

УДК 004.046(047)

Б.Б. Мўминов

FSV ТЕХНОЛОГИЯСИНИНГ ДОЛЗАРБЛИК БЎЙИЧА ТЕСКАРИ АЛОҚА СТРАТЕГИЯЛАРИ

Ушбу мақолада FSV технологиянинг излаш тизимидаги долзарблик тескари алоқа стратегияларининг асосий амалга ошириш усулларини локал тамойил асосида ошириш баён қилинган. Унда долзарблик ва тасодифий (сохта, алдамчи) долзарблик бўйича тескари алоқа, долзарблик бўйича тескари алоқа учун Роккио алгоритми, долзарблик бўйича эҳтимоли бўлган тескари алоқа, корпоратив тармоқда долзарблик бўйича тескари алоқа, Долзарблик бўйича ошқормас тескари алоқа шунингдек долзарблик бўйича тескари алоқа стратегиясини баҳолаш асослари очиб берилган.

Калит сўзлар: Маълумотларни излаш, излаш тизими, FSV технология, долзарблик, тескари алоқа, сўров, Роккио алгоритми, центроид.

Кириш. Бугунги кунда маълумотларнинг ҳажми ошиб бориш маълумотларни излаш тизимларига (МИТ) бўлган эҳтиёжни кучайтирмоқда. Ҳар ахборот тизими маълумотлар базасига асослангандек, маълумотларни излаш тизимларига ҳам асосланади[1]. Шунинг учун ахборот тизимларида маълумотларни интеллектуал излаш ва қайта ишлаш моделлари ва усулари асосида FSV технологияси яратилди. FSV технология бу махсус маълумотларни излаш тизимини яратишга мўлжалланган платформа бўлиб ҳисоблади. *FSV технологияси* – бу корпоратив тармоқларнинг ахборот муҳитида МИҚИ моделлари, усуллари ва алгоритмларини интеграция ва модификацияловчи, сервер иловали мижоз-сервер ахтitekturасига асосланган инструментал дастурий платформа[1].

Кўп жиҳатли корпоратив тармоқларда FSV технологияси асосида яратилган излаш модуларининг 3 та базавий босқичларини бўлди. Улар:

1. (F) - сўровни тақдим этиш усули, яъни тизим фойдаланувчисининг ахборот эҳтиёжларини ифодалашнинг шакллаштириш усули.
2. (S) - электрон ресурсларнинг сўровга мувофиқлик функцияси, яъни сўровнинг ва топилган электрон ресурсларнинг мувофиқлиги даражаси.
3. (V) – электрон ресурсларни тақдим қилиш усули.

Сараланган маълумотлар тўпламида излаш тизимининг ўқитиш учун долзарблик бўйича тескари алоқани ҳисоблаш муҳим ҳисобланади. Бу FSV технологияси амалга оширишда мавжуд ва такомиллаштирилган долзарбликка тескари алоқани ҳисоблаш моделлари, усуллари ва ёндашувлар асосида стратегияларини аниқлаш муҳимдир.

Ахборот тизимларининг излаш тизимларида бир тушунча турли сўзлар билан ифодаланиши мумкин. Бу ҳодиса синоним (synonym) каби машҳур, кўпгина МИТларнинг излаш тўлиқлик мезонига таъсир қилади. Масалан, фойдаланувчилар [шум] сўровига шунингдек, [шўх] сўзи (лекин фақат инсон маъносида), хабар – мос келувчи контекстда шум хабар сўзи мос келишини хошлашади.

Фойдаланувчилар кўпинча сўровларни аниқлаш-тирган ҳолда бу муаммони мустақил ечишга уринишади. Бу мақолада излаш тизимнинг ўзи ё автоматик, ё фойдаланувчи иштироки билан сўров қайта аниқлаши мумкин бўлган усулларни кўриб чиқамиз.

Асосий қисм. Бу масалани ечиш усуллари икки асосий категорияга бўлинади: глобал ва локал. Глобал усуллар сўров ва қайтарилган натижалардан қатъи назар сўровни кенгайтириш ёки янгидан шакллантирилишини назарда тутаяди, шунинг учун сўров шаклланишидаги ўзгаришлар бошқа семантик яқин терминларга мос келувчи янги сўровнинг ҳосил бўлишига олиб келади. Глобалларга қуйидаги усуллар тегишли.

- Махсус тезаурус ёки WordNet тезаурсини ёрдамида сўровнинг кенгайтириш/янгидан ташкил қилиниши.

- Тезауруснинг автоматик генерациялаши ёрдамида сўровнинг кенгайтириши.

- Нотўғри ёзувларни тўғрилаш усулларига ўхшаш усуллар.

Локал усуллар дастлабки сўров бўйича топилган ЭРларни ҳисобга олган ҳолда сўровни ўзгартиради. Локал усулларга қуйидагилар киради.

- Долзарблик бўйича тескари алоқа.

