

МОНГОЛ УЛСЫН ИХ СУРГУУЛЬ
МЭДЭЭЛЛИЙН ТЕХНОЛОГИ, ЭЛЕКТРОНИКИЙН СУРГУУЛЬ
МЭДЭЭЛЭЛ, КОМПЬЮТЕРЫН УХААНЫ ТЭНХИМ

Мөнхжаргалын Билгүүндалай

ТӨРИЙН ЦАХИМ ҮЙЛЧИЛГЭЭ САНАЛ
БОЛГОХ СИСТЕМИЙН ЗАГВАР
ХӨГЖҮҮЛЭЛТ

(Model development of recommendation system for digital
public services)

Мэдээллийн технологи (D061304)
Бакалаврын судалгааны ажил

Улаанбаатар

2025 оны 04 сар

МОНГОЛ УЛСЫН ИХ СУРГУУЛЬ
МЭДЭЭЛЛИЙН ТЕХНОЛОГИ, ЭЛЕКТРОНИКИЙН СУРГУУЛЬ
МЭДЭЭЛЭЛ, КОМПЬЮТЕРЫН УХААНЫ ТЭНХИМ

ТӨРИЙН ЦАХИМ ҮЙЛЧИЛГЭЭ САНАЛ БОЛГОХ
СИСТЕМИЙН ЗАГВАР ХӨГЖҮҮЛЭЛТ
(Model development of recommendation system for digital
public services)

Мэдээллийн технологи (D061304)
Бакалаврын судалгааны ажил

Удирдагч: _____ Др. Г.Амарсанаа

Гүйцэтгэсэн: _____ М.Билгүүндалай (21B1NUM1559)

Улаанбаатар

2025 оны 04 сар

Зохиогчийн баталгаа

Миний бие Мөнхжаргалын Билгүүндалай ”ТӨРИЙН ЦАХИМ ҮЙЛЧИЛГЭЭ САНАЛ БОЛГОХ СИСТЕМИЙН ЗАГВАР ХӨГЖҮҮЛЭЛТ” сэдэвтэй судалгааны ажлыг гүйцэтгэсэн болохыг зарлаж дараах зүйлсийг баталж байна:

- Ажил нь бүхэлдээ эсвэл ихэнхдээ Монгол Улсын Их Сургуулийн зэрэг горилохоор дэвшүүлсэн болно.
- Энэ ажлын аль нэг хэсгийг эсвэл бүхлээр нь ямар нэг их, дээд сургуулийн зэрэг горилохоор оруулж байгаагүй.
- Бусдын хийсэн ажлаас хуулбарлаагүй, ашигласан бол ишлэл, зүүлт хийсэн.
- Ажлыг би өөрөө (хамтарч) хийсэн ба миний хийсэн ажил, үзүүлсэн дэмжлэгийг дипломын ажилд тодорхой тусгасан.
- Ажилд тусалсан бүх эх сурвалжид талархаж байна.

Гарын үсэг: _____

Огноо: _____

ГАРЧИГ

УДИРТГАЛ.....	1
1. МЭДЛЭГИЙН ГРАФД ТҮШИГЛЭСЭН САНАЛ БОЛГОХ СИСТЕМ ...	2
1.1 Мэдлэгийн граф	2
1.2 Санал болгох системийн аргууд	4
1.3 Мэдлэгийн графт анхааралтай сүлжээ.....	8
2. ЦАХИМ ҮЙЛЧИЛГЭЭГ САНАЛ БОЛГОХ ЗАГВАР	14
2.1 E-Mongolia портал	14
2.2 Цахим үйлчилгээний өгөгдөл.....	15
2.3 Мэдлэгийн графын загвар	16
3. ТУРШИЛТ.....	18
3.1 Үнэлэх аргачлал	18
3.2 Өгөгдлийг загварын сургалт болон туршилтад хуваах	19
3.3 Үр дүн.....	20
4. ДҮГНЭЛТ	24
5. КОДЫН ХЭРЭГЖҮҮЛЭЛТ	26

ЗУРГИЙН ЖАГСААЛТ

1.1	Мэдлэгийн граф	3
1.2	DKN ерөнхий бүтэц	6
1.3	MKR ерөнхий бүтэц	8
1.4	Хамтын мэдлэгийн граф	9
1.5	KGAT ерөнхий бүтэц	11
2.1	Мэдлэгийн графын загвар	15
3.1	Ирмэгийн төрөл нэмэгдэх үеийн өөрчлөлт	21

ХҮСНЭГТИЙН ЖАГСААЛТ

2.1	E-Mongolia порталын өгөгдлийн бүрдэл	15
2.2	Туршилтуудад ашиглагдсан мэдлэгийн графын хэмжээ	17
3.1	K=20 байх туршилтын үр дүн	20
3.2	Өгөгдлүүдийн үр дүнгийн харьцуулалт	22

Кодын жагсаалт

5.1	Хэрэглэгчийн өмнө авсан үйлчилгээний түүхээс сургалтын болон туршилтын хоёр хэсэгт граф үүсгэх код	26
5.2	Үйлчилгээ үзүүлдэг байгууллагын мэдээллийг агуулсан мэдлэгийн граф үүсгэх код	28
5.3	Үйлчилгээ үзүүлдэг байгууллага + үйлчилгээний хамаарах төрөл мэдээллийг агуулсан мэдлэгийн граф үүсгэх код	29
5.4	Хэрэглэгчийн хамаарах насны бүлэг мэдээллийг агуулсан мэдлэгийн граф үүсгэх код	31
5.5	Хэрэглэгчийн хамаарах насны бүлэг + үйлчилгээ үзүүлдэг байгууллага + үйлчилгээний хамаарах төрөл + үйлчилгээг хамгийн их авдаг сар мэдээллийг агуулсан мэдлэгийн граф үүсгэх код	33

УДИРТГАЛ

Граф ашигласан санал болгох системийн судалгаа эрчимжиж бараа бүтээгдэхүүн, сурталчилгаа, эсвэл кино, хөгжим зэргийг санал болгоход ихэвчлэн нэвтрүүлж байна. Харин төрийн цахим үйлчилгээний тухайд судалгаа цөөн. Энэ ажлаар Монгол Улсын төрийн цахим үйлчилгээний нэгдсэн портал E-Mongolia системийн хэрэглэгчийн үйлчилгээ авсан түүхэн өгөгдлөөс мэдлэгийн графт анхааралтай сүлжээ аргачлалаар цахим үйлчилгээг санал болгох граф машин сургалтын загварыг туршиж үнэлсэн. Өгөгдлийн бүрдлээс хоорондоо ялгаатай мэдлэгийн графуудыг үүсгэн санал болгох загварын гүйцэтгэлд үзүүлэх нөлөөг харьцуулсан. Туршилтын үр дүнд мэдлэгийн графыг үйлчилгээнд чухал нөлөөлөх агуулгаар баяжуулахад таамаглах хувийг нэмэгдүүлж байгааг олж мэдлээ. Нийтдээ 50,000 хэрэглэгчийн авсан үйлчилгээний өгөгдлийг ашиглан сургасан загварын $recall@20$, $ndcg@20$ хэмжигдэхүүнүүдийн хувьд тус тус 0.2597, 0.3097 утгатай гарсан болно.

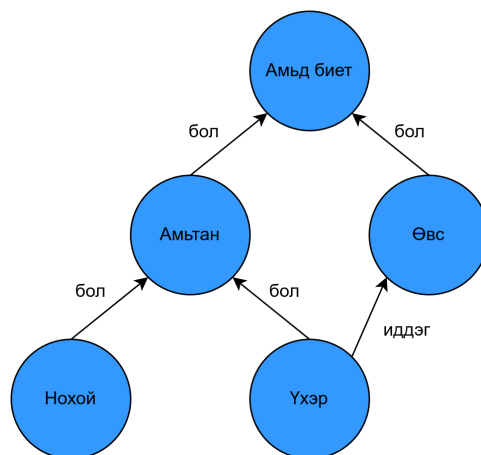
1. МЭДЛЭГИЙН ГРАФД ТҮШИГЛЭСЭН САНАЛ БОЛГОХ СИСТЕМ

Интернэтийн хөгжил, өргөжилттэй зэрэгцэн мэдээллийн хэт их өсөлтөөс шалтгаалан хэрэглэгчид маш олон сонголт үүсэж, тэдгээр сонголтуудаас өөрийн сонирхож буй зүйлийг олох нь улам бүр төвөгтэй болсоор байна. Энэ асуудлыг шийдвэрлэж, хэрэглэгчийн туршлагыг сайжруулахын тулд дуу, кино, худалдан авалтын бүтээгдэхүүн санал болгодог санал болгох системүүд өргөнөөр ашиглагдах болсон. Санал болгох систем нь ямар ч төрлийн өгөгдлийн санд түшиглэж ажиллах боломжтой. Харин мэдлэгийн графт түшиглэсэн санал болгох системийн давуу тал нь хэрэглэгчид санал болгосон дуу, кино, эсвэл бүтээгдэхүүнийг ямар шалтгаанаар санал болгосон болохыг тайлбарлах боломжтой байдаг явдал юм [1, 2, 3, 4, 5].

