

Бүтээлийн Жагсаалтаас Хамтын Граф Байгуулах Зарим Алгоритм

Э.Гэрэлмаа, Ц.Ганбат

ШУТИС, Мэдээлэл, холбооны технологийн сургууль,

Компьютерийн ухааны салбар

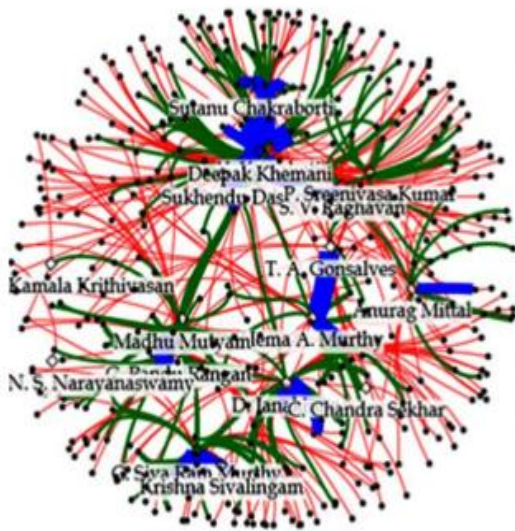
Мэдээллийн технологийн салбар

erdenechimeggerelmaa@gmail.com, ganbat_tsend@yahoo.com

Хураангуй— Эрдэмтэн, судлаачдын бүтээлийн жагсаалт нь зохиогчид, бүтээлээс бүрдсэн жагсаалт юм. Бүтээлийн жагсаалтаас зохиогчийн болон зохиогчдын хамтын сүлжээг графаар гаргах тухай энд өгүүлнэ. Хамтын сүлжээг гаргахын тулд бүтээлийн жагсаалтыг графаар илэрхийлэх, графыг компьютерт дүрслэх арга, зарим статистик гаргах алгоритм боловсруулах, түүний үндсэн дээр хамтын сүлжээг графаар илэрхийлэх, графыг дүрслэх, харуулах, графыг гаргах алгоритмуудыг боловсруулсан тухай өгүүлнэ.

I. УДИРТГАЛ

Графийн хэрэглээ нь харилцаа холбоо, биотехнологи, компьютерийн салбар, зам тээвэр, гэрлэн дохио, нийтийн тээвэр, цэрэг арми, нислэгийн хуваарь гэх мэт олон салбарт хэрэглэгдэж байна. Мөн том хэмжээний графууд олон улсад хэрэглэгддэг. Тухайлбал холливудийн жүжигчид хамтарч кинонд төглөсөн бол хоорондоо холбогдох hollywood граф. Энэхүү граф нь 2011 онд 1,5 сая оройтой байсан юм. Мөн Унгарын математикч Erdos гэдэг эрдэмтний бусад судлаачтай хамтарсан хамтралын графыг дурдаж болно. Энэ хүн амьдралдаа 500 хүнтэй хамтарч нийт 1525 өгүүлэл бичсэн байна. Мөн Энэтхэгийн технологийн институт (ИТ) салбар сургуулийн эрдэм шинжилгээний хамтын графийг байгуулжээ. [7]



Зураг 1. ИТ Madras Co-authorship Graph

Монголд эрдэмтдийн хамтын графийн байгуулах хэрэглээ гарч байна. ШУТС-д эрдэмтдийн хамтын графийг байгуулж байна. Хамтын графаас эрдэмтдийн хамтарсан байдал, тухайн эрдэмтний өгүүллийн тоо, тухайн өгүүллийн зохиогчийн тоо, эрдэм шинжилгээний салбаруудын интеграцийг харуулах зэрэг олон тооны үзүүлэлтийг гаргаж авч болох юм.

II. БҮТЭЭЛИЙН ЖАГСААЛТЫГ ДҮРСЛЭХ

A. Бүтээлийн жагсаалт

Эрдэмтэн, судлаачдын бүтээлийн жагсаалт гэдэг нь бүтээлийн зохиогчид, бүтээлээс бүрддэг.