- Долзарблик бўйича кўп тескари алоқа каби машҳур тасодифий долзарблик бўйича тескари алоқа.

- Долзарблик бўйича (глобал) норавшан (неявная, ошқормас) тескари алоқа.

Бу мақолада FSV технологиясидаги долзарблик бўйича тескари алоқанинг барча ёндашувларга тўхталаямиз, лекин энг кўп тарқалган ва муваффақиятга эришган долзарблик бўйича тескари алоқа усулига эътиборимизни қаратамиз.

1. Долзарблик ва тасодифий (сохта, алдамчи) долзарблик бўйича тескари алоқа. Долзарблик бўйича тескари алоқа (Relevance Feedback - RF) натижаларнинг якуний рўйхатини яхшилаш учун излаш жараёнига фойдаланувчини чақаришдан иборат. Кўпинча, фойдаланувчи тизимга ЭРлар долзарблгини натижаларнинг дастлабки рўйхатида

хабар қилади. Қисқача кўринишда бу жараён куйидагича содир бўлади.

- Фойдаланувчи сўровни амалга оширади (қиска, оддий).

- Тизим топилган ЭРларнинг дастлабки рўйхатини қайтаради.

- Фойдаланувчи топилган ЭРларнинг баъзиларини долзарб ёки долзарбмас сифатида белгилайди.

- Тизим фойдаланувчи билан тескари алоқага асосланган ҳолда ахборот эҳтиёжининг яхшиланган кўринишини аниқлайди.

- Тизим экранга топилган ЭРларнинг аниқланган тўпламини чиқаради.

RF усули бир ёки бир нечта итерацияларни кўзда тутиши мумкин. Бу жараённинг асосида фойдаланувчи тўплам таркибини яхши билмаган ҳолда аниқ сўровни шакллантиришга қодир эмас, лекин ЭРларни баҳолаши мумкин деган ғоя ётади. Шунинг учун сўровни аниқлаштириш учун бундай итерацияларнинг бир нечтасини бажариш мақсадга мувофиқ. Бу сенарий доирасида RF усули фойдаланувчининг ахборот эҳтиёжи эволюциясига шароит яратиши мумкин. Айрим ЭРларни кўриб чиқиш фойдаланувчига у кидираётган маълумот ҳақида тасаввурни аниқлашга ёрдам беради.

RF усулига ёрқин мисол расмлар кидирилиши ҳисобланади. Расмлар излаши натижасини кўриб чиқиш осон, лекин айнан шу соҳада фойдаланувчига сўровни сўзлар билан шакллантириш қийин, лекин қайси расмлар долзарб ёки долзарбмас эканлигини кўрсатиш осон. Фойдаланувчи МИТ саҳифасида сўровни киритгандан сўнг, у натижаларнинг бирламчи рўйхатини олади. Фойдаланувчи долзарб деб ҳисоблайдиган расмни танлади. Улар сўровни аниқлаш учун ишлатилади, бу вақтда эса бошқа расмлар сўровнинг янги шаклига таъсир қилмайди. RF усули бўйича итерация бажарилганидан сўнг сараланган янги натижалар кўрсатилди.

2. Долзарблик бўйича тескари алоқа учун Роккио алгоритми. Роккио алгоритми[3] (Rocchio algorithm) – RF усулини амалга ошириш учун классик алгоритм бўлиб ҳисобланади. У долзарблик бўйича тескари алоқа моделини вектор тўплам моделига инкорпорациялайди.

Долзарб ЭРларга максимал яқин ва долзарбмас ЭРларга минимал ўхшаш q сўров векторини топишни хоҳлаймиз. Агар C_r – бу долзарб ЭРларнинг тўплами, C_{nr} эса – долзарбмаслар тўплами бўлса, унда излашининг занжири куйидаги вектор бўлади.

$$\vec{q}_{op} = \arg \max_q [sim(\vec{q}, C_r) - sim(\vec{q}, C_{nr})] \quad (1)$$

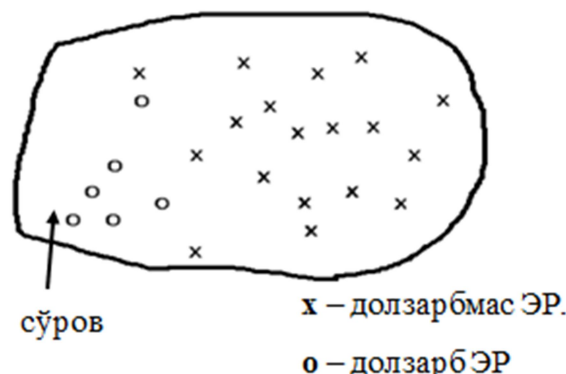
Бу ерда sim катталиги тенглик куйидагича аниқланади.