1.1 Мэдлэгийн граф

Мэдлэгийн граф нь бодит амьдрал дээрх юмс, нэгж-объектыг, хооронд нь холбох холбоосоор гурвал үүсгэж, жишээ нь, *[толгой объект]—[холбоо хамаарал]→[сүүл объект]* хэлбэрээр илэрхийлсэн чиглэлтэй граф бүтэцтэй өгөгдлийн бүрдэл юм [5]. Жишээлбэл, *[Иргэн]—[авсан үйлчилгээ]→[Ориин суугаа газрын тодорхойлолт]* гэсэн гурвал нь тухайн иргэн, иргэний авсан үйлчилгээ, мөн энэ хоёр объектын хоорондын холбоосыг агуулж буй. Мэдлэгийн граф дахь орой нь объектыг, ирмэг нь объектуудын хоорондын холбоо хамаарлыг илэрхийлдэг. Мэдлэгийн граф нь нэгээс олон төрлийн зангилаа болон ирмэгүүдийг агуулдаг тул олон талт граф юм. Ийм төрлийн граф нь объектын олон төрлийн шинж чанаруудыг граф дахь ялгаатай ирмэгүүдийг даган олж авах боломжийг олгодог тул илэрхийлэх чадамж өндөртэй байдаг. Мөн ирмэгүүдийн хоорондын харилцааг ашиглан объектуудын өндөр түвшний холбоо хамаарлыг илүү нарийн ойлгох боломжтой. Зураг [1.1] -с мэдлэгийн графын

ерөнхий бүтэц болон "нохой бол амьтан", "үхэр өвс иддэг", "өвс бол амьд биет" зэрэг мэдээллүүдийг харж болно.



Зураг 1.1: Мэдлэгийн граф

Мэдээллийг хүснэгтэд хадгалдаг холбоо хамааралт өгөгдлийн санд илэрхийлэх, олж харах боломжгүй холбоо хамаарал болон утга зүйг мэдлэгийн графд илэрхийлэх боломжтой. Жишээлбэл, "шар" гэдэг үгийг авч үзье. Мэдлэгийн граф ашиглах үед энэ үг нь "шар" өнгө болон "шар" үхэр гэсэн хоёр өөр утгыг илэрхийлэх тусдаа зангилаанууд байдлаар хадгалагдана. Өгөгдлийг графаар дүрслэх нь "шар" гэдэг үгийн олон утгыг илүү хялбар бөгөөд тодорхой байдлаар ялгаж ойлгох боломжийг олгодог. Харин холбоо хамааралт өгөгдлийн сан ашиглан өгөгдлийг илэрхийлэх тохиолдолд "шар" гэдэг үгийн утгыг ялгахын тулд нэг нь өнгөтэй холбоотой, нөгөө нь амьтантай холбоотой хоёр өөр хүснэгтэд хадгалж болно. Гэвч энэ тохиолдолд үгийн утгыг ялгаж ойлгоход төвөгтэй болно. Санал болгох системээс гадна, мэдлэгийн графт түшиглэсэн хиймэл оюун ухаан, машин сургалтын загварын том давуу талуудын нэг нь тухайн загварын гаргасан шийдвэрийг яагаад, ямар шалтгаанаар гаргасан болохыг тайлбарлах боломжтой [1, 2]. Үүний үр дүнд хэрэглэгчид болон судлаачид хиймэл оюун ухааны шийдвэр гаргалтын үйл явцыг гүнзгий судалж,

алгоритмын ажиллагааг илүү сайн ойлгох боломжтой болно.

1.2 Санал болгох системийн аргууд

Төрийн цахим үйлчилгээг санал болгох загварын гол зорилго нь хэрэглэгчид ирээдүйд шаардлагатай болох магадлалтай үйлчилгээг урьдчилан таамаглаж санал болгох явдал юм. Мэдлэгийн графын үүднээс авч үзвэл энэхүү асуудал нь хэрэглэгч болон төрийн үйлчилгээний хооронд үүсэж болох шинэ ирмэг буюу холбоо хамаарлыг илрүүлэх асуудал болж хувирна [1, 2, 3]. Мэдлэгийн графт түшиглэсэн санал болгох системийг хэрэгжүүлэх үндсэн 3 арга байдаг [1, 2].

- **Хамтын шүүлтүүр**

Энэ арга нь хэрэглэгчийн өмнө нь сонирхож байсан зүйлсийн түүхийг ашиглан таамаглал гаргадаг бөгөөд хэрэглэгч нь ижил төстэй сонирхолтой бусад хэрэглэгчдийн сонгосон зүйлсийг сонирхох магадлалтай гэж үздэг.

- **Агуулгад суурилсан шүүлтүүр**

Энэ арга нь мэдлэгийн граф дахь хэрэглэгч, зүйлсийн агуулга дээр үндэслэн таамаглал гаргадаг бөгөөд хэрэглэгч нь өмнө нь сонирхож байсан зүйлстэйгээ ижил төстэй зүйлсийг сонирхох магадлалтай гэж үздэг.

- **Хосолсон**

Энэ төрлийн арга нь хамтын шүүлтүүр болон агуулгад суурилсан шүүлтүүр аргуудыг хослуулан ашиглаж, тус тусын хязгаарлалтыг нөхөн сайжруулснаар илүү оновчтой санал болголт хийх боломжийг олгодог.

Мэдлэгийн графт түшиглэсэн, дахин ашиглах боломжтой, хамгийн сүүлийн үеийн санал болгох системийн загвар сургах аргуудыг судалж, E-Mongolia порталын өгөгдөлд тохирохыг нь сонгон авч туршилт хийсэн. Арга, алгоритм сонгох явцад судалсан аргуудыг дараах дэд хэсгүүдэд дэлгэрэнгүй тайлбарлана.

1.2.1 Мэдлэгийн графт түшиглэсэн гүн сүлжээ

Уламжлалт мэдээ санал болгох аргууд нь мэдээний гарчиг дахь объектууд болон агуулагдаж буй утга зүйг хангалттай хэмжээнд илрүүлж чаддаггүйгээс болж, гарчгуудын хоорондын холбоог зөв тодорхойлох чадвар нь хязгаарлагддаг. Мэдлэгийн графт түшиглэсэн гүн сүлжээ (DKN) загвар нь мэдлэгийн граф ашиглан мэдээ санал болгох системд тулгардаг асуудлыг шийдэж, үр дүнг сайжруулах зорилгоор боловсруулсан, агуулгад суурилсан гүн сургалтын арга юм. DKN нь мэдээний гарчиг дахь үгийн утга зүйн талаарх ойлголтыг сайжруулахын тулд гаднын буюу нэмэлт мэдлэгийн граф ашигладаг. Энэ нь санал болгох системийн загварт мэдээний утга зүйг илүү сайн тусгаж, мэдээ санал болголтын нарийвчлалыг нэмэгдүүлдэг [4].

DKN-ын гол бүрэлдэхүүн хэсгүүд:

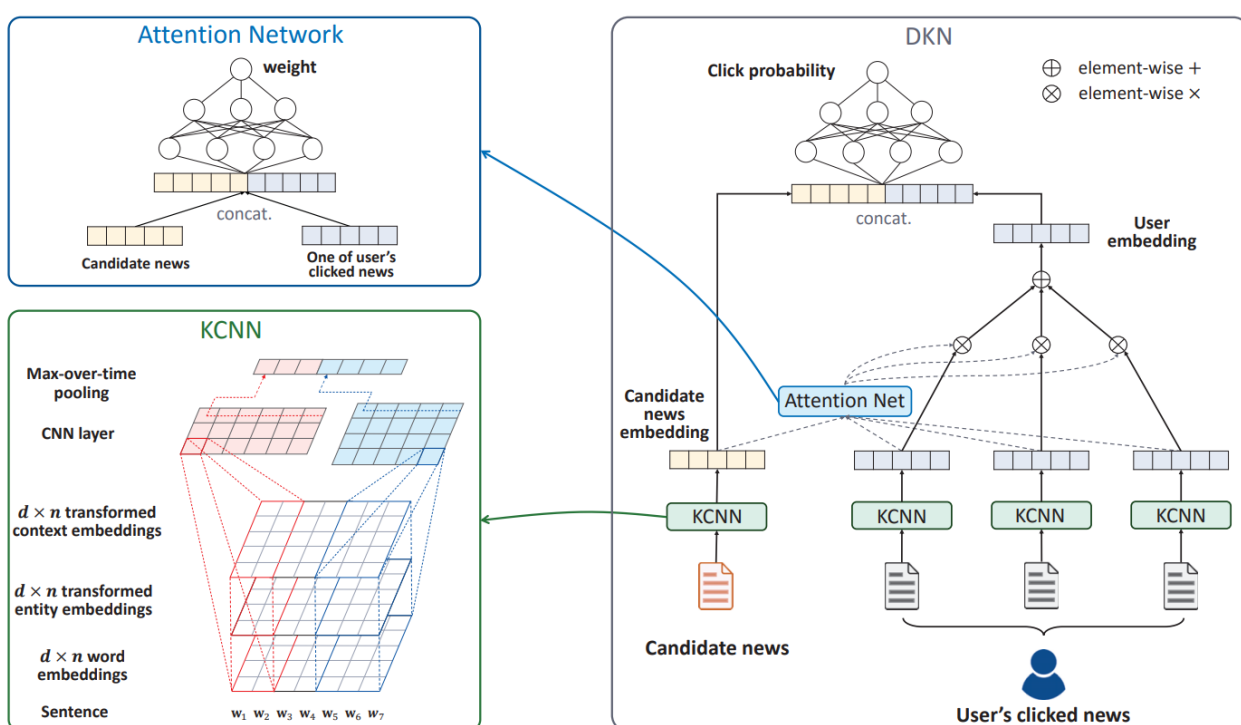
- **Мэдлэгт суурилсан хуримтлагдсан нейрон сүлжээ (KCNN):**

Хүн мэдээний гарчиг уншиж ойлгодог шиг гарчиг дахь үгсийн утгыг ойлгохын тулд шинжилж, утга зүйн түвшний илэрхийлэмжийг тодорхойлдог. Мэдээний гарчиг дахь үгийг дэлгэрэнгүй тайлбарлахын тулд нэмэлт мэдлэгийн граф буюу гаднын эх сурвалжаас мэдээлэл авч, мэдлэгийн түвшний илэрхийлэмжийг тодорхойлдог. KCNN ажиллахдаа ялгаатай түвшний илэрхийлэмжүүдийг ялгаатай сувгаар зэрэг тооцоолол хийж, тодорхойлох боломжтой бөгөөд энэ зарчим нь тоон зураг улаан, ногоон, цэнхэр гурван сувгаар дамжин боловсруулагддаг үйл явцтай төстэй гэж үзэж болно. KCNN-д тодорхойлогдсон утга зүйн түвшний болон мэдлэгийн түвшний хоёр өөр төрлийн мэдээлэл буюу илэрхийлэмжүүд нэгтгэгдэн санал болгох системд мэдээний гарчгийг илүү сайн ойлгоход тусалдаг.