#	Authors	Paper	Other info
1	Khuder A, Ganbat Ts	N-gram analysis of a Mongolian Text	2011, MUST
2	Dolgorsuren B, Ganbat Ts	Development of Computer law syllabus	2011, ШУТИС
3	Munkhbuyan B, Ganbat Ts, Batzolboo B	Enterprise Architecture Development strategy in Higher Education	2011, EICTHE
4	Zolboo D, Ganbat Ts	The survey on online service at metropolis State	2012, FITAT
5	Zolboo D, Ganbat Ts	Survey of comparison of the governmental online service	2012, MUST
6	Dolgorsuren B, Ganbat Ts	Some translation problems on software technology term	2012, MUST
7	Khuder A, Ganbat Ts	Part of Speech Tagging Experiments on Mongolian Language	2012, MUST
8	Dolgorsuren B, Ganbat Ts	Development of computer model to calculate statistics about traffic jam	2013, MMT

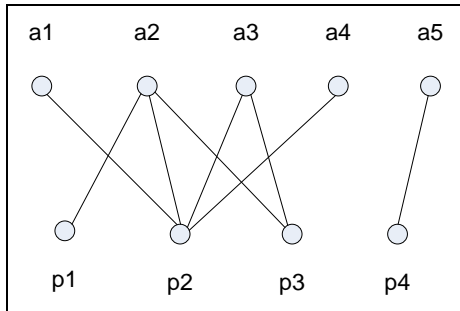
Зураг 2. Бүтээлийн жагсаалт

B. Бүтээлийн жагсаалтыг графаар дүрслэх

$A = \{a_1, a_2, a_3, \dots, a_N\}$ гэсэн n зохиогчтой A олонлог, $P = \{p_1, p_2, p_3, \dots, p_M\}$ гэсэн m өгүүлэлтэй P олонлогийн тусламжтайгаар бүтээлийн жагсаалтыг $PS = \{(a_1, p_2), (a_1, p_5), (a_2, p_2), \dots, (a_N, p_7)\}$ олонлогоор дүрслэж болно [1]. Түүнийг G_1 граф гэж үзвэл $A \cap P = \emptyset$ байх хоёр хэсэгт $A \cup P$ олонлог, PS олонлогоос бүрдэх $G_1(A \cup P, PS)$ чиглэлгүй граф болно

Тухайлбал $A = \{a_1, a_2, a_3, a_4, a_5\}$ зохиогчдын олонлог, $P = \{p_1, p_2, p_3, p_4\}$ өгүүллийн олонлогоор

бүтээлийн жагсаалтыг $PS = \{(a1,p2), (a2,p1), (a2,p2), (a2,p3), (a3,p2), (a3,p3), (a4,p2), (a5,p4)\}$ олонлогоор дүрсэлж болно. Графаар дүрсэлвэл:



Зураг 3. Бүтээлийн граф

n эрдэмтэн, m бүтээлтэй $G1$ графийг компьютерт дүрслэх матрицын хэд хэдэн арга байдаг. Тэдгээрийн нэг болох матрицын мөрөнд зохиогчдыг, баганад бүтээлийг тэмдэглэж мөр, баганы огтолцолд зохиолч бүтээлийг бичвэл 1, бичээгүй бол 0 тэмдэглэх аргыг хэрэглэе [3].

Бүтээлийн матриц нь n мөр, m баганатай.

$$C = \begin{bmatrix} c_{1,1} & c_{1,2} & c_{1,3} & \dots & c_{1,m} \\ c_{2,1} & c_{2,2} & c_{2,3} & \dots & c_{2,m} \\ c_{3,1} & c_{3,2} & c_{3,3} & \dots & c_{3,m} \\ \dots & \dots & \dots & \dots & \dots \\ c_{n,1} & c_{n,2} & c_{n,3} & \dots & c_{n,m} \end{bmatrix} \quad \text{Үүнд}$$

$$c_{i,j} = \begin{cases} 1 & (a_i, p_j) \in AP \\ 0 & \text{үгүй} \end{cases} \quad (1)$$

Зураг 3-т дүрслэсэн графийг $C = |c(i,j)|$ матрицаар $i=1..n, j=1..m$ дүрслэвэл

$$C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{Үүнд}$$

$$c(i,j) = \begin{cases} 1 & (a,p) \in PS \\ 0 & \text{үгүй бол} \end{cases} \quad (2)$$

C. Бүтээлийн жагсаалтыг графаар дүрслэх

Матрицын мөр бүрийн нийлбэр нь зохиогчийн өгүүллийн тоо болно. i -р зохиогчийн өгүүллийн тоог $paper_count(i)$ гэж тэмдэглэвэл түүнийг дараах алгоритмаар олно.

Хүснэгт 1 Зохиогчийн өгүүллийн тоог олох алгоритм, ГАРСАН ҮР ДҮН

for i=1 to n	a1	1
t=0	a2	3
for j=1 to m	a3	2
t=t+c(i,j)	a4	1
paper_count(i)=t	a5	1

Матрицын багана бүрийн нийлбэр нь өгүүллийн зохиогчдын тоо болно. j -р өгүүллийн зохиогчдын тоог $author_count(j)$ гэж тэмдэглэвэл түүнийг дараах алгоритмаар олно.

Хүснэгт 2 Өгүүллийн зохиогчдын тоог олох алгоритм, ГАРСАН ҮР ДҮН

for j=1 to m	p1	1
t=0	p2	4
for i=1 to n	p3	2
t=t+c(i,j)	p4	1
author_count(j)=t		

Эдгээрээс гадна матрицаас хамгийн олон өгүүлэлтэй зохиогч, хамгийн олон зохиогчтой өгүүллийг олох алгоритмууд байж болно.

III. ЗОХИОГЧДЫН ХАМТЫН ГРАФ

A. Бүтээлийн олонлогоос хамтын олонлог олох

Зохиогчдын өгүүллийн олонлог PS -ээс зохиогчдын хамтын CS олонлогийг хэрхэн олох вэ? Зохиогч ба түүний өгүүлэлтэй холбогдсон нумыг шууд холболт гэж нэрлэе. Харин зохиогчоос эхлэлтэй түүний өгүүллээр дайрч өөр зохиогчтой холбогдсон нумыг дам холболт гэе. Тэгвэл зохиогчоос эхтэй дам холболт бүрийг зам гэвэл

a1-ын хувьд

$$z1=a1p2a2 \quad z2=a1p2a3 \quad z3=a1p2a4$$

3 замтай

a2-ын хувьд

$$z4=a2p2a1 \quad z5=a2p2a3 \quad z6=a2p2a4 \quad z7=a2p3a3 \quad 4$$

замтай

a3-ын хувьд

$$z8=a3p2a1 \quad z9=a3p2a2 \quad z10=a3p2a4 \quad z11=a3p3a2 \quad 4$$

замтай

a4-ын хувьд

$$z12=a4p2a1 \quad z13=a4p2a2 \quad z14=a4p2a3 \quad 3$$

замтай

a5-ын хувьд зам байхгүй 0 замтай

Зохиогч бүрийн замын тоо хамтралын тоо, замын төгсгөл нь хамтарсан зохиогч болно.

Зам бүрээс дундын оройг хасвал хамтын олонлогийн элементүүд болно

$$CS = \{(a1,a2), (a1,a3), (a1,a4), (a2,a1), (a2,a3), (a2,a4), (a3,a1), (a3,a2), (a3,a4), (a4,a1), (a4,a2), (a4,a3)\}$$

Үүнийг зохиогчдын хамтын олонлог гэе. Харин a_i зохиолчийн хамтын олонлог нь a_i -ээр эхэлсэн элементүүдтэй дэд олонлог болно.

B. Хамтын матриц байгуулах алгоритм

Зохиогчдын хамтын матриц гэдэг нь n зохиолч, n зохиолчтой хамтарч бүтээл тууригдсан эсэхийг илэрхийлнэ. Ерөнхий бүтцийг харуулбал

$$CM = \begin{bmatrix} cm_{1,1} & cm_{1,2} & cm_{1,3} & \dots & cm_{1,n} \\ cm_{2,1} & cm_{2,2} & cm_{2,3} & \dots & cm_{2,n} \\ cm_{3,1} & cm_{3,2} & cm_{3,3} & \dots & cm_{3,n} \\ \dots & \dots & \dots & \dots & \dots \\ cm_{n,1} & cm_{n,2} & cm_{n,3} & \dots & cm_{n,n} \end{bmatrix}$$

Үүнд:

$$cm_{i,j} = \begin{cases} 1 & (a_i, a_j) \in CS \\ 0 & \text{үгүй} \end{cases} \quad (3)$$

Бүтээлийн матриц C-ээс CM-ыг гаргахын тулд эхлээд бүх $cm(i,j)=0$ болгоно. Үүнд $i=1..n, j=1..n$

Дараа нь дараах алгоритмаар CM матрицын элемент $cm(i,j)$ -ыг олно

Хүснэгт 3 Хамтын матриц байгуулах алгоритм

```

for i=1 to n
  for j=1 to n
    cm(i,j)=0
  for i=1 to n
    for j=1 to m
      if c(i,j)=1 then
        for k=1 to n
          if i<>k and c(k,j)=1 then cm(i,k)= cm(i,k)+1
    
```

Үүний үр дүнд $CM=|cm(i,j)|$ хамтын матриц үүснэ. Үүнд $i=1..n, j=1..n$

$$CM = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 2 & 1 & 0 \\ 1 & 2 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (4)$$

Мөр бүрийн 0 биш нүдний тоо нь зохиогч бүрийн хамтарсан зохиогчдын тоо болно.

Хүснэгт 4 Хамтарсан зохиогчдын тоог олох алгоритм, гарсан үр дүн

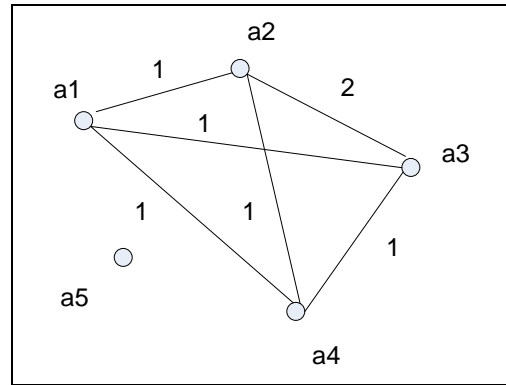
For i=1 to n	A1	3
k=0	A2	3
For j=1 to n	A3	3
If $cm(i,j)>0$ then $k=k+1$	A4	3
collab_authors(i)=k	A5	0

Харин мөрийн нийлбэр нийт хамтарсан тоо болно.

Хүснэгт 5 Хамтарсан нийт тоог олох алгоритм, гарсан үр дүн

For i=1 to n	A1	3
s=0	A2	4
For j=1 to n	A3	4
If $cm(i,j)>0$ then $s=s+cm(i,j)$	A4	3
collab_total(i)=s	A5	0

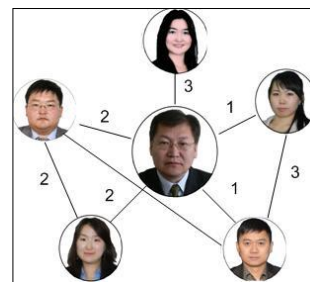
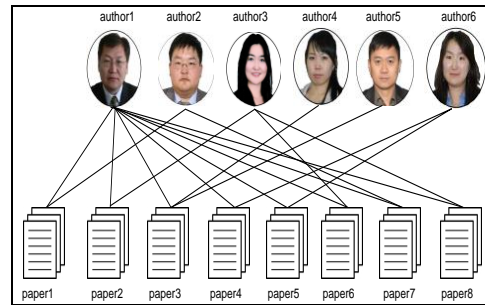
Хамтын CM матрицаас хамтын CG граф байгуулбал:



Зураг 4. Зохиогчдын хамтын граф

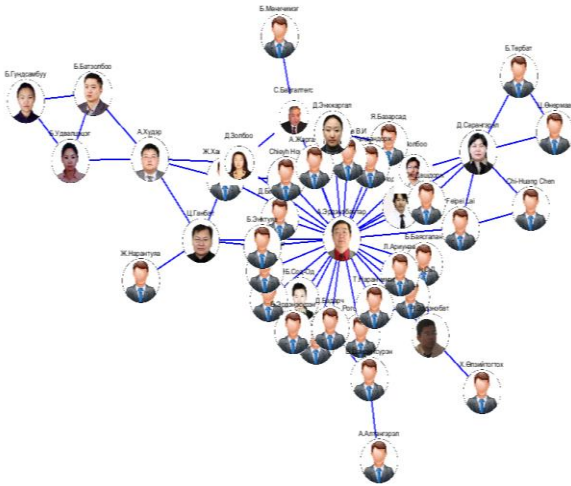
Энд a5 нь хэнтэй ч хамтраагүй тул тусгаар орой болно.

Бүтээлийн жагсаалтаас зохиогчдын хамтын граф гаргахаас гадна өгөгдсөн зохиолчийн хамтын графыг гаргах боломжтой.



Зураг 5. Зохиогчийн хамтын граф

Дээрх аргачлал ашиглан хамтын графыг харуулах програм хөгжүүлсэн. Програмын үр дүнг доор харуулав.



Зураг 6. GraphStream ашиглан байгуулсан хамтын граф

ДҮГНЭЛТ

Бүтээлийн жагсаалтыг компьютерт илэрхийлэх, харуулах, бүтээлийн жагсаалтаас зарим статистик олох, бүтээлийн жагсаалтаас хамтын граф байгуулах, дүрслэх алгоритмуудыг боловсруулаа.

Үүнийг өөрийн салбарын зарим судлаачдын бүтээлийн жагсаалт дээр туршиж үзлээ.

Цаашид бүтээлийн жагсаалтанд зохиогч болон өгүүлэл нэмэх, бүтээлийн жагсаалт шинэчлэгдэхэд хамтын графийг шинэчлэх ажлуудыг хийнэ.

АШИГЛАСАН МАТЕРИАЛ

- [1] Kenneth H. Rosen. Discrete mathematics and its applications. 7th edition. McGraw-Hill. 2012.
- [2] Deo Narsingh. System simulation with digital computer, ISBN 978-81-203-0028-6
- [3] Deo Narsingh. Graph theory with applications to engineering and computer science. ISBN: 978-81-203-0145-0
- [4] Ganbat Tsend. System simulation using computer. MUST edition. Ulaanbaatar. 2004. 2nd edition. In mongolian
- [5] Ganbat Tsend, Baatarkhuu Enkhtuya. Devepment of portal for Research publication. Project report. 2015. Ulaanbaatar. MEDS. Mongolian Science Technology Foundation
- [6] Ganbat Tsend. Developing a collaboration model between researchers using their publication bipartite graphs, the 2016 Summer School on ICT proceedings
- [7] Tasleem Arif, Rahid Ali, M.Asger. Scientific Co-authorship Social Networks: A Case Study of Computer Science Scenario in India. August 2012.