$$sim(q_j, q) = \frac{\vec{d}^{(j)} \cdot \vec{q}}{|\vec{d}^{(j)}| \cdot |\vec{q}|} = \frac{\sum_{i=1}^n w_i^{(j)} w_i^q}{\sqrt{\sum_{i=1}^n (w_i^{(j)})^2} \sqrt{\sum_{i=1}^n (w_i^q)^2}}$$

Агар яқинлик ўлчами сифатида ўхшашликнинг косинус ўлчамидан фойдаланилса, унда долзарб ва долзарбмас ЭРларни ажратиш учун \vec{q}_p оптимал вектор сўрови куйидаги тенглик билан аниқланади.

$$\vec{q}_p = \frac{1}{|C_r|} \sum_{d \in C_r} \vec{d} - \frac{1}{|C_{nr}|} \sum_{d_n \in C_{nr}} \vec{d}_n \quad (2)$$

Бошқача қилиб айтганда, оптимал сўров – бу долзарб ва долзарбмас ЭРлар центроидлари орасидаги фарк (1- расмга қаранг). Бироқ бу кузатув унчалик ҳам фойдали эмас, чунки долзарб ЭРларнинг тўлиқ тўплами аниқ эмас.

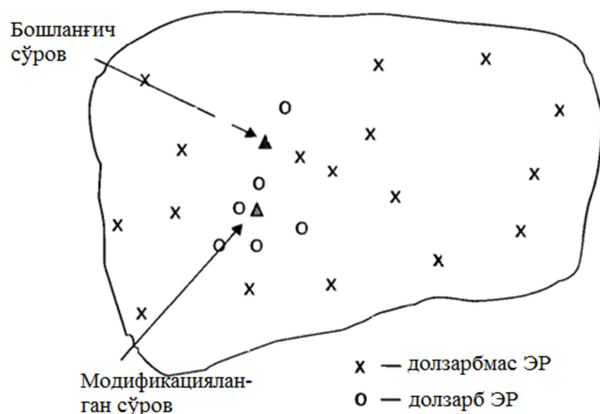


1- расм. Долзарб ва долзарбмас ЭРлар

Роккио алгоритми. Таърифланган долзарблик бўйича тескари алоқа механизми 1970 йилда Дж. Солтоннинг SMART тизимида амалга оширилган ва шу тизим эвазига машҳур бўлган[2]. Фойдаланувчи сўрови ва қисман ЭРларнинг долзарблиги ҳақида билимлар бўлсин. Алгоритм модификацияланган \vec{q}_n сўровдан фойдаланишни таклиф қиламиз.

$$\vec{q}_n = \alpha \vec{q}_0 + \beta \frac{1}{|D_{nr}|} \sum_{d_n \in D_{nr}} \vec{d}_n - \gamma \frac{1}{|D_r|} \sum_{d \in D_r} \vec{d} \quad 3$$

Бу ерда \vec{q}_0 – сўровнинг асл вектори, D_r ва D_{nr} – мос равишда маълум долзарб ва долзарбмас ЭРларнинг тўпламлари, α , β ва γ эса – ҳар бир қўшилувчининг вазни. Бу катталиклар баҳоланган ЭРлар ва дастлабки сўров тўпламларининг ҳиссасини бошқаради. Агар бизда баҳоланган ЭРлар кўп бўлса, унда биз β ва γ ни қийматини оширишимиз мумкин. \vec{q}_0 сўровдан бошлаб янги сўров долзарб ЭРлар центроиди йўналишида ва долзарбмас ЭРлар центроидидан маълум масофага кўчади. Бу янги сўровни вектор тўпламнинг стандарт моделида ЭРларни излаш учун ишлатиш мумкин.



2- расм. Тескари алоқанинг долзарблик бўйича қўлланилиш эффеќти

Долзарбмас ЭР векторини айирган ҳолда мусбат квадрат вектор тўпланидан кечишимиз мумкин. Роккио алгоритмида терминларнинг манфий вазнлари эътиборга олинмайди. Бошқача қилиб айтганда, бундай терминнинг вазни нолга тенг қилиб олинади. Тескари алоқанинг долзарблик бўйича қўлланилиш самараси 2-расмда кўрсатилган.

Долзарблик бўйича тескари алоқа ҳам тўлиқлик, ҳам аниқлик мезонини яхшилаши лозим[2]. Бироқ амалиётда тўлиқлик мушкул аҳволда бўлган ҳолатларда тўлиқликни ошириш учун энг фойдали ҳисобланиши мавжуд. Кўпинча бу Роккио усули – фойдаланиш контексти билан сўровни кенгайтириши (яъни, янги терминлар қўшилади) билан тушунтирилади. Фойдаланувчилар юқори тўлиқликка эришишни хоҳлаганларида, улар натижаларни кўриш ва излашни қайта бажаришга вақт сарфлашга тайёр бўлишади.