- **Анхаарлын модуль:**

Анхаарлын модуль нь санал болгох системд хэрэглэгч бүрийн онцлог, ялгаатай байдлыг харгалзан, тухайн хэрэглэгчид хамгийн ач холбогдолтой мэдээг санал

болгох боломжийг бүрдүүлдэг. Хэрэглэгч бүр өөр өөр сонирхолтой төдийгүй, тэдний сонирхдог мэдээний төрөл нь хугацааны явцад өөрчлөгдөх боломжтой байдаг. Анхаарлын модуль нь хэрэглэгчийн өмнө нь уншиж, сонирхсон мэдээнүүдийг анализ хийж, тэдгээртэй утга зүйн болон сэдвийн хувьд ижил төстэй байж болох мэдээг санал болгох зарчмаар ажилладаг. Энэ механизм нь системд хэрэглэгчдийн хугацааны явцад өөрчлөгдөж буй сонирхолд дасан зохицож хэрэглэгчийн одоогийн сонирхолд нийцсэн мэдээг санал болгох боломжийг олгодог.



Зураг 1.2: DKN ерөнхий бүтэц

1.2.2 Олон үүрэгт шинж чанарын сургалт

Хамтын шүүлтүүр арга нь хэрэглэгч өөртэйгөө төстэй бусад хэрэглэгчдийн өмнө нь сонирхсон зүйлсийг сонирхох магадлалтай гэж үздэг. Гэвч энэхүү арга нь шинэ хэрэглэгчид үйлчилгээ санал болгох буюу өгөгдөл хангалтгүй үед үр дүн багатай байх асуудалтай тулгардаг [1, 2, 5]. Энэхүү асуудлыг шийдвэрлэхийн тулд олон үүрэгт шинж чанарыг сургалт

арга боловсруулагдсан бөгөөд хэрэглэгч болон зүйлсийн харилцаа болон мэдлэгийн граф хоёроос нь сурах замаар дурдсан асуудлыг багасгах зорилготой [5]. MKR-ын бүрэлдэхүүн хэсгүүд:

- **Санал болгох модуль:**

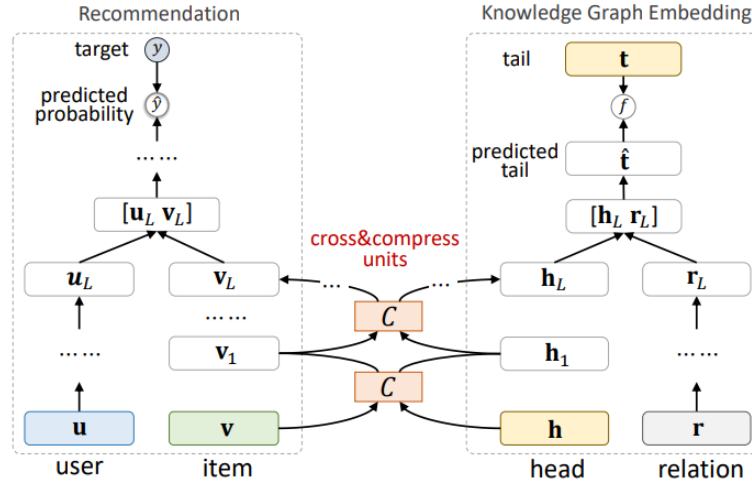
Энэ модуль нь оролтод хэрэглэгч болон зүйлийн өгөгдлийг авч, multi-layer perceptron (MLP) ашиглан тэдгээрийн онцлог шинж чанаруудыг ялгаж, боловсруулдаг. Дараа нь эдгээр шинж чанарууд дээр тооцоолол хийж, тухайн хэрэглэгч тухайн зүйлийг сонирхох магадлалыг таамаглах үүрэгтэй.

- **Мэдлэгийн графын ембеддинг модуль:**

Энэ модуль нь мэдлэгийн граф дахь объектуудын хоорондын бүтэцлэгдсэн холбоо хамаарлыг сурч, мэдлэгийн граф дахь гурвалуудаас хэрэгтэй шинж чанаруудыг ялган авах үүрэгтэй. Ингэснээр мэдлэгийн граф дотор шууд илэрхийлэгдээгүй харилцаануудыг таамаглах боломжийг бий болгодог. Өөрөөр хэлбэл, байхгүй холбоо хамаарлыг таамаглаж, мэдлэгийн графт байхгүй "сүүл" объектыг тодорхойлох боломжтой.

- **Cross&Compress Unit:**

Энэ нь MKR аргын гол бүрэлдэхүүн хэсэг бөгөөд санал болгох модуль болон мэдлэгийн графын ембеддинг модулиудыг холбож, харилцан мэдээлэл дамжуулах боломжийг бүрдүүлдэг. Үүний үр дүнд нэг модуль нөгөөгийнхөө сургалтад туслан, илүү нарийвчлалтай таамаглал гаргах боломжтой болдог.



Зураг 1.3: MKR ерөнхий бүтэц

1.3 Мэдлэгийн графт анхааралтай сүлжээ

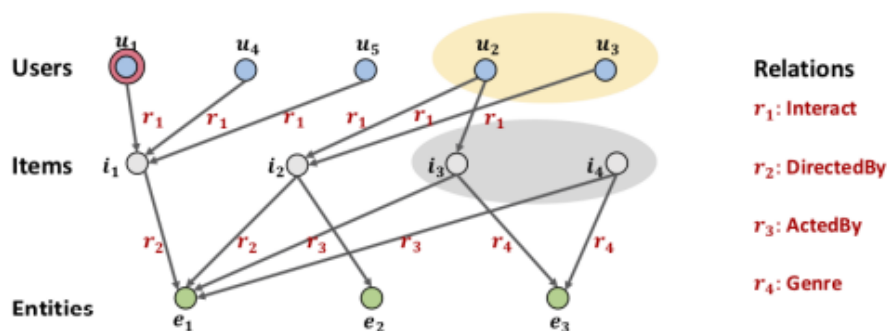
Мэдлэгийн графт суурилсан санал болгох системийн загвар хөгжүүлэх Мэдлэгийн графт анхааралтай сүлжээ (KGAT) арга нь MKR аргатай ижлээр хамтын шүүлтүүр аргын сул талыг шийдвэрлэх зорилгоор загварчлагдсан. Хамтын шүүлтүүрийн арга нь хэрэглэгч болон үйлчилгээ хоорондын харилцаанд тулгуурлан оновчтой таамаглал гаргах чадвартай боловч үйлчилгээ болон хэрэглэгчийн шинж чанар, холбоо хамаарлыг илэрхийлэх нэмэлт өгөгдлийг ашиглах боломж багатай байдаг [1, 2, 3]. Үүнээс шалтгаалан хэрэглэгчийн өмнө нь авсан үйлчилгээний түүх хомс үед энэ төрлийн аргуудын гүйцэтгэл буурах сул талтай байдаг. KGAT арга нь зөвхөн хоёр талт граф бус хамтын мэдлэгийн граф ашиглаж хэрэглэгч болон үйлчилгээ хоорондын өндөр зэрэглэлийн холболтуудыг илрүүлж загварын санал болголтын чанарыг сайжруулах замаар дээрх сул талыг шийдэхээр загварчлагдсан арга юм. Энэхүү аргачлал нь мэдлэгийн граф дахь өндөр зэрэглэлийн холболт буюу олон орой дамжсан уялдааг тооцоолон илрүүлдэг [3].

1.3.1 Томьёолол

Энэ хэсэгт Мэдлэгийн графт анхааралтай сүлжээ аргын хүрээнд ашиглагдах мэдлэгийн графтай холбоотой ойлголтуудыг дэлгэрэнгүй тайлбарлана.

- **Хоёр талт граф:**

Санал болгох системийн үндсэн суурь бүтцүүдийн нэг бол хоёр талт граф (bipartite graph) юм [3, 5]. Энэ нь хоёр төрлийн оройтой граф бөгөөд ихэвчлэн хэрэглэгч ба үйлчилгээг илэрхийлдэг. Хэрэглэгчид зөвхөн үйлчилгээтэй, үйлчилгээ зөвхөн хэрэглэгчтэй холбогдсон байдлаар хоёр оройн хооронд ирмэг үүсдэг. Энэ граф нь хэрэглэгчийн өмнө авсан үйлчилгээний түүхийг илэрхийлэхэд тохиромжтой бөгөөд мэдлэгийн графт түшиглэсэн санал болгох загваруудад өргөн хэрэглэгддэг. Гэсэн хэдий ч энэхүү граф нь хэрэглэгч болон үйлчилгээний нэмэлт шинж чанарууд, холбоо хамаарал зэргийг илэрхийлэх боломжгүй учир өндөр зэрэглэлийн холболт болон утга агуулгатай тайлбар шаардсан загваруудад хангалтгүй байдаг [3].



