллаа.

Түлхүүр үг— аналитик геометр, хээ, вектор график үүсгэх

5G - ийн дээд зурваст ажиллах тархинаас мэдрэлийн өгөгдлийг өндөр хурдаар дамжуулах суулгац антены загварчлал

Батбаатар ЗАГАРСҮРЭН[1], Болормаа ОЮУН-ЭРДЭНЭ[1], Баярсайхан ПҮРЭВЦЭРЭН[1], Пүрэвдорж УУГАНБАЯР[1a]
[1]Холбооны салбар, Мэдээлэл, Холбооны Технологийн Сургууль
Шинжлэх Ухаан Технологийн Их Сургууль Улаанбаатар, Монгол улс
zagarsuren0109@gmail.com[1], oune0789@gmail.com[1], purevtseren_b@must.edu.mn[1], p.uuganbayar@must.edu.mn[1a]

Хураангуй - Энэхүү өгүүлэлд Тархи-Компьютерийн Интерфейс (BCI/BMI) системд зориулсан тархины суулгац хавтгай урвуу F (PIFA) хэлбэрийн антены загварыг танилцууллаа. Антеныг 5G-ийн 25–26 ГГц (mmWave) давтамжийн зурваст 2.2 mm x 3.9 mm x 0.324 mm маш жижиг хэмжээтэй загварчлав. Антены гүйцэтгэлийг бодит орчинд үнэлэхийн тулд тархины биологийн бүтцийг төлөөлөх 7 давхар тархины загварыг ашиглан симуляцийг хийж гүйцэтгэв. Симуляцийн үр дүнгээс харахад антены ойлтын коэффициентийн утга -36.85 дБ-ээс -13 дБ байгаа нь $S11 < -10$ дБ байх техникийн шаардлагыг хангаж байна. Мөн Тусгай шингээлтийн харьцаа (Specific Absorption Rate, SAR)-ны утга 1.57^{-5}Вт/кг гарсан нь олон улсын зөвшөөрөгдөх хязгаар (1.6 Вт/кг)-аас хэд дахин бага байгаа бөгөөд хэвийн ажлын горимд аюулгүй гэдгийг харуулж байна. Иймээс санал болгож буй антен нь BCI системд тархины нейроны мэдээллийг бодит цагийн горимд найдвартай, тогтвортой дамжуулах боломжтой юм.

Түлхүүр үг— Тархи-Компьютерийн Интерфейс (BCI), 5G mmWave, Хавтгай урвуу F (PIFA) антен, Мэдрэлийн өгөгдөл дамжуулалт.

ЕБС-ийн багш, сурагчдын кибер аюулгүй байдлын мэдлэгийг тодорхойлох судалгаа

Баярмагнайгийн Ундрал
Кибер аюулгүй байдлын тэнхим
ШУТИС, МХТС
Улаанбаатар, Монгол улс undral.b@must.edu.mn

Хураангуй— Сүүлийн жилүүдэд боловсролын салбарт мэдээллийн технологийн хэрэглээ эрчимтэй нэмэгдэж, багш, сурагчдын дижитал орчны оролцоо өссөнтэй холбоотойгоор кибер аюулгүй байдлын эрсдэл нэмэгдэж байна. Энэхүү судалгаагаар ерөнхий боловсролын сургуулийн багш, сурагчдын кибер аюулгүй байдлын мэдлэг, хандлага, зан үйлийг (KAB) тодорхойлж, тэдгээрийн хоорондын хамаарлыг шинжилсэн. Судалгаанд HAIS-Q аргачлалд суурилсан асуулга ашиглан 398 сурагч, 387 багшийг хамруулсан. Үр дүнгээс харахад сурагчид багш нараас илүү өндөр үзүүлэлттэй байсан бөгөөд мэдлэг, хандлага, зан үйлийн хооронд эерэг хамаарал илэрсэн. Судалгааны үр дүнд үндэслэн боловсролын салбарт кибер аюулгүй байдлын сургалтыг системтэйгээр хэрэгжүүлэх шаардлагатайг тодорхойллоо.

Түлхүүр үг— Кибер аюулгүй байдал, ерөнхий боловсролын сургууль, мэдлэг, хандлага, зан үйл, HAIS-Q загвар, багш, сурагч, кибер аюулгүй байдлын боловсрол, сургалтын хөтөлбөр

AI RED teaming аргачлалд суурилсан хиймэл оюуны суурь загварын эмзэг байдлын судалгаа

Нямаагийн ДАВААСҮРЭН[1]
Кибер аюулгүй байдлын тэнхим
ШУТИС, МХТС
Улаанбаатар, Монгол улс
davaasuren.n@must.edu.mn

Алтайгийн ЛХАГВАОЧИР[2]
Кибер аюулгүй байдлын тэнхим
ШУТИС, МХТС
Улаанбаатар, Монгол улс
lkhagvaochir@must.edu.mn

Батбаатарын ПҮРЭВСҮРЭН[2]
Кибер аюулгүй байдлын тэнхим
ШУТИС, МХТС
Улаанбаатар, Монгол улс
purevsuren@must.edu.mn