Мусбат тескари алоқа шунингдек, манфий тескари алоқага қараганда қимматли ҳисобланади, шунинг учун кўпгина МИТларида параметрларни $\gamma < \beta$ кўринишида ўрнатиш қабул қилинган. Куйидаги параметрларни танлаш мақсадга мувофиқ: $\alpha = 1, \beta = 0.75, \gamma = 0.15$. Кўп тизимлар, масалан расмларни излаш тизими, фақат ижобий тескари алоқага йўл қўяди, бу $\gamma = 0$ ўрнатмага эквивалент. Салбий тескари алоқа сифатида фақат юқори рангли долзарбмас ЭРлардан фойдаланиш мумкин (яъни (1) тенгликда $|D_{nr}| = 1$). Долзарблик бўйича тескари алоқанинг турли вариантларини таққослашнинг кўп сонли тажрибавий натижалари якуний хулоса қилиш имконини бермаслигига қарамадан, баъзи тадқиқотчилар, Ide desc-hi номини олган айнан шу вариант энг эффеќтли ёки ҳеч бўлмаганда қарама-қарши бўлмаган вариант деб ҳисоблашади.

3. Долзарблик бўйича эҳтимоли бўлган тескари алоқа. Вектор тўпланида сўров термини вазнини ўзгартиришдан ташқари долзарблик бўйича тескари алоқани ташкил қилишнинг бошқа ёндашувларини таклиф қиламиз. Масалан, агар фойдаланувчи бир нечта долзарб ва долзарбмас ЭРларни кўрсатган бўлса, унда синфловчини қуриш

мумкин. Кўпинча, синфловчининг асоси содда Байеса эҳтимол модели бўлиши мумкин. R – ЭРнинг долзарблигини акс эттирувчи мантқиқий идентификатор ўзгарувчиси. Шунда биз $P(x_i = 1)$ катталиқни, яъни ЭР долзарб ёки йўқлигига қараб ЭРда t термин учраш эҳтимолини баҳолашимиз мумкин.

$$\hat{P}(x_i = 1 | R = 1) = \frac{|VR_t|}{|VR|} \quad (4)$$

$$\hat{P}(x_i = 0 | R = 0) = \frac{(df_t - |VR_t|)}{(N - |VR|)}$$

Бу ерда N – ЭРларнинг умумий сони, $df_t - t$ терминга эга ЭРлар сони, VR – маълум долзарб ЭРлар тўплами ва $VR - t$ терминга эга, VR лардан ташкил топган ЭРлар тўплани остиси. Гарчи маълум долзарб ЭРлар тўплами барча долзарб ЭРларнинг тўплани остиси ҳисобланади, агар долзарб ЭРларнинг тўплами барча ЭРларнинг кичик тўпланини ўзидан намоён этади деб тахмин қилинса, унда юқорида келтирилган баҳолаш ҳақиқатга яқин. Бу баҳолаш асосида сўров терминлари вазнини ўзгартиришнинг янги усулини таклиф қилиш мумкин. Ҳозирча тенгликлар (4) термин вазнини ўлчаш учун етарли эмаслигига эътибор берамиз. Бу тенгликлар тўплани бўйича статистика ва долзарб ҳисобланган, ЭРлар таркибидаги терминларнинг тақсимланиши ҳақида ахборотдан фойдаланади, лекин маълум сўров ҳақидаги маълумотдан эмас.

Долзарблик бўйича тескари алоқа қачон самарали бўлади? – деган савол туғилади. RF усулининг муваффаќияти маълум тахминларга боғлиқ.

Биринчидан, фойдаланувчи қидирилатган ЭРларга қандайдир маънода яқин дастлабки сўровни шакллантириш учун етарлича билимларга эга бўлиши керак[2]. Бу шарт ҳар қандай МИТда зарур ҳисобланади, лекин RF усулининг ўзи йўқота олмайдиган бир нечта муаммоларни айтиб ўтиш лозим.

Нотўғри ёзув. Агар фойдаланувчи сўров терминини тўпланининг бирорта ҳам ЭРида ҳам учрамайдиган кўринишда ёзса, унда RF усули эффеќтли бўлмаслиги аниқ. Бу муаммони хатоларни тўғрилаш усуллари ёрдамида ечиш мумкин.

Кўп тилли МИТлари. Турли тиллардаги ЭРлар вектор тўпланида бир-биридан анча йироқда жойлашади, бир хил тилдаги ЭРлар эса бир-бирига яқин гуруҳланади.

Фойдаланувчи сўровси ва луѓатининг келиш-маслиги. Агар фойдаланувчи лаптоп терминини юборса, барча ЭРлар эса нотбук сўзига эга бўлса, унда сўров натижа бермайди ва долзарблик бўйича тескари алоқа эффеќтсиз бўлади.