Зураг 1.4: Хамтын мэдлэгийн граф

- **Хамтын мэдлэгийн граф:**

Хамтын мэдлэгийн граф (collaborative knowledge graph) нь хэрэглэгчийн өмнө авсан үйлчилгээний түүхийг илэрхийлэх хоёр талт графыг түүний объектуудын нэмэлт шинж чанар, жишээлбэл, үйлчилгээний ангилал, үйлчилгээг үзүүлдэг байгууллага гэх

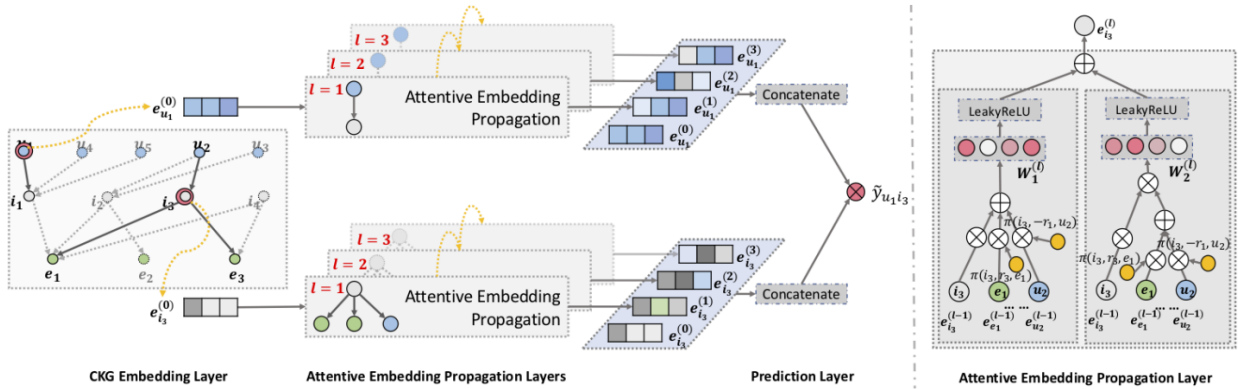
зэргийг агуулах мэдлэгийн графтай нэгтгэн өргөтгөсөн нэг төрлийн граф бүтэц юм. Санал болгох системийн загварт хамтын мэдлэгийн графыг ашигласнаар хэрэглэгчийн нэг үйлчилгээг авсан мэдээллээр дамжуулан түүнтэй холбогдсон бусад үйлчилгээг санал болгох өндөр боломжтой болдог [3]. Жишээлбэл: (хэрэглэгч1 - үйлчилгээ1 - ангилалА - үйлчилгээ2) холбоо нь хэрэглэгч1 хэрэглэгчид үйлчилгээ2-г санал болгоход ашиглагдаж болно.

- **Өндөр түвшний холболт:**

Өндөр түвшний холболт нь хэрэглэгч болон зүйлийн хоорондох илүү гүн, олон алхам бүхий холбоо хамаарлыг илрүүлэх замаар санал болгох системийн үр дүнг нэмэгдүүлдэг. Зөвхөн шууд холбогдсон харилцан үйлдлийг шалгахын оронд, энэ нь олон орой дамжсан холболтыг дагаж, хэрэглэгчдийг ижил төстэй сонирхол эсвэл зүйлийн шинж чанараар холбодог. Жишээлбэл, Хэрвээ хэрэглэгч нэг кино үзсэн, тэр кино нь өөр нэг кинотой ижил найруулагчтай, тухайн хэрэглэгч өмнө нь үзээгүй кино байж таарвал систем хэрэглэгчид хоёр дахь киног санал болгож болно. Нуугдмал холбоо хамаарлуудыг илрүүлдэг учраас санал болгох нарийвчлалыг нэмэгдүүлдэг.

1.3.2 Аргачлал

Энэ хэсэгт Мэдлэгийн графт анхааралтай сүлжээ (KGAT) аргын үндсэн бүтэц, бүрэлдэхүүн хэсгүүдийн талаар тайлбарлана.



Зураг 1.5: KGAT ерөнхий бүтэц

- **Эмбеддинг давхарга:**

Энэ давхарга нь хэрэглэгч, үйлчилгээ болон бусад оройнуудыг тоон вектор болгон хөрвүүлж граф дахь бүтцийг машин сургалтад тохиромжтой хэлбэрт оруулдаг. KGAT арга нь хамтын мэдлэгийн графыг эмбеддинг хийхдээ холбоо хамаарал тус бүрийг ялгаатай орон зайд тусгаарлаж тооцоолдог TransR алгоритмыг ашигладаг

- **Тархаах давхарга:**

Энэ давхарга нь өөр оройнуудад олон удаа хийгдэж өндөр зэрэглэлийн холболтуудыг илрүүлж хэрэглэгч бүрд тохирсон хувийн санал болгох боломжийг олгодог. Энэ давхарга дараах 3 дэд алхмаас бүрддэг:

- Мэдээлэл цуглуулах:

Энэ алхамд орой бүрийг илэрхийлэх векторыг шинэчлэхийн тулд хөрш оройнуудаас мэдээлэл цуглуулдаг. Хэрэв хэрэглэгчийг илэрхийлсэн орой байвал тухайн хэрэглэгчийн шинж чанарууд болон өмнө авсан үйлчилгээний мэдээллийг, харин үйлчилгээг илэрхийлсэн байвал тухайн үйлчилгээний шинж чанарууд болон өмнө тухайн үйлчилгээг авсан хэрэглэгчийн мэдээллийг тус тус цуглуулна. Энэхүү мэдээлэл цуглуулах механизм нь графаар дамжин холбоо хамаарлуудыг илүү тодорхой болгож санал болголтын чанарыг сайжруулах

үүрэгтэй. Ингэснээр зөвхөн шууд холбогдсон оройнуудаас бус олон орой дамжсан холбоо хамаарлуудын нөлөөг харгалзан үзэж илүү оновчтой таамаглал гаргах боломжтой болдог.

– Анхаарал механизм хэрэгжүүлэх:

Тухайн нөхцөл байдлаас шалтгаалан бүх холбоо хамаарлууд ижил чухал биш тул хамгийн чухал хөршүүдийг онцлохын тулд анхаарлын механизм ашигладаг. Энэ алхам нь өөр өөр холбоо хамаарлуудад өөр өөр жин оноож хүчтэй холбоо хамаарлуудыг санал болголтод илүү их нөлөөлөх боломжийг олгодог. Жишээлбэл, хэрэв хэрэглэгч тогтмол нэг зохиолчийн бичсэн номыг сонирхдог бол энэ хэрэглэгчийн тухайн зохиолчтой холбоотой хамаарлыг илүү өндөр жинтэй гэж тооцож хамааралгүй холбооснуудаас илүү чухалчилдаг. Ингэснээр санал болгох үйл явц хэрэглэгчийн зан төлөвийг нарийвчлалтай ойлгох боломжтой болдог.

– Мэдээлэл нэгтгэх:

Мэдээлэл цуглуулж чухал холбоо хамаарлуудыг тодорхойлсны дараа орой бүрийг илэрхийлэх векторыг шинэчлэн тодорхойлохын тулд мэдээлэл нэгтгэх алхмаар дамжуулан нэгтгэдэг. Энэ алхам нь өндөр хамааралтай шинж чанаруудыг ялгаж шаардлагагүй холбоо хамаарлуудыг шүүж хасах замаар санал болголтын чанарыг сайжруулдаг. Ингэснээр хамгийн үнэ цэнтэй мэдээллийг хадгалж илүү оновчтой таамаглал гаргах боломжтой болдог.

• **Таамаглах давхарга:**

Олон удаагийн тархаах давхаргууд ажилласны дараа оролтын хэрэглэгч болон үйлчилгээг илэрхийлсэн вектор илүү нарийн, боловсронгуй болж сайжирдаг [1, 2]. Таамаглах давхарга нь эдгээр эцсийн эмбеддингүүдийг нэгтгэн тухайн хэрэглэгч тухайн үйлчилгээг авах магадлалыг тооцоолно.

Энэ судалгааны ажлын хүрээнд цахим төрийн үйлчилгээ санал болгох системийн загвар

хөгжүүлэлтэд ашиглах аргаар мэдлэгийн граф дахь өндөр түвшний холболтуудыг загварчлах чадамжтай тул Мэдлэгийн графт анхааралтай сүлжээ аргыг сонгосон. DKN нь санал болгох зүйлээ тайлбарлахын тулд нэмэлт гаднын мэдлэгийн граф ашигладаг тул E-Mongolia шиг бүтэцчилсэн, текстэн бус өгөгдөлд тохиромжгүй. KGAT нь олон үсрэлтийн холболтуудаар дамжуулан ембеддингүүдийг дахин давтан дамжуулах замаар мэдлэгийн графын бүтцийг шууд ашигладаг. Энэ нь хэрэглэгч, үйлчилгээ, болон нэгжүүдийн хооронд илүү гүн гүнзгий, утга төгөлдөр холбооснуудыг илрүүлэх боломжийг олгодог. MKR нь санал болголт болон мэдлэгийн графын сургалтыг тусдаа сүлжээнүүдээр хийдэг. KGAT нь анхаарлын механизм ашигласнаар хөрш орших оройнуудын чухлыг оновчтой үнэлэх боломжийг бүрдүүлж, илүү нарийн, тайлбарлах боломжтой зөвлөмжийг үүсгэдэг. Эдгээр давуу талууд нь KGAT-г E-Mongolia порталын өгөгдөлд тохиромжтой арга болгож байна.

2. ЦАХИМ ҮЙЛЧИЛГЭЭГ САНАЛ БОЛГОХ ЗАГВАР

2.1 E-Mongolia портал

Бакалаврын судалгааны ажлаар E-Mongolia порталын өгөгдлийг ашиглан мэдлэгийн графт түшиглэсэн санал болгох системийн загварын туршилт хийсэн. Энэ бүлэгт E-Mongolia порталын товч танилцуулга, цахим үйлчилгээний өгөгдөл, болон хүснэгтэн өгөгдлөөс мэдлэгийн граф үүсгэх загварын талаар танилцуулна. E-Mongolia нь иргэд төрийн үйлчилгээг нэгдсэн, хялбар байдлаар цахимаар авах боломжийг олгодог Монгол улсын албан ёсны цахим платформ юм. Энэхүү системийг төрийн үйлчилгээний үр ашгийг нэмэгдүүлэх, хүнд суртлын саад бэрхшээлийг бууруулах, иргэдийн цаг хугацаа, зардлыг хэмнэх зорилгоор нэвтрүүлсэн. Тус портал нь хэрэглэгчдэд төрийн олон төрлийн үйлчилгээг цахимаар авах боломжийг олгож, биечлэн очих шаардлагагүй болгож байна. E-Mongolia нь хувийн бичиг баримт, бизнесийн үйлчилгээ, боловсрол, болон эрүүл мэнд зэрэг төрийн байгууллагуудын 2000 гаруй төлбөртэй болон төлбөргүй хэлбэрийн цахим үйлчилгээг үзүүлдэг.