Ганболдын НАСАНБАТ[2]
Кибер аюулгүй байдлын тэнхим
ШУТИС, МХТС
Улаанбаатар, Монгол улс
nasanbat@must.edu.mn

Хураангуй: Хиймэл оюун ухаан (AI), ялангуяа их өгөгдөлтэй хэлний загварууд (LLM) нь орчин үеийн дижитал дэд бүтцийн үндсэн бүрэлдэхүүн болж, Generative AI-ийн хурдацтай нэвтрэлт нь кибер аюулгүй байдлын шинэ хэмжээст сорилтыг үүсгэж байна. Уламжлалт программ хангамжийн эмзэг байдлаас ялгаатай нь AI загваруудын сул тал нь prompt injection, jailbreak, data poisoning, мэдрэмтгий өгөгдөл гарган авах (model inversion) зэрэг өвөрмөц халдлагын векторуудад илэрдэг. Энэхүү судалгаагаар AI системд тусгайлан зориулсан "Улаан баг" (AI Red Teaming)-ийн арга зүй, эмзэг байдлын ангилал, илрүүлэлт болон хамгаалалтын боломжийг онолын болон туршилтын түвшинд гүйцэтгэсэн. Судалгаанд хосолсон арга зүй ашиглаж, олон төрлийн LLM загвар (Qwen2.5:7B, DeepSeek-V3) дээр 50 гаруй jailbreak техник, 15 төрлийн prompt injection туршилтыг бодит орчинд явуулсан. Туршилтын үр дүнгээс харахад одоогийн хамгаалалтын механизмууд (Guardrails, Moderation API) нь jailbreak оролдлогын 72%-ийг илрүүлдэг боловч үлдсэн 28% нь амжилттай халдлага болж, хориотой агуулга үүсгэх эрсдэлийг бий болгож байна.

Түлхүүр үг: A*I Red Teaming, Jailbreak, Prompt injection, Үндсэн загварын хамгаалалт*

Олон үүлэн орчны хандалтын бодлого дахь халдлагын замын графыг загварчлах нь

Баасанхүүгийн ЯНЖИНЛХАМ
Кибер аюулгүй байдлын тэнхимийн оюутан
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар, Монгол улс
yanjinlhamb0515@gmail.com

Эрдэнэжаргалын САНЧИРМАА
Кибер аюулгүй байдлын тэнхимийн оюутан
Кибер аюулгүй байдлын судалгааны баг
ШУТИС, МХТС
Улаанбаатар, Монгол улс
sancra2004@gmail.com

Баярмагнайн УНДРАЛ
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн багш
ШУТИС, МХТС
Улаанбаатар, Монгол улс
undral.b@must.edu.mn

Бат-Эрдэнийн МӨНХБАЯР
Кибер аюулгүй байдлын судалгааны баг
Кибер аюулгүй байдлын тэнхимийн эрхлэгч,
доктор (Ph.D.), дэд профессор
ШУТИС, МХТС
Улаанбаатар, Монгол улс
munkhbayar.b@must.edu.mn

Хураангуй—Орчин үеийн мэдээллийн технологийн салбарт байгууллагууд үйл ажиллагаагаа олон үүлэн орчинд зэрэг байршуулах болсон нь хандалтын хяналтын удирдлагыг улам бүр нарийн төвөгтэй болгож байна. Клауд орчны өргөжилт болон ижилслийн тоо нэмэгдэхийн хэрээр уламжлалт статик бодлогын шинжилгээний аргууд нь систем хоорондын логик холбоосуудаас үүдэлтэй далд эрсдэлийг илрүүлж чадахгүй байгааг салбарын судалгаанууд харуулж байна. Иймд энэхүү судалгааны ажил нь графийн онол болон эрсдэлийн математик тооцооллын модулийг ашиглан олон үүлэн орчны хандалтын бодлого дахь халдлагын замыг илрүүлэх шинэлэг системийг боловсруулсан болно. Бид судалгаандаа Amazon Web Services, Microsoft Azure болон Google Cloud Platform дээрх салангид ижил хэрэгслүүдийн харилцан хамаарлыг тогтоох каноникал схемын загварыг ашигласан.

Туршилтад JSON, YAML, CSV болон TSV форматаар ирсэн олон мянган хандалтын өгөгдөл болон бодлогын баримт бичгүүд дээр Neo4j граф өгөгдлийн санг ашиглан халдлагын замыг зураглан шинжилсэн. Системийн үндсэн логик нь давуу эрх шилжилт болон систем хоорондын итгэмжлэх эрхийн хамаарлыг гашиглан халдлагын боломжит векторуудыг өндөр нарийвчлалтайгаар тодорхойлж байгааг туршилтын үр дүн харуулж байна. Туршилтаар систем нь клауд дамжсан нууц түлхүүрийн ашиглалт болон аюултай хослолуудыг илрүүлж, тэдгээрт тохирох эрсдэлийн оноог математик загварчлалаар бодитой тооцож байгаа нь практик дээр батлагдсан.