Иккинчидан, долзарблик бўйича тескари алоқа усули долзарб ЭРлар бир-бирига ўхшаш, яъни

кластерлар ҳосил қилишини талаб қилади[4]. Идеал ҳолатда терминнинг барча долзарб ЭРлар бўйича тақсимланиши фойдаланувчи томонидан белгиланган барча ЭРларда термин тақсимланишига ўхшаш бўлиши керак, барча долзарбмас ЭРлар бўйича термин тақсимланиши эса долзарб ЭРлардаги термин тақсимланишидан фарқ қилиши керак. Усул яхши ишлайди, агар барча долзарб ЭРлар алоҳида прототип ёки турли прототиплар мавжуд бўлса атрофида кластер ҳосил қилади, долзарб ЭРларнинг луғатлари кучли ёпилади ҳамда долзарб ва долзарбмас ЭРлар ўртасидаги ўхшашлик кичик. Роккио модели долзарб ЭРларни центроид ёрдамида моделлаштирадиган алоҳида кластер сифатида интерпретациялайди. Бу ёндашув ишламайди, агар долзарб ЭРлар мултимодал синфни ташкил қилса, яъни вектор соҳасида бир нечта кластердан ташкил топса. Бу қуйидаги ҳолатларда содир бўлиши мумкин.

Тўпланишнинг сифатли тайёрланган ЭРлар таркиби кўпинча бу муаммони ечишда ёрдам бериши мумкин. Масалан, Фирмадаги ҳолатга турли гуруҳларнинг муносабати ҳақидаги мақола турли томонлар ишлатадиган терминологияга эга бўлиши мумкин: натижада ЭРлар кластери орасида алоқа ҳосил бўлади.

Долзарблик бўйича тескари алоқа фойдаланувчиларга доим ҳам ёқавермайди. Улар кўпинча аён тескари алоқани рад этишади ёки давом эттиришни умуман хоҳлашмайди. Бундан ташқари, кўпинча тескари алоқа асосида излаш натижалари бўйича маълум ЭР нима учун топилганлигини тушуниш қийин.

Бундан ташқари, долзарблик бўйича тескари алоқа бир нечта амалий муаммоларни келтириб чиқаради. *RF* усули қўлланилиши натижасида генерацияланадиган узун сўровлар оддий МИТларида эффектсиз бўлади. Бу катта ҳисоблаш сарфларига олиб келади ва фойдаланувчи сўровига жавоб қайтариш вақтини оширади. Бу муаммонинг қисман ечимини долзарб ЭРда энг муҳим терминлар вазнини оширган ҳолда эришиш мумкин, масалан, энг кўп учрайдиган биринчи йигирмата термин. Баъзи тажрибалар натижасига кўра терминларнинг чекланган сонидан фойдаланиш анча яхши натижалари бериши мумкин, гарчи бошқа оммада муаллиф терминларнинг катта сонидан фойдаланиш топилган ЭРлар сифатини ошириши ҳақида айтади.

4. Корпоратив тармоқда долзарблик бўйича тескари алоқа. Корпоратив тармоқ излашнинг баъзи тизимлари ўхшаш-боғлиқ ЭРларнинг излашига функционаллик беради: фойдаланувчи натижалар рўйхатида унинг фикрига кўра ахборот эҳтиёжига мос келувчи ЭРни кўрсатади ва унга ўхшаш бошқа ЭРларни сўрайди. Бундай функционалликни долзарблик бўйича тескари алоқанинг соддалаштирилган вариант сифатида кўриб чиқиш мумкин. Бироқ умуман олганда *RF* коди Корпоратив тармоқда камдан-кам ишлатилади. Олдин долзарблик бўйича тўлиқ тескари алоқа

берган (Excite Корпоратив тармоқ-излаш тизими бундан мустасно). Бироқ вақт ўтиши билан бу имкониятдан воз кечишди, чунки у фойдаланувчилар талабига мос келмади. Корпоратив тармоқда камдан-кам одамлар излашнинг кенгайтирилган имкониятидан фойдаланади, кўпчилик ягона сўров билан чекланишни маъқул кўради. Бу ҳолат икки сабаб билан тушунтирилиши мумкин: биринчидан, *RF* усулининг мақсадини оддий фойдаланувчига тушунтириш қийин, ва иккинчидан, бу усулнинг асосий мақсади қандага кўра, Корпоратив тармоқ-излаш тизими фойдаланувчиларини камдан-кам қизиқтирадиган излаш тўлиқлигини ошириш ҳисобланади.

2000 йилда Спик ва бошқалар Excite излаш тизимида *RF* усулидан фойдаланиш натижаларини оммалаштирди. Бу усул фақат тахминан 4% излаш сессияларида қўлланилган, бунда улардан кўпчилиги ҳар бир натижа билан борадиган "More like this" ("Ўхшаш ЭРлар") созланишидан фойдаланиш билан чекланган. Фойдаланувчиларнинг кўпчилиги натижаларнинг биринчи саҳифасини кўришган ва кейинги саҳифаларда ЭРларни излашмаган. *RF* усулидан фойдаланилганда натижалар тахминан учтадан икки ҳолатда яхшилланган.