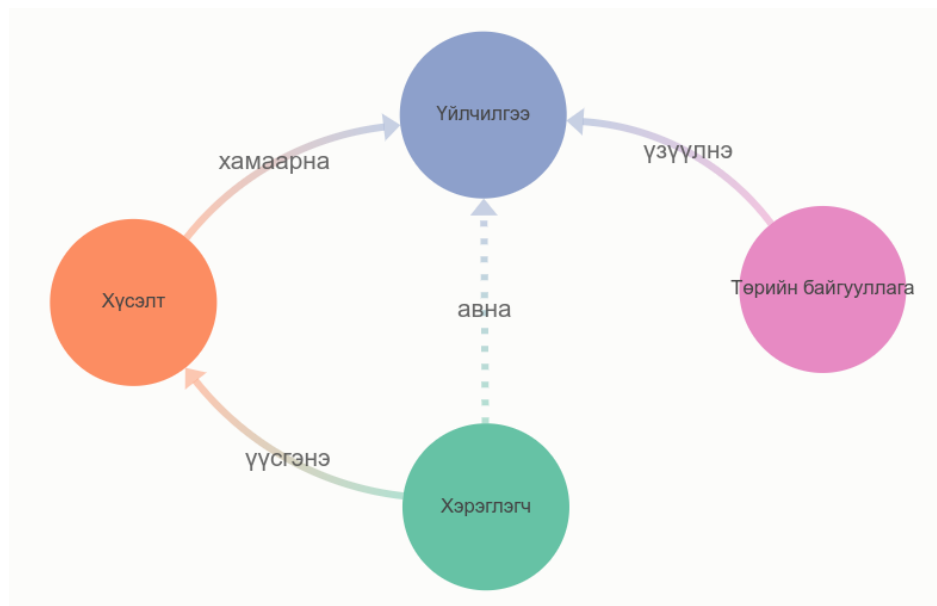
Энэ платформыг гар утасны аппликейшн болон веб хэлбэрээр ашиглах боломжтой бөгөөд хэрэглэгчид регистрийн дугаар, тоон гарын үсэг, банкны эрх зэргийг ашиглан аюулгүй байдлыг хангасан байдлаар нэвтэрч болно. E-Mongolia нь зөвхөн дотоодын иргэд төдийгүй гадаадын иргэдэд хүртээмжтэй байх үүднээс монгол болон англи хэлний дэмжлэгтэй ажилладаг. E-Mongolia нь Монгол улсын цахим засаглалд томоохон хувь нэмрээ оруулж, бүх иргэдэд төрийн үйлчилгээг илүү хүртээмжтэй, хялбар болгож байна.

2.2 Цахим үйлчилгээний өгөгдөл

Туршилтад ашиглах E-Mongolia порталын цахим үйлчилгээний өгөгдөл нь таван хүснэгтээс бүрдэнэ. Аюулгүй байдлын үүднээс зарим утга нь E-Mongolia системийн жинхэнэ утгууд бус таних боломжгүй болгосон, нэргүйжүүлсэн утгууд болно. Хүснэгт 2.1-т үзүүлсэн E-Mongolia порталын өгөгдлийн бүрдлээс Зураг 2.1-т үзүүлсэн мэдлэгийн графыг үүсгэх боломжтой.

Хүснэгт 2.1: E-Mongolia порталын өгөгдлийн бүрдэл

Нэр	Мөрийн тоо	Баганын тоо
Хэрэглэгч	50,000	4
Үйлчилгээ	2,620	7
Хэрэглэгчийн авсан үйлчилгээний түүх	2,163,544	7
Нэгж үйлчилгээний үйл явц	1,048,575	4
Төрийн үйлчилгээний байгууллага	656	5



Зураг 2.1: Мэдлэгийн графын загвар

2.3 Мэдлэгийн графын загвар

Энэхүү туршилтын гол зорилго нь ашиглагдах мэдлэгийн графыг туршилт бүрээр өргөжүүлж, цахим төрийн үйлчилгээг санал болгох системийн загварт янз бүрийн холбооснууд хэрхэн нөлөөлж байгааг тодорхойлох явдал юм. Энэ бүлэгт туршилт бүрийн хувьд ашигласан мэдлэгийн графын загварын онцлог, түүний бүтцийн талаар тайлбарлана. Мэдлэгийн графт түшиглэсэн санал болгох системийн загвар сургах туршилтад дахин ашиглах боломжтой Мэдлэгийн графт анхааралтай сүлжээ (KGAT) аргыг сонгосон. E-Mongolia системийн өгөгдлийг сонгосон аргадаа тохируулан бэлтгэсэн ба KGAT арга нь хэрэглэгчдийн өмнө нь авсан үйлчилгээний түүх буюу хэрэглэгч болон цахим төрийн үйлчилгээ хоорондын харилцааг илэрхийлэх хоёр хэсэгт граф болон түүнд нэмэлт ойлголт, баяжуулалт өгөх мэдлэгийн графыг оролтдоо ашиглан хамтын мэдлэгийн граф үүсгэн загварыг сургахад ашигладаг. Туршилт бүрд ашиглагдах хоёр хэсэгт граф буюу хэрэглэгчийн өмнө авсан үйлчилгээний түүх ижил байх ба нийт 50000 хэрэглэгч 2785 үйлчилгээнээс авсан түүхийг илэрхийлнэ. Зураг 2.1 дахь үүсгэх боломжтой ерөнхий загвараас оройн төрөл, ирмэгийн төрөл, гурвалын хэмжээгээр ялгаатай мэдлэгийн графуудыг үүсгэн туршсан ба туршилтуудыг ашиглагдсан мэдлэгийн графаар нь нэрлэн дараах байдлаар тодорхойлсон:

1. Иргэн-үйлчилгээ + байгууллага:

Энэ туршилтад ашигласан мэдлэгийн граф нь зөвхөн нэг төрлийн ирмэг болох $[Үйлчилгээ]—[үзүүлдэг]→[Байгууллага]$ холбоог хамаарлыг агуулна.

2. Иргэн-үйлчилгээ + байгууллага + төрөл:

Энэ туршилтад хоёр хэсэгт граф дахь үйлчилгээг илэрхийлсэн оройнуудад нэмэлт мэдээлэл болох $[Үйлчилгээ]—[үзүүлдэг]→[Байгууллага]$, $[Үйлчилгээ]—[төрлийн]→[Үйлчилгээний төрөл]$ хоёр төрлийн ирмэгтэй мэдлэгийн графыг ашиглан туршсан.

3. Иргэн-үйлчилгээ + насны бүлэг:

Эхний туршилттай адил энэ туршилтад ашиглагдах мэдлэгийн граф нь нэг төрлийн ирмэгтэй. Энэ туршилтад үйлчилгээтэй бус хэрэглэгчтэй холбоотой *[Хэрэглэгч]—[хамаарна]→[Насны бүлэг]* ирмэгийг нэмэлт мэдээлэл болгон мэдлэгийн графт оруулсан. 50,000 хэрэглэгчийн насны төвүүдийг олж 27-оос доош, 27→34, 35→41, 42→50, 51→60, 60-аас дээш гэсэн 6 насны бүлэгт хуваасан.

4. Иргэн-үйлчилгээ + насны бүлэг + байгууллага + төрөл + сар:

Энэ туршилтад ирмэгийн төрлийн тоогоор хамгийн том мэдлэгийн графыг ашиглан туршсан бөгөөд граф нь дараах дөрвөн төрлийн ирмэгүүдийг агуулна.

- *[Үйлчилгээ]—[үзүүлдэг]→[Байгууллага]*
- *[Үйлчилгээ]—[төрлийн]→[Үйлчилгээний төрөл]*
- *[Үйлчилгээ]—[хамгийн их авдаг]→[сар]*
- *[Хэрэглэгч]—[хамаарна]→[Насны бүлэг]*

Дөрвөн ялгаатай бүтэц бүхий мэдлэгийн графууд буюу дөрвөн удаагийн туршилтын мэдлэгийн графын оройн тоо, ирмэгийн төрлийн тоо, гурвалын тоог Хүснэгт 2.2-д харуулав.

Хүснэгт 2.2: Туршилтуудад ашиглагдсан мэдлэгийн графын хэмжээ

Туршилт	Оройн тоо	Ирмэгийн төрлийн тоо	Гурвалын тоо
Иргэн-үйлчилгээ + байгууллага	3,227	1	2,620
Иргэн-үйлчилгээ + байгууллага + төрөл	3,235	2	5,240
Иргэн-үйлчилгээ + насны бүлэг	50,007	1	50,000
Иргэн-үйлчилгээ + насны бүлэг + байгууллага + төрөл + сар	53,254	4	56,024

3. ТУРШИЛТ

E-Mongolia порталын өгөгдлөөс 4 ялгаатай бүтэц бүхий мэдлэгийн граф үүсгэн тус бүрд суурилсан Мэдлэгийн графт анхааралтай сүлжээ (KGAT) аргыг ашиглан сургасан санал болгох загварын үр дүнг $recall@k$, $precision@k$, $hit@k$, $ndcg@k$ хэмжигдэхүүнүүдээр үнэлэн дараах гурван асуултад хариулт олохыг зорив.

1. Хамгийн олон буюу дөрвөн ирмэгийн төрөл агуулсан мэдлэгийн граф ашигласан туршилт нь хамгийн сайн үр дүн үзүүлэх үү?
2. Мэдлэгийн граф дахь ирмэгийн төрөл (холбоо хамаарал) нэмэгдэх тусам загварын таамаглах чадвар сайжрах уу?
3. Мэдлэгийн граф дахь ялгаатай ирмэгүүдийн төрөл нь санал болгох үр дүнд ялгаатай нөлөө үзүүлэх үү?