Энэхүү судалгааны хүрээнд боловсруулсан Клауд Ижилслийн Граф Шинж программ нь аюулгүй байдлын мэргэжилтнүүдэд халдлагын гадаргууг бодитоор харах, эрсдэлийг бууруулах зөвлөмжийг шуурхай авах боломжийг олгож буйгаараа практик ач холбогдолтой юм.

Түлхүүр үг—Олон үүлэн орчин, хандалтын бодлого, илүүдэл эрх, эрсдэлийн үнэлгээ, халдлагын зам, IAM

Өсвөр насныхны сэтгэцийн эрүүл мэнд ба асран хамгаалагчийн хүн ам зүйн байдлын хоорондын хамаарлын загварчлал

Цагаан Баатархүү
Мэдээллийн технологи
ШУТИС, МХТС
Монгол, Улаанбаатар
baatarkhuu@must.edu.mn

Лувсан-Иш Ажнай
Физик, Мэдээлэл зүй
АШУҮИС, Био-Анагаах
Монгол, Улаанбаатар
ajnai@mnums.edu.mn

Цэнд Жавзмаа
Физик, Мэдээлэл зүй
АШУҮИС, Био-Анагаах
Монгол, Улаанбаатар
javzmaa.ts@mnums.edu.mn

Лувсанцэрэн Пүрэвдолгор
Физик, Мэдээлэл зүй
АШУҮИС, Био-Анагаах
Монгол, Улаанбаатар хот
purevdolgor@mnums.edu.mn

Чилхаасүрэн Баасандорж
Физик, Мэдээлэл зүй
АШУҮИС, Био-Анагаах
Монгол, Улаанбаатар хот
baasandorj@mnums.edu.mn

Ахыт Тилеубай
Физик, Мэдээлэл зүй
АШУҮИС, Био-Анагаах
Монгол, Улаанбаатар хот
Akhit@mnums.edu.mn

Болдбаатар Одгэрэл
Физик, Мэдээлэл зүй
АШУҮИС, Био-Анагаах
Монгол, Улаанбаатар
odgerel.b@mnums.edu.mn

Энхжаргал Энх-Үрэл
Физик, Мэдээлэл зүй
АШУҮИС, Био-Анагаах Монгол,
Улаанбаатар хот
enkh-urel@mnums.edu.mn

Оюунбилэг Бат-Энх
Мэдээллийн технологи МҮИС,
Мэдээлэл Технологи Дизайны Сургууль
Монгол улс, Улаанбаатар
batenkh@mnu.edu.mn

Хураангуй — Өсвөр насныхны хөгжлийн орчинд гэр бүл нь хамгийн суурь нөлөө үзүүлэгч хүчин зүйл бөгөөд гэр бүлийн бүтэц, харилцаа, тогтвортой байдал нь хүүхдийн сэтгэлзүйн төлөвшилд шууд нөлөөлдөг. Гэр бүлийн орчны ялгаа нь өсвөр үеийнхний зан төлөв, сэтгэл хөдлөлийн хариу үйлдэлд тусгалаа олж, стрессийг даван туулах чадвар болон нийгэмших үйл явцад чухал үүрэг гүйцэтгэдэг. Ялангуяа асран хамгаалагчийн нас, өрхийн ам бүлийн тоо, гэр бүлийн байдал, эдийн засгийн нөхцөл зэрэг үзүүлэлтүүд нь өсвөр насныхны сэтгэлзүйн тогтвортой байдалд нөлөөлөх гол хүчин зүйлсийн нэг гэж үзэгддэг. Эдгээр нөлөөллийг үнэлэхэд сэтгэл хөдлөл, зан үйлийн өөрчлөлтийг илрүүлэх стандарт хэрэгсэл болох Strengths and Difficulties Questionnaire (SDQ)-ийг өргөн ашигладаг бөгөөд уг арга нь өсвөр насныхны сэтгэлзүйн эрсдэлийг олон талт үзүүлэлтээр хэмжих боломжийг олгодог. Иймээс гэр бүлийн орчны үзүүлэлтүүдийг SDQ үнэлгээтэй уялдуулан судлах нь өсвөр үеийнхний сэтгэцийн эрүүл мэндийн асуудлыг эрт илрүүлэх, урьдчилан сэргийлэхэд чухал ач холбогдолтой юм. Энэхүү судалгаанд Говь-Алтай аймгийн 6–12 дугаар ангийн сурагчид, тэдний асран хамгаалагчид болон багш нарыг хамруулсан нийт 10,600 гаруй оролцогчоос SDQ асуумж ашиглан өгөгдөл цуглуулж, сурагч, асран хамгаалагч болон ангийн багшийн сэтгэл хөдлөл, зан үйлийн үнэлгээний өгөгдлийн санг бүрдүүлсэн. Асран хамгаалагчийн санд машин сургалтын аргуудыг гэр бүлийн үзүүлэлтүүдэд нь тулгуурлан өсвөр үеийнхний сэтгэцийн эрүүл мэндийн эрсдэлийг илрүүлэх загваруудыг боловсруулсан.

Түлхүүр үг— mental, SDQ, Random Forest