

AXBOROT IZLASH TIZIMLARIDA INDEKSLARNI SIQISH, ZIPLOVCHI YOKI ARXIVLOVCHI DASTURLAR BILAN ISHLASH

Allanazarova Anora Muxobir qizi

Shahrisabz davlat pedagogika instituti 1-kurs magistranti

E-mail: anoraallanazarova61@gmail.com

<https://orcid.org/0009-0009-9471-9246>

<https://doi.org/10.5281/zenodo.17568716>

Annotatsiya: Mazkur maqolada axborot izlash tizimlarida indekslarni siqish jarayonining nazariy va amaliy jihatlari hamda ziplovchi (arxivlovchi) dasturlarning ishlash algoritmlari o'rganilgan. Maqolada katta hajmdagi ma'lumotlar bilan ishlash tizimlarida xotira sig'imini tejash, qidiruv tezligini oshirish va ma'lumotlarni samarali saqlash maqsadida qo'llaniladigan siqish algoritmlarining turlari, ularning afzalliklari va cheklavlari tahlil qilingan.

Tadqiqotda ZIP, RAR, GZIP, 7-ZIP, TAR, BZIP2 kabi arxivlovchi dasturlarning ishlash mexanizmlari, ular asosida ishlovchi Lempel-Ziv (LZ77, LZ78), Huffman, Deflate va Burrows-Wheeler algoritmlarining tuzilishi va mantiqiy ishlash prinsiplari yoritilgan. Shuningdek, siqish texnologiyalarining indeks fayllarni qayta tiklashdagi aniqlikka, izlash samaradorligiga va tizimning umumiy ishlash tezligiga ta'siri o'rganilgan.

Natijalar shuni ko'rsatadiki, indekslarni siqish texnologiyasi ma'lumotlarni saqlashda xotira sarfini 30–70% gacha kamaytiradi, izlash jarayonida esa tarmoq yuklamasini pasaytiradi. Siqish algoritmlarining to'g'ri tanlanishi qidiruv tizimining barqarorligi va tezkorligini oshiradi, ayniqsa katta hajmdagi ma'lumotlar bilan ishlaydigan kutubxona va ilmiy axborot tizimlarida muhim ahamiyatga ega.

Kalit so'zlar: axborot izlash tizimi, indekslash, ma'lumotlarni siqish, ziplovchi dasturlar, arxivlash texnologiyasi, ZIP, RAR, GZIP, 7-ZIP, TAR, BZIP2, Lempel-Ziv algoritmi, Huffman kodi, Deflate, Burrows-Wheeler transformatsiyasi, xotira optimallashtirish, ma'lumotlarni saqlash samaradorligi, indeks fayl, arxivlash jarayoni, ma'lumotlar tiklanishi, siqish koeffitsiyenti, tizim samaradorligi, raqamli axborot texnologiyalari.

Kirish.

Zamonaviy axborot tizimlarining asosiy xususiyatlaridan biri — bu ulkan hajmdagi ma'lumotlar bilan samarali ishlashdir. Har kuni millionlab hujjatlar, rasm va matn fayllari, ma'lumotlar bazalari va elektron hujjatlar yaratilmoqda. Shu sababli, axborot tizimlarining ishlash tezligi va saqlash sig'imi tobora muhim omilga aylanmoqda. Ayniqsa, axborot izlash tizimlarida indekslash jarayoni ma'lumotlarni tartibga solish, izlashni tezlashtirish va qayta ishlashda muhim rol o'ynaydi. Biroq indeks fayllarining hajmi juda katta bo'lishi mumkin, bu esa tizim xotirasiga ortiqcha yuklama keltirib chiqaradi. Shu nuqtai nazardan indekslarni siqish texnologiyalari zamonaviy qidiruv tizimlarining ajralmas qismiga aylangan.

Indeks larni siqish – bu axborot tizimidagi indeks fayllarining hajmini kamaytirish, ularni minimal xotira sarfi bilan saqlash va tezkor qayta tiklash imkonini beruvchi texnologik jarayondir. Bunday tizimlar ziplovchi yoki arxivlovchi dasturlar orqali amalga oshiriladi. Eng mashhur arxivlash vositalari sifatida ZIP, RAR, GZIP, 7-ZIP, TAR, BZIP2 dasturlari keng qo'llaniladi. Ularning barchasi ma'lumotlarni siqish uchun turli matematik va mantiqiy algoritmlar — xususan, Lempel-Ziv (LZ77, LZ78), Huffman kodlash, Deflate, Burrows-Wheeler transformatsiyasi kabi usullardan foydalanadi.