Сўнги вақтларда анча муҳим йўналиш фойдаланувчиларнинг излаш натижалари бўйича босишининг таҳлили бўлди, бу долзарблик бўйича аён бўлмаган тескари алоқани таъминлайди. Бундай маълумотлардан фойдаланиш Ёахимс ишларида батафсил ўрганилган. Шунингдек, аён бўлмаган тескари алоқа кўриниши сифатида кўриб чиқиш мумкин бўлган, гарчи бу алоқа саҳифанинг ўқувчилари билан эмас, унинг муаллифлари томонидан ўрнатилсада (амалиётда кўпчилик муаллифлар шунингдек ўқувчилар ҳисобланади), Корпоратив тармоқда гиперҳавоалар тузилишидан фойдаланиш жуда муваффақиятли бўлди.

5. Тасодифий (сохта, алдамчи) долзарблик бўйича тескари алоқа. Сохта долзарблик бўйича тескари алоқа[3] (pseudo relevance feedback), ёки долзарблик бўйича кўп тескари алоқа[3] (blind relevance feedback), - бу автоматик локал таҳлил усули. У *RF* нинг қўлда бажариладиган қисмини автоматлаштириш имконини беради, шундай экан фойдаланувчи тизим билан кўшимча алоқага киришмасдан излаш сифатини оширади. Бу усул доирасида олдин излаш бажарилади ва энг катта рангларга эга бўлган биринчи *q* ЭРлар долзарб деб ҳисобланадиган, кейин эса уларга бу тахминни ҳисобга олган ҳолда *RF* усули қўлланиладиган энг долзарб ЭРларнинг дастлабки бирлашмаси топилади.

Долзарблик бўйича ошқормас тескари алоқа. Қайта алоқа учун база сифатида долзарбликнинг ошқор баҳоларидан кўра билвосита гувоҳликлардан фойдаланиш мумкин. Бу усул кўпинча долзарблик бўйича ошқормас тескари алоқа (implicit relevance

feedback) деб аталади. Ошкормас тескари алоқа ишончсизроқ, лекин фойдаланувчилар фикрини ҳисобга олмайдиган сохта долзарблик бўйича тескари алоқага қараганда анча фойдали. Шу билан бирга фойдаланувчилар кўпинча ошкор тескари алоқада иштирок этишдан бош тортишади, катта тизим, масалан Корпоратив тармоқ-излаш тизими ҳолатида ноаниқ тескари алоқа ҳақида катта ҳажмдаги маълумотларни йиғиш эса қийин эмас.

Контекстда ғоясини фойдаланувчилар кўриб чиқиш учун кўпроқ танлайдиган ЭРларга анча юқори ранг беришдан иборат бўлган DirectHit усули таклиф қилинган. Бошқача қилиб айтганда, саҳифаларга ҳаволалар бўйича босишлар бу саҳифаларнинг сўровларга долзарблик кўрсаткичи ҳисобланади деб тахмин қилинади. Бу ёндашувда бир нечта тахминлар қилинади, масалан излаш натижаларидаги ЭРлар снippetлари (уларнинг ёрдамида фойдаланувчилар қайси саҳифаларга ўтиш кераклигини танлашади) бу ЭРларнинг долзарблик индикаторлари ҳисобланади. Асл DirectHit излаш тизимида босишлар ҳақида маълумотлар глобал йиғилган ва аниқ фойдаланувчи ва сўровлар билан боғлиқ бўлмаган. Бу усул босишларни (clickstream mining) таҳлил қилишнинг умумий ёндашуви турларидан бири ҳисобланади. Ҳозирги вақтда шунга жуда ўхшаш усул Корпоратив тармоқдаги излаш сўровларига мос келувчи реклама эълонларини ранжирлаш учун фойдаланилади.

Резюме усули. Тадиқотлар кўрсатдики, *RF* усули натижалар долзарблигини жуда масарали ошириш имконини беради. Ундан муваффақиятли фойдаланиш учун етарлича кўп долзарб ЭРлар мавжуд бўлган сўровлар зарур[2]. Долзарблик бўйича тўлиқ тескари алоқа фойдаланувчи учун оғир ҳисобланади, унинг амалга оширилиши эса кўпгина МИТларида унчалик самарали эмас. Кўп ҳолатларда шунга ўхшаш яхшиланишга камроқ куч сарфлаган ҳолда интерактив излашнинг бошқа усуллари ёрдамида эришиш мумкин.

Ихтиёрий сўров бўйича асосий излаш сенарийсидан ташқари долзарблик бўйича тескари алоқа қуйидаги ҳолатларда ишлатилади:

- Ўзгарадиган ахборот эҳтиёжларини кузатиш (масалан, фойдаланувчи қизиқадиган автомобил маркаси вақт ўтиши билан ўзгаради).

- Ахборот филтрларини қўллаш (масалан, янгиликлар олиш учун).

- Фаол ўқиш (active learning) (ўқитиладиган танлов шаклланиши учун харажатларни қисқартириш учун қўлда синфлаш фойдали бўлган мисолларни топиш).

6. Долзарблик бўйича тескари алоқа стратегиясини баҳолаш. Долзарблик бўйича интерактив тескари алоқа излаш сифатида сезиларли ютуққа эришиш имконини беради. Эмперик нуқтаи назардан долзарблик бўйича ҳатто бир тескари алоқа цикли кўпинча жуда фойдали бўлади. Икки цикл баъзан кам фойдали бўлади. *RF* дан муваффақиятли фойдаланиш учун баҳоланган ЭРларнинг

етарлича катта сони талаб қилинади, бўлмаса жараён нотурғун бўлади ва фойдаланувчининг ахборот эҳтиёжидан бошқа томонга кетиши мумкин. Мос равишда, бештадан кам бўлмаган баҳоланган ЭРларга эга бўлиш тавсия этилади.

RF усулининг эффе́ктивлигини тўлиқ ва аён баҳолаш қийин. Маълумки, биринчи стратегия сифатида дастлабки q_0 сўровдан бошлаш ва “аниқлик-тўлиқлик” графигини қуриш мумкин. Фойдаланувчидан тескари алоқанинг бир циклидан сўнг модификацияланган q_m сўровни ҳисоблаймиз ва яна “аниқлик-тўлиқлик” графигини тузамиз. Иккала цикл давомида тизим ишининг сифати оддий таққослашга йўл қўядиган тўпламдаги барча ЭРлар бўйича баҳоланади. Бу ҳолатда таасурот қолдирадиган ютуққа эга бўлиш мумкин: MAP кўрсаткичининг тахминан 50% га ўсиши. Афсуски, бу ёлгон. Бу ютуқ кўпинча маълум долзарб ЭРлар (фойдаланувчи томонидан баҳоланган) энди анча юқори долзарбликка эга бўлиши билан тушунтирилади. Сифатни анча аниқ баҳолаш учун таққослашни фойдаланувчига маълум бўлмаган ЭРлар бўйича ўтказиш керак.

Иккинчи ғоя, иккинчи цикл давомида қолдиқ тўпламдаги ЭРлардан, яъни долзарб сифатида баҳоланган ЭРлардан ташқари барча ЭРлардан фойдаланишдан иборат. Бундай баҳо ҳақиқатга анча яқин бўлиб чиқади. Афсуски, ўлчанган эффе́ктивлик кўпинча дастлабки сўровдагидан кўра пастда бўлиб қолади. Бу самарали долзарб ЭРларнинг кичик микдори бўлганда ва уларнинг сезиларли ҳиссаси фойдаланувчи томонидан биринчи цикл давомида баҳоланганда айниқса, аниқ намён бўлади. Долзарблик бўйича тескари алоқа турли усулларининг нисбатан эффе́ктивлигини аниқ баҳолаш мумкин, лекин тизимлар сифатини долзарблик бўйича ва усиз тескари алоқа билан асосли таққослаш жуда қийин, чунки тескари алоқадан олдин ва кейин тўплам ўлчами ва долзарб ЭРлар сони фарқ қилади.

Шундай қилиб, бу усуллардан бирортаси тўлиқ қониқарли ҳисобланмайди. Усулда бири дастлабки сўров ва долзарбликни баҳолаш учун, иккинчиси эса – таққослаш баҳоси учун ишлатиладиган икки тўпламлар билан ишлашни кўзда тутати. q_0 ва q_m сўровларни қайта ишлаш сифатини иккинчи тўплам ёрдамида аниқ таққослаш мумкин.

Эҳтимол, *RF* усулининг фойдалилигини баҳолашнинг энг яхши усули фойдаланувчилар иштирокидаги тажрибалар ҳисобланса керак, масалан, вақтинча кўрсаткичларни таққослаш йўли билан: фойдаланувчи бошқа стратегия билан (масалан, сўровни янги шакллантириш билан) таққослаганда *RF* ёрдамида долзарб ЭРларни қанчалик тез топмоқда ёки вақтнинг маълум қисмида фойдаланувчи қанчалик кўп долзарб ЭРларни топмоқда. Фойдалиликнинг бундай баҳолари энг аниқ ва тизимдан фойдаланишга реал яқин ҳисобланади.

Хулоса

МИТларида долзарбликни тескари алоқа асосида ҳисоблаш бу ҳар бир ахборот тизимининг фойдаланувчилари ва уларнинг реал билимлари асосида аниқланади. Аммо бу долзарблик ҳамма вақт ҳам ўзини оқламайди. Бундай ҳолатларда МИТларининг фойдаланувчилари сони ва тизимдаги фойдали ЭРларни кўпайтириш лозим. Шунингдек, ахборот тизимини қўллаб қувватлаш, бойитишни ҳам амалга ошириш мумкин. FSV технологиясида бу ёндашувлар учун барча амалга ошириш, МИТни ўқитиш модкллари мавжуд. Буни кенгайтириш тизимга боғлиқдир. Бу ёндашувларни махсус инструмент (Tools for calc_r) ни созлаш, метод (event) ва хусусиятларни (properties) ўзгартириш мумкин.

Адабиётлар

1. Мўминов Б.Б. Маълумотларни излаш тизими. –Т.: Фан ва технология. 2016. –210 б.
2. Шокин Ю.И. Проблемы поиска информации / Ю.И. Шокин, А.М. Федотов, В.Б. Барахнин. Новосибирск: Наука, 2010. – 220 с.
3. Page L., Brin S., Motwani R., Winograd T.: The PageRank Citation Ranking: Bringing Order to the

Web. Stanford Digital Libraries Working Paper, Stanford University, 1998. -17 p. <http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf>

4. Anand Singh Kunwar, The PageRank Citation Ranking: Bringing Order to the Web Paper Review. <http://home.iitk.ac.in/~anandk/cs300/3B.pdf>

Мўминов Баходир Болтаевич

Техника фанлари доктори, Ахборот-коммуникацион технологиялари кафедраси, Ўзбекистон Республикаси Миллий гвардия харбий-техник институти.

E-mail: mbbahodir@gmail.com

Relevance feedback strategies in FSV technology

Abstract: This article presents the basic principles of the main feedback strategies in the search engines of FSV technology. It discusses the relevance and coincidence (false, deceptive) relevance of feedbacks, the Rocky algorithm for the relevance of feedback, current feedback, relevance in the corporate network, the relevance of feedbacks on forecasting, and the evaluation of the feedback strategy.

Key words: data search, search engine, FSV technology, relevance, feedback, query, Rocky algorithm, centroid.

УДК 532.5

Р.М. Мадрахимов, Д.С. Яхшибоев, Ш.Қ. Отаев К.Р. Мадрахимов.

МАТРИЦА АРГУМЕНТЛИ ГОЛОМОРФ ФУНКЦИЯЛАР УЧУН ТЕЙЛОР ҚАТОРИНИНГ АНАЛОГИ

Ушбу мақолада матрица аргументли голоморф функциялар учун Тейлор қаторининг аналогини ўрганилади. В статье изучается аналог ряда Тейлора для голоморфных функций матричного аргумента. In this article we study analogue of the Taylor series holomorphic functions of the matrix argument.

Ушбу мақолада матрица аргументли голоморф функциялар учун Тейлор қаторининг аналогини келтирилган.

1-таъриф: [1]. $A \in C[m \times m]$ бўлсин. У ҳолда A матрицанинг ўнг (Right) абсолют қиймати (чап (Left) абсолют қиймати) деб қуйидаги микдорга айтамыз $|A|_R = \sqrt{A^*A}$ ($|A|_L = \sqrt{AA^*}$).

Абсолют қийматнинг баъзи хоссаларини тақидлаб ўтамыз. Биз ўнг абсолют қиймат учун хоссаларни келтирамыз, булар чап абсолют қиймат учун ҳам ўринли бўлади.

1. $\forall \alpha \in C, |\alpha A|_R = \alpha |A|_R$
2. $|A|_R = \theta \Leftrightarrow A = \theta$ бу ерда θ нол матрица.
3. Умумий ҳолда. а) $|AB|_R \neq |A|_R |B|_R$ мисол

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \text{ ва } B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

- б) $|A + B|_R \neq |A|_R + |B|_R$ мисол

$$\sigma_3 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$|(\sigma_3 + 1) + (\sigma_1 - 1)| \neq |(\sigma_3 + 1)| + |(\sigma_1 - 1)|$$

(Нильсон мисоли [1] 241,бет)

4. Матрицанинг чап ва ўнг абсолют қийматлари учун қуйидаги тенглик ўринли $\sqrt{A^*A} = U\sqrt{AA^*}$ (бу ерда U унитар матрица), яъни $\sqrt{A^*A}$ ва $\sqrt{AA^*}$ лар унитар матрица аниқлигида тенг.

5. Агар A ва B операторлар диагональ матрицалар бўлса, у ҳолда

$|AB|_R = |A|_R |B|_R$, аммо $\|AB\|_S = \|A\|_S \|B\|_S$ бўлиши ҳар доим шарт эмас. Бу ерда $\|\cdot\|_S$ спектрал норма. Мисол $A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}, B = \begin{bmatrix} 7 & 0 \\ 0 & 0 \end{bmatrix}$

6. Агар норма қийматини диагональ матрица деб тасавур қилсак, у ҳолда буларни таққослаш мумкин. Ушбу $\begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}$ матрицани қараймиз. Унинг