3.1 Үнэлэх аргачлал

Туршилтын үр дүнгээр санал болгох системийн загварын гүйцэтгэлийг үнэлж аль мэдлэгийн графын хувилбар илүү сайн үр дүн үзүүлж байгааг дараах хэмжигдэхүүнүүдээр харьцуулан үнэлсэн:

- $Recall@K$:

Загвар K ширхэг үйлчилгээ санал болгох үед тухайн хэрэглэгчид санал болгох ёстой байсан үйлчилгээнүүдээс хэд нь багтсан байгааг илэрхийлэх харьцаа. Энэ нь системийн хамрах чадварыг хэмждэг.

- $Precision@K$:

Загварын санал болгосон K ширхэг үйлчилгээний хэдэн хувь нь зөв байсан болохыг илэрхийлдэг. Энэ нь системийн нарийвчлалыг илтгэнэ.

- **Hit@K:**

Санал болгосон K үйлчилгээ дотор ядаж нэг зөв үйлчилгээ орсон эсэхийг шалгадаг үнэлгээний үзүүлэлт. Энэ нь систем хэрэглэгчид хэрэгтэй мэдээллийг ерөнхийд нь хүргэж чадаж байгааг илтгэнэ.

- **NDCG@K:**

Санал болгох ёстой үйлчилгээнүүдэд эрэмбэ тогтоож, K ширхэг үйлчилгээ санал болгох явцад тухайн эрэмбийн дагуу хэр оновчтой санал болгосныг хэмжих үнэлгээ. Зөв үйлчилгээнүүд санал болголтын эрэмбийн дээд хэсэгт байршсан эсэхийг тооцдог. Энэ нь зөвхөн зөв санал болгох чадавхыг бус, тэдгээрийг эрэмбэлэх оновчтой байдлыг харуулна.

3.2 Өгөгдлийг загварын сургалт болон туршилтад хуваах

Уламжлалт машин сургалтын загварууд, тухайлбал объект илрүүлэлт зэрэгт сургалт болон туршилтын өгөгдлийг тодорхой харьцаагаар (жишээлбэл 80:20, 70:30 гэх мэт) санамсаргүй байдлаар хувааж ашигладаг. Энэ нь тухайн загварын чадварыг үнэлэх нийтлэг аргачлал бөгөөд нэгэн төрлийн өгөгдлийн тархалттай үед үр дүнтэй. Гэвч хэрэглэгчийн хэрэгцээ шаардлага, үйлчилгээний сонголт зэрэг цаг хугацаанаас хамаарсан өгөгдөлд энэ аргачлал төдийлөн тохиромжгүй байдаг. Мэдлэгийн графт суурилсан санал болгох системүүд нь хэрэглэгч ба үйлчилгээний хоорондын хамаарлыг хоёр хэсэгт граф хэлбэрээр загварчилдаг. Өгөгдлийг санамсаргүй хуваах арга нь графын бүтэц, холбоосыг хэт задлан, загварын сургалтын үр дүнд сөргөөр нөлөөлөх магадлалтай. Графт суурилсан, цаг хугацаанаас хамаарсан загваруудыг сургах, туршихдаа хэрэглэгч бүрийн үйлчилгээний түүхийг цаг хугацааны дараалалд үндэслэн хуваадаг. Хоёр хэсэгт графыг хэрэглэгч бүрийн хувьд хугацааны дарааллыг харгалзан үзэж тодорхой харьцаагаар хуваан ихэнх хэсгийн сургалтад ашигладаг. Харин үлдсэн хугацааны агшин дахь хэрэглэгчийн авсан үйлчилгээний түүхийг туршилтад ашигладаг бөгөөд өнгөрсөн хугацааны агшин дахь графт

суурилан сурсан загвар туршилтын хугацааны агшинд хамаарах үйлчилгээг зөв таамаглан санал болгох чадварыг нь үнэлдэг. Өгөгдлийн энэ төрлийн хуваалт нь хэрэглэгчийн шаардлагыг илүү сайн загварчлах, мөн загварын оновчтой таамаглах чадварыг илүү нарийн үнэлэх давуу талтай. Энэхүү судалгааны ажлын E-Mongolia порталын өгөгдлөөс үүсгэсэн хоёр хэсэгт графыг энэ хуваалтын аргаар хугацааны хувьд 80:20 харьцаатайгаар хуваасан. Мэдлэгийн графт анхааралтай сүлжээ арга нь хоёр хэсэгт графаас гадна нэмэлт мэдээлэл бүхий мэдлэгийн графыг ашиглаж хамтын мэдлэгийн графыг үүсгэдэг. Мэдлэгийн граф нь хугацааны агшнаас хамааран өөрчлөгдөхгүй тул сургалт болон туршилтад хуваагдахгүй.

3.3 Үр дүн

Загварыг үнэлэх хэмжигдэхүүнүүдийн $k=20$ байх үеийн 5 туршилтын үр дүнг Хүснэгт 3.1-д харуулав.

Table 3.1: $K=20$ байх туршилтын үр дүн

Туршилт	Recall	Precision	Hit	NDCG
Иргэн-үйлчилгээ + байгууллага	0.2577	0.0663	0.6591	0.3068
Иргэн-үйлчилгээ + байгууллага + төрөл	0.2581	0.0664	0.6569	0.3082
Иргэн-үйлчилгээ + насны бүлэг	0.2666	0.0654	0.6515	0.3159
Иргэн-үйлчилгээ + насны бүлэг + байгууллага + төрөл + сар	0.2597	0.0669	0.6626	0.3097

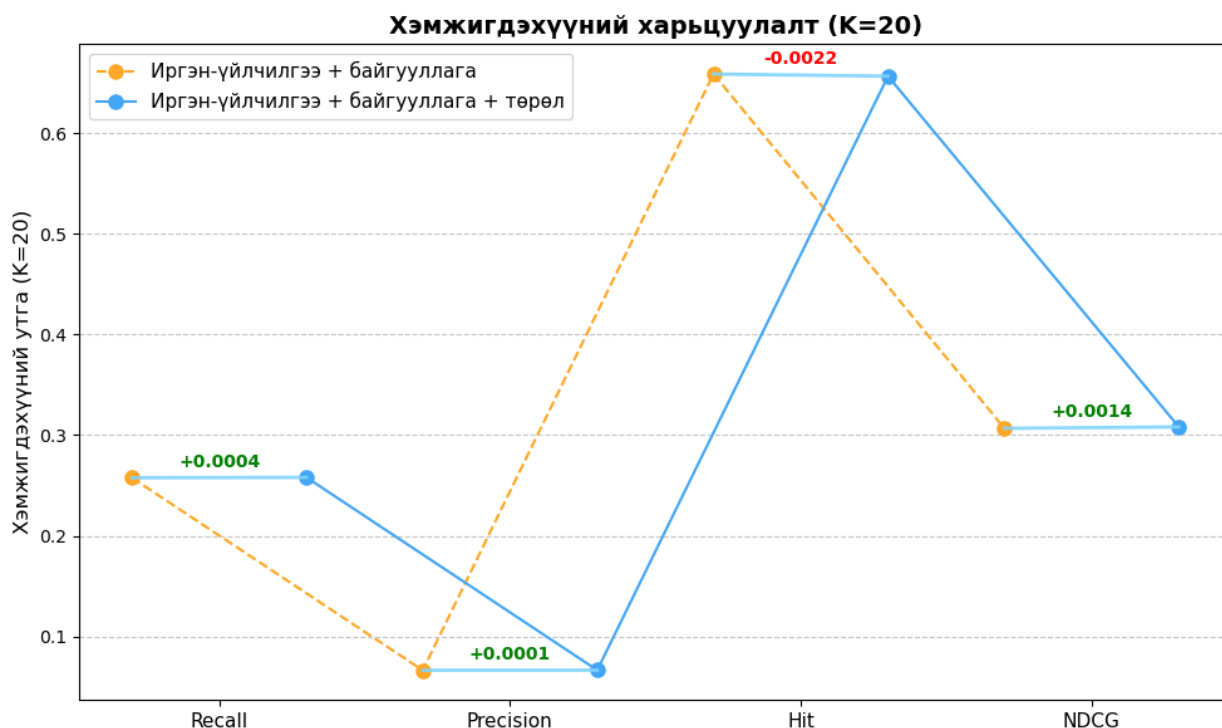
CA1: Хамгийн олон ирмэгийн төрөл агуулсан мэдлэгийн граф

Мэдлэгийн графын аль хувилбарыг ашигласан санал болгох загвар нь хамгийн оновчтой таамагладаг, сайн үр дүнтэй болохыг Хүснэгт 3.1-д харуулсан дөрвөн удаагийн туршилтын $K=20$ байх үеийн загварыг үнэлэх хэмжигдэхүүнүүдийн утгаас ажигласан. Туршилтууд дундаас хамгийн олон буюу 4 ирмэгийн төрөл агуулсан мэдлэгийн графыг ашиглан туршсан *[Иргэн-үйлчилгээ + байгууллага + насны бүлэг + төрөл + сар]* туршилтын загварын pre-

cision@20, hit@20 утгууд нь хамгийн өндөр байна. Харин recall@20, ndcg@20 утгууд нь [Иргэн-үйлчилгээ + насны бүлэг] туршилтаас бага зэрэг доогуур байна. Энэ нь олон холбоост мэдлэгийн графыг нэмэлт мэдээлэл болгох нь KGAT аргыг өндөр зэрэглэлийн холбоо хамаарлыг үр дүнтэй илрүүлэх, санал болгох нарийвчлал болон зөв таамаглах чадварыг нэмэгдүүлдэг болохыг илтгэж байна.

CA2: Ирмэгийн төрөл нэмэгдэх үеийн нөлөө

[Иргэн-үйлчилгээ + байгууллага] туршилтад ашиглагдсан мэдлэгийн графт үйлчилгээний хамаарах төрлийг илэрхийлэх ирмэгийн төрөл нэмж баяжуулж [Иргэн-үйлчилгээ + байгууллага + төрөл] туршилтад ашигласан. Нэг ирмэгийн төрөл нэмж баяжуулснаар мэдлэгийн графт 8 орой, 2620 гурвал нэмэгдсэн. Бага тоо хэмжээгээр өргөжсөн хэдий ч загварыг үнэлэх recall@20, precision@20, ndcg@20 хэмжигдэхүүнүүд жигд өсөлттэй байгаа нь мэдлэгийн графт ирмэгийн төрөл нэмж баяжуулах нь санал болгох загварыг гүйцэтгэлд эергээр нөлөөлдөг болох нь харагдаж байна.



Зураг 3.1: Ирмэгийн төрөл нэмэгдэх үеийн өөрчлөлт

СА3: Ялгаатай ирмэгийн төрлийн нөлөө

Мэдлэгийн графт анхааралтай сүлжээ аргаар хөгжүүлсэн санал болгох загварын гүйцэтгэл мэдлэгийн граф дахь ирмэгийн төрлийг нэмж баяжуулахад сайжирч байгаа нь эхний хоёр туршилтын үр дүнгээс харагдсан. *[Иргэн-үйлчилгээ + насны бүлэг]* туршилт нь зөвхөн нэг ирмэгийн төрөл агуулсан граф ашигласан боловч харгалзах recall@20 болон ndcg@20 хэмжигдэхүүн нь хамгийн өндөр байна. Гэсэн хэдий ч энэ графыг хоёр ирмэгийн төрлөөр өргөжүүлсэн туршилтын үр дүн нь зөвхөн хэрэглэгчийн хамаарах насны бүлгийг илэрхийлэх ирмэгийн төрлийг агуулсан туршилтаас recall@20, ndcg@20 хэмжигдэхүүний утгууд нь бага байгаа нь мэдлэгийн графт агуулагдаж байгаа ирмэгийн төрлүүдийн хоорондын нийцтэй байдал, ямар нэмэлт мэдээллийг өгөх ирмэгийн төрөл зэргээс хамааран санал болгох загварын үр дүнд ялгаатай нөлөө үзүүлэх ба зарим тохиолдолд сөрөг нөлөө үзүүлэх боломжтой нь харагдаж байна.

Өгөгдлийн харьцуулалт

KGAT аргын зохиогчид тус аргыг Amazon-book, Last-FM, Yelp2018 зэрэг нээлттэй өгөгдлийн сангуудад туршиж үр дүнг харьцуулсан. Энэхүү судалгаанд E-Mongolia порталаас авсан өгөгдөл дээр суурилан сургасан KGAT загварын гүйцэтгэлийг дээрх өгөгдлийн сангуудын үр дүнтэй харьцуулсан бөгөөд Хүснэгт 3.2-д үзүүлэв.

Table 3.2: Өгөгдлүүдийн үр дүнгийн харьцуулалт

	Amazon-book	Last-FM	Yelp2018	E-Mongolia
Recall@20	0.1489	0.0870	0.0712	0.2597
ndcg@20	0.1006	0.1325	0.0867	0.3097

Хүснэгтээс харахад E-Mongolia өгөгдөлд сургасан загварын ndcg@20 үзүүлэлт нь бусад өгөгдлийн санд суурилан сургасан загваруудтай харьцуулахад хамгийн өндөр байгаа нь санал болгосон үйлчилгээний эрэмбийг хэрэглэгчийн хэрэгцээнд нийцүүлэн оновчтой тодорхойлж байгааг харуулж байна. Recall@20 үзүүлэлт нь өндөр байгаа нь

хэрэглэгчид хамааралтай үйлчилгээнүүдийг хангалттай хэмжээнд санал болгож байгааг илтгэж байна. Мөн `ndcg@20` үзүүлэлт өндөр байгаа нь хэрэглэгчид хамааралтай үйлчилгээнүүдийг зөв эрэмбийн дагуу санал болгож байгааг илэрхийлнэ. Эдгээр үр дүн нь E-Mongolia порталын өгөгдлөөс үүсгэсэн мэдлэгийн граф нь сайн холбогдсон, мэдээлэл сайн байгаагаас шалтгаалан KGAT аргын анхаарлын механизм хэрэгжүүлэн өндөр зэрэглэлийн холболтуудыг илрүүлэхэд тохиромжтой болж загварын үр дүнд нөлөөлж байгааг харууллаа.

4. ДҮГНЭЛТ

Туршилтын үр дүнгээс мэдлэгийн графт түшиглэсэн санал болгох системийн загвар сургахад ашиглах мэдлэгийн графыг өргөжүүлэх буюу нэмэлт мэдээллээр хангаж хэрэглэгч болон үйлчилгээ хоорондын холбоо хамаарлыг илүү тайлбарлах тусам загварын гүйцэтгэл сайжрах ба тэр дундаа хэрэглэгчид шаардлагатай үйлчилгээг илүү олон таамаглах чадвар нэмэгдэнэ. Нэмэлт мэдээлэл хэлбэрээр ашиглагдах мэдлэгийн граф дахь холбоо хамаарал буюу ирмэгийн төрөл бүр нь загварын гүйцэтгэлд ялгаатай нөлөө үзүүлдэг ба зарим тохиолдолд сөргөөг нөлөөлөх боломжтой учраас мэдлэгийн графыг баяжуулах ирмэгийн төрлийг зөв сонгох нь чухал. Судалгааны үр дүн дэх хамгийн сайн гүйцэтгэлтэй KGAT загварын Recall@20 үзүүлэлт 0.2597 буюу хэрэглэгч бүрийн хувьд системийн санал болгосон 20 үйлчилгээ дотор тухайн хэрэглэгчид хамааралтай үйлчилгээний 25 хувийг дунджаар зөв таамаглаж байгааг илтгэнэ. Өөрөөр хэлбэл, хэрэглэгчийн ирээдүйд авах боломжтой бүх үйлчилгээнүүдээс тодорхой хэсгийг систем амжилттай санал болгож байгааг илэрхийлэх хэмжигдэхүүн юм. Мөн E-Mongolia порталын өгөгдөлд суурилан KGAT аргыг ашигласан санал болгох загварын хувьд мэдлэгийн граф нь хэрэглэгчээс илүүтэй үйлчилгээг тодотгосон нэмэлт мэдээллийг агуулах нь үр дүнд эерэг нөлөө үзүүлэх нь харагдсан. Цаашид санал болгох загварын сургалтад ашиглах мэдлэгийн графыг зөв холбоосоор баяжуулж загварын гүйцэтгэлийг улам сайжруулах боломжтой.

Bibliography

- [1] J. Liu and L. Duan, “A survey on knowledge graph-based recommender systems,” in *2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 2021, pp. 2450–2453.
- [2] Q. Guo, F. Zhuang, C. Qin, H. Zhu, X. Xie, H. Xiong, and Q. He, “ A Survey on Knowledge Graph-Based Recommender Systems ,” *IEEE Transactions on Knowledge & Data Engineering*, vol. 34, no. 08, pp. 3549–3568, Aug. 2022. [Online]. Available: <https://doi.ieeecomputersociety.org/10.1109/TKDE.2020.3028705>
- [3] X. Wang, X. He, Y. Cao, M. Liu, and T.-S. Chua, “Kgat: Knowledge graph attention network for recommendation,” in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery Data Mining*, ser. KDD ’19. ACM, Jul. 2019, p. 950–958. [Online]. Available: <http://dx.doi.org/10.1145/3292500.3330989>
- [4] H. Wang, F. Zhang, X. Xie, and M. Guo, “Dkn: Deep knowledge-aware network for news recommendation,” 2018. [Online]. Available: <https://arxiv.org/abs/1801.08284>
- [5] H. Wang, F. Zhang, M. Zhao, W. Li, X. Xie, and M. Guo, “Multi-task feature learning for knowledge graph enhanced recommendation,” 2019. [Online]. Available: <https://arxiv.org/abs/1901.08907>

5. КОДЫН ХЭРЭГЖҮҮЛЭЛТ

```
1 import pandas as pd
2
3 citizen_df = pd.read_excel('../..//excel-files/citizen.xlsx')
4 service_df = pd.read_excel('../..//excel-files/emon.service.xlsx')
5 request_df_sheets = pd.read_excel('../..//excel-files/request.xlsx',
    sheet_name=None)
6
7 request_df = pd.concat(request_df_sheets.values())
8 request_df.rename(columns={"createc_date": "created_date", "service_id"
    : "serviceid"}, inplace=True)
9 service_df.rename(columns={"_id": "serviceid"}, inplace=True)
10 request_df["created_date"] = pd.to_datetime(request_df["created_date"])
11
12 all_users = set(citizen_df['userid']).union(set(request_df['userid']))
13 all_services = set(service_df['serviceid']).union(set(request_df['
    serviceid']))
14
15 print(len(all_users), len(all_services))
16
17 user_map = {user_id: idx for idx, user_id in enumerate(sorted(all_users
    ))}
18 service_map = {service_id: idx for idx, service_id in enumerate(sorted(
    all_services))}
19
20 with open('user_list.txt', 'w') as f:
```

```

21     for user, idx in user_map.items():
22         f.write(f"{user} {idx}\n")
23
24 with open('item_list.txt', 'w') as f:
25     for service, idx in service_map.items():
26         f.write(f"{service} {idx}\n")
27
28 request_df['userid'] = request_df['userid'].map(user_map)
29 request_df['serviceid'] = request_df['serviceid'].map(service_map)
30
31 request_df = request_df.sort_values(by='created_date')
32
33 split_idx = int(len(request_df) * 0.8)
34 train_df = request_df.iloc[:split_idx]
35 test_df = request_df.iloc[split_idx:]
36
37 missing_users = set(test_df['userid']) - set(train_df['userid'])
38 if missing_users:
39     extra_train_data = test_df[test_df['userid'].isin(missing_users)]
40     train_df = pd.concat([train_df, extra_train_data])
41     test_df = test_df[~test_df['userid'].isin(missing_users)]
42
43 def save_format(filename, data):
44     grouped = data.groupby('userid')['serviceid'].apply(lambda x: ' '.
45         join(map(str, x)))
46     with open(filename, 'w') as f:
47         for user, items in grouped.items():
48             f.write(f"{user} {items}\n")

```

```

48 save_format('train.txt', train_df)
49 save_format('test.txt', test_df)
50
51 print(" Preprocessing complete! Files saved: user_list.txt, item_list.
      txt, train.txt, test.txt")

```

Код 5.1: Хэрэглэгчийн өмнө авсан үйлчилгээний түүхээс сургалтын болон туршилтын хоёр хэсэгт граф үүсгэх код

```

1 num_users = len(user_map)
2 num_items = len(service_map)
3
4 all_agency = set(service_df['govAgencyId'])
5 kg_entity_map = {agency_id: idx for idx, agency_id in enumerate(sorted(
      all_agency))}
6
7 entity2id = {}
8
9 for service, idx in service_map.items():
10     entity2id[service] = idx
11
12 offset = len(service_map) + len(user_map)
13 for other_entity, idx in kg_entity_map.items():
14     entity2id[other_entity] = idx + offset
15
16 relation2id = {
17     "provided_by": 0,
18 }
19

```

```

20 kg_triples = []
21
22 for _, row in service_df.iterrows():
23     service_id = entity2id[row['serviceid']]
24     agency_id = entity2id[row['govAgencyId']]
25     kg_triples.append(f"{service_id} {relation2id['provided_by']} {
        agency_id}")
26
27 with open("kg_final.txt", "w") as f:
28     for triple in kg_triples:
29         f.write(triple + "\n")
30
31 with open("entity_list.txt", "w") as f:
32     for ent, eid in entity2id.items():
33         f.write(f"{ent} {eid}\n")
34
35 with open("relation_list.txt", "w") as f:
36     for rel, rid in relation2id.items():
37         f.write(f"{rel} {rid}\n")

```

Код 5.2: Үйлчилгээ үзүүлдэг байгууллагын мэдээллийг агуулсан мэдлэгийн граф үүсгэх код

```

1 num_users = len(user_map)
2 num_items = len(service_map)
3
4 all_agency = set(service_df['govAgencyId'])
5 all_service_types = set(service_df["serviceType"])
6 kg_entity_map = {agency_id: idx for idx, agency_id in enumerate(sorted(
    all_agency))}

```

```

7 kg_entity_map.update({service_type: idx + len(all_agency) for idx,
    service_type in enumerate(sorted(all_service_types))})
8
9 entity2id = {}
10
11 for service, idx in service_map.items():
12     entity2id[service] = idx
13
14 offset = len(service_map) + len(user_map)
15 for other_entity, idx in kg_entity_map.items():
16     entity2id[other_entity] = idx + offset
17
18 relation2id = {
19     "provided_by": 0,
20     "type_of": 1,
21 }
22
23 kg_triples = []
24
25 for _, row in service_df.iterrows():
26     service_id = entity2id[row['serviceid']]
27     agency_id = entity2id[row['govAgencyId']]
28     service_type_id = entity2id[row['serviceType']]
29     kg_triples.append(f"{service_id} {relation2id['provided_by']} {
        agency_id}")
30     kg_triples.append(f"{service_id} {relation2id['type_of']} {
        service_type_id}")
31

```

```

32 with open("kg_final.txt", "w") as f:
33     for triple in kg_triples:
34         f.write(triple + "\n")
35
36 with open("entity_list.txt", "w") as f:
37     for ent, eid in entity2id.items():
38         f.write(f"{ent} {eid}\n")
39
40 with open("relation_list.txt", "w") as f:
41     for rel, rid in relation2id.items():
42         f.write(f"{rel} {rid}\n")

```

Код 5.3: Үйлчилгээ үзүүлдэг байгууллага + үйлчилгээний хамаарах төрөл мэдээллийг агуулсан мэдлэгийн граф үүсгэх код

```

1 num_users = len(user_map)
2 num_items = len(service_map)
3
4 all_age_group = set(["unknown", "under_27", "27_34", "35_41", "42_50",
5     "51_60", "over_60"])
6
7 kg_entity_map = {age_group: idx for idx, age_group in enumerate(sorted(
8     all_age_group))}
9
10 entity2id = {}
11
12 for user, idx in user_map.items():
13     entity2id[user] = idx + len(service_map)
14
15 offset = len(user_map) + len(service_map)

```

```

13 for other_entity, idx in kg_entity_map.items():
14     entity2id[other_entity] = idx + offset
15
16 def assign_age_group(age):
17     if pd.isna(age) or not isinstance(age, (int, float)):
18         return "unknown"
19     if age < 27:
20         return "under_27"
21     elif age < 35:
22         return "27_34"
23     elif age < 42:
24         return "35_41"
25     elif age < 51:
26         return "42_50"
27     elif age < 61:
28         return "51_60"
29     else:
30         return "over_60"
31 citizen_df['age_group'] = citizen_df['age'].apply(assign_age_group)
32
33 relation2id = {
34     "belongs_to": 0,
35 }
36
37 kg_triples = []
38
39 for _, row in citizen_df.iterrows():
40     user_id = entity2id[row['userid']]

```

```

41     age_group_id = entity2id[row['age_group']]
42     kg_triples.append(f"{user_id} {relation2id['belongs_to']} {
        age_group_id}")
43
44 with open("kg_final.txt", "w") as f:
45     for triple in kg_triples:
46         f.write(triple + "\n")
47
48 with open("entity_list.txt", "w") as f:
49     for ent, eid in entity2id.items():
50         f.write(f"{ent} {eid}\n")
51
52 with open("relation_list.txt", "w") as f:
53     for rel, rid in relation2id.items():
54         f.write(f"{rel} {rid}\n")

```

Код 5.4: Хэрэглэгчийн хамаарах насны бүлэг мэдээллийг агуулсан мэдлэгийн граф үүсгэх код

```

1 num_users = len(user_map)
2 num_items = len(service_map)
3
4 all_agency = set(service_df['govAgencyId'])
5 all_service_types = set(service_df["serviceType"])
6 all_months = set({"month1", "month2", "month3", "month4", "month5", "
    month6", "month7", "month8", "month9", "month10", "month11", "
    month12"})
7 all_age_group = set(["unknown", "under_27", "27_34", "35_41", "42_50",
    "51_60", "over_60"])

```



```

8 len(all_agency), len(all_service_types), len(all_months), len(
    all_age_group)
9
10 kg_entity_map = {agency_id: idx for idx, agency_id in enumerate(sorted(
    all_agency))}
11 kg_entity_map.update({service_type: idx + len(all_agency) for idx,
    service_type in enumerate(sorted(all_service_types))})
12 kg_entity_map.update({month: idx + len(all_agency) + len(
    all_service_types) for idx, month in enumerate(sorted(all_months))})
13 kg_entity_map.update({age_group: idx + len(all_agency) + len(
    all_service_types) + len(all_months) for idx, age_group in enumerate(
    sorted(all_age_group))})
14
15 entity2id = {}
16
17 for service, idx in service_map.items():
18     entity2id[service] = idx
19
20 user_offset = len(service_map)
21 for user, idx in user_map.items():
22     entity2id[user] = idx + user_offset
23
24 offset = len(service_map) + len(user_map)
25 for other_entity, idx in kg_entity_map.items():
26     entity2id[other_entity] = idx + offset
27
28 service_df = service_df.merge(most_requested, on='serviceid', how='left
    ')

```

```

29
30 def assign_age_group(age):
31     if pd.isna(age) or not isinstance(age, (int, float)):
32         return "unknown"
33     if age < 27:
34         return "under_27"
35     elif age < 35:
36         return "27_34"
37     elif age < 42:
38         return "35_41"
39     elif age < 51:
40         return "42_50"
41     elif age < 61:
42         return "51_60"
43     else:
44         return "over_60"
45 citizen_df['age_group'] = citizen_df['age'].apply(assign_age_group)
46
47 relation2id = {
48     "provided_by": 0,
49     "type_of": 1,
50     "most_requested": 2,
51     "belongs_to": 3,
52 }
53
54 kg_triples = []
55
56 for _, row in service_df.iterrows():

```

```

57     service_id = entity2id[row['serviceid']]
58     agency_id = entity2id[row['govAgencyId']]
59     service_type_id = entity2id[row['serviceType']]
60
61     kg_triples.append(f"{service_id} {relation2id['provided_by']} {
        agency_id}")
62     kg_triples.append(f"{service_id} {relation2id['type_of']} {
        service_type_id}")
63
64     if pd.notna(row['month']):
65         month_id = entity2id['month' + str(int(row['month']))]
66         kg_triples.append(f"{service_id} {relation2id['most_requested
        ']} {month_id}")
67
68 for _, row in citizen_df.iterrows():
69     user_id = entity2id[row['userid']]
70     age_group_id = entity2id[row['age_group']]
71     kg_triples.append(f"{user_id} {relation2id['belongs_to']} {
        age_group_id}")
72
73 with open("kg_final.txt", "w") as f:
74     for triple in kg_triples:
75         f.write(triple + "\n")
76
77 with open("entity_list.txt", "w") as f:
78     for ent, eid in entity2id.items():
79         f.write(f"{ent} {eid}\n")
80

```

```
81 with open("relation_list.txt", "w") as f:
82     for rel, rid in relation2id.items():
83         f.write(f"{rel} {rid}\n")
```

Код 5.5: Хэрэглэгчийн хамаарах насны бүлэг + үйлчилгээ үзүүлдэг байгууллага + үйлчилгээний хамаарах төрөл + үйлчилгээг хамгийн их авдаг сар мэдээллийг агуулсан мэдлэгийн граф үүсгэх код