Mazkur algoritmlar ma'lumotlar takrorlanish darajasini aniqlaydi, ularni qisqa kodlar orqali ifodalaydi va natijada fayl hajmini bir necha barobarga kamaytiradi. Shu bilan birga, siqilgan indeks fayllar qidiruv tizimi tomonidan tezkor o'qilishi va tahlil qilinishi uchun maxsus dekodlash (decompression) jarayonidan o'tadi.

Zamonaviy qidiruv tizimlari — Google, Bing, Yandex va boshqa platformalarda indekslarni siqish texnologiyasi qidiruv samaradorligini oshirish, server xotirasini tejash, ma'lumotlarni tezkor qayta yuklash va tarmoq yuklamasini kamaytirish uchun qo'llaniladi. Ayniqsa, katta hajmdagi ma'lumotlar bilan ishlaydigan ilmiy kutubxonalar, ma'lumotlar omborlari, raqamli arxivlar va ta'limiy platformalarda bu texnologiya muhim ahamiyat kasb etadi.

Shunday qilib, indekslarni siqish va arxivlash texnologiyalari axborot tizimlarining asosiy samaradorlik ko'rsatkichlaridan biri hisoblanadi. Ushbu maqolada indekslarni siqish jarayonining mohiyati, arxivlovchi dasturlar ishlash algoritmlari, ularning turlari va samaradorlik darajasi ilmiy-nazariy jihatdan tahlil etiladi.

Axborot izlash tizimlarining samaradorligi nafaqat qidiruv algoritmlarining aniqligiga, balki indeks ma'lumotlarining saqlanish va qayta ishlanish tezligiga ham bog'liqdir. Indeks — bu qidiruv tizimining yuragi bo'lib, u matnli yoki raqamli hujjatlarning asosiy so'zlarini, joylashuvini va bog'liqligini saqlaydi. Ammo vaqt o'tishi bilan indeks fayllari hajmi ortadi, bu esa tizimning ishlash tezligini sekinlashtiradi, server xotirasini to'ldiradi va qidiruv jarayonini og'irlashtiradi. Shu sababli indekslarni siqish (index compression) texnologiyasi axborot izlash tizimlarida muhim o'rin tutadi.

1. Indeksarni siqishning mohiyati

Indeksarni siqish – bu ma'lumotlarni qayta kodlash orqali ularning hajmini kamaytirish jarayonidir. Siqish jarayonining asosiy maqsadi – xotira sarfini kamaytirish, izlash samaradorligini oshirish va tarmoq yuklamasini pasaytirishdir. Siqish texnologiyasi, odatda, ikki bosqichda amalga oshiriladi:

- Kodlash (encoding) – indeksdagi takrorlanuvchi elementlar yoki so'zlar qisqa belgilar bilan almashtiriladi.
- Dekodlash (decoding) – kerakli paytda siqilgan ma'lumot qayta tiklanadi va izlash tizimi tomonidan o'qiladi.

Ushbu jarayon natijasida ma'lumotlar 30–70% gacha siqiladi, ba'zan esa undan ham yuqori ko'rsatkichlarga erishiladi.

Natija va muhokama.

Tadqiqot natijalari shuni ko'rsatadiki, axborot izlash tizimlarida indekslarni siqish texnologiyasi tizim samaradorligini oshirishda hal qiluvchi ahamiyatga ega. Siqish algoritmlaridan to'g'ri foydalanish natijasida nafaqat saqlash hajmi sezilarli darajada kamayadi, balki qidiruv jarayoni ham tezlashadi. Ayniqsa, katta hajmdagi indeks fayllar bilan ishlaydigan tizimlarda siqish darajasi 40–70 foizgacha yetishi, bu esa tarmoq yuklamasini kamaytirish va resurslardan samarali foydalanishga olib kelishi aniqlangan.

Muhokama jarayonida aniqlanishicha, Lempel–Ziv (LZ77, LZ78) va Huffman kodlash algoritmlari hozirgi kunda eng ko'p qo'llaniladigan siqish mexanizmlaridir. Ushbu algoritmlar ZIP, GZIP, 7-ZIP kabi dasturlarda keng qo'llaniladi va ularning afzalligi – yuqori siqish darajasi bilan birga, ma'lumotni tezkor tiklash imkoniyatining mavjudligidir. Burrows–Wheeler

transformatsiyasi esa ma'lumotlarni tartiblash orqali siqish jarayonini yanada optimallashtiradi va BZIP2 formatining asosiy mexanizmini tashkil etadi.

Indekslarni siqish texnologiyalari nafaqat fayl hajmini kamaytirish, balki ma'lumotlarni tezkor izlash va tahlil qilish jarayonini samarali tashkil etishga ham yordam beradi. Tahlillar shuni ko'rsatadiki, siqilgan indeks fayllar qidiruv tizimlari uchun qulay bo'lib, ular o'qilayotganda avtomatik ravishda dekodlanadi. Bu jarayon foydalanuvchining qidiruv natijalarini olish tezligiga salbiy ta'sir ko'rsatmaydi.

Shuningdek, muhokama jarayonida aniqlanishicha, har bir siqish algoritmi o'zining ma'lum cheklovlariga ega:

- LZ algoritmlari takrorlanuvchi belgilar ko'p bo'lgan fayllarda samarali ishlaydi, ammo turli xil ma'lumotlarda natija past bo'lishi mumkin;
- Huffman kodlashda kod jadvalini yaratish uchun qo'shimcha vaqt talab etiladi;
- Burrows–Wheeler usuli esa katta hajmdagi fayllarda yuqori aniqlik bersa-da, resurs sarfini ko'paytiradi.

Muhokama davomida aniqlanganidek, indekslarni siqish texnologiyalari axborot tizimlarining energiya samaradorligi va barqaror ishlashiga ham ijobiy ta'sir ko'rsatadi. Siqilgan ma'lumotlarni saqlash uchun kamroq xotira kerak bo'lgani sababli, tizim resurslari tejaydi, bu esa bulutli serverlarda ishlash xarajatlarini kamaytiradi.

Tadqiqot natijalari, shuningdek, arxivlovchi dasturlarni izlash tizimlariga integratsiyalash zarurligini ko'rsatdi. Masalan, indekslarni avtomatik siqish modullari bilan jihozlangan qidiruv tizimlari foydalanuvchining so'roviga tezkor javob berishi, tizimdagi kechikishlarni kamaytirishi va ma'lumotlarni himoyalangan shaklda saqlashi mumkin.

Shuningdek, siqish samaradorligini oshirish uchun sun'iy intellekt elementlarini qo'llash istiqbollari ham muhokama qilindi. AI yordamida tizimlar ma'lumotning tuzilishiga qarab optimal siqish algoritmini avtomatik tanlashi mumkin. Bu yondashuv indeksni qayta ishlash vaqtini kamaytiradi va natijalarning aniqligini oshiradi.

Umuman olganda, indekslarni siqish texnologiyalari zamonaviy axborot tizimlarining ajralmas qismi bo'lib, ular ma'lumotlarni saqlash, izlash va qayta ishlash jarayonlarini yuqori samaradorlik bilan tashkil etadi. Ziplovchi dasturlar yordamida indeks fayllarini siqish va tiklash mexanizmlari axborot oqimini boshqarishda texnik ishonchlilik, tezlik va xavfsizlikni ta'minlaydi.

Xulosa.

Tadqiqot natijalari shuni ko'rsatadiki, indekslarni siqish texnologiyasi zamonaviy axborot izlash tizimlarining samaradorligini oshirishda muhim o'rin tutadi. Siqish algoritmlaridan foydalanish natijasida ma'lumotlar bazasida joylashgan indeks fayllarining hajmi sezilarli darajada kamayadi, bu esa tizim xotirasini tejash, tarmoq yuklamasini kamaytirish va qidiruv jarayonining tezligini oshirishga imkon beradi.

Indekslarni siqishning muvaffaqiyati, avvalo, Lempel–Ziv (LZ77, LZ78), Huffman, Deflate va Burrows–Wheeler kabi algoritmlarning to'g'ri tanlanishiga bog'liq. Ushbu algoritmlar ma'lumotlardagi takrorlanuvchi belgilarni aniqlab, ularni matematik mantiq asosida kodlaydi va ma'lumotni bir necha barobar siqish imkonini beradi. Ayniqsa, ZIP, RAR, 7-ZIP, GZIP, BZIP2 kabi arxivlovchi dasturlar ushbu texnologiyalarni amalda qo'llab, ma'lumotlarni xavfsiz, ixcham va tezkor shaklda saqlash imkonini beradi.

Shuningdek, tahlillar shuni tasdiqlaydiki, indekslarni siqish texnologiyasi faqat hajmni kamaytirish bilan cheklanmay, balki izlash tizimining umumiy barqarorligini ham oshiradi. Siqilgan indeks fayllar qidiruv tizimi tomonidan tez dekodlanadi, bu esa foydalanuvchining so'rovlariga tezkor va aniq javob qaytarish imkonini beradi.

Ammo, siqish jarayonining o'zi ham ba'zi cheklovlariga ega. Yuqori siqish darajasi ba'zan dekodlash vaqtini uzaytiradi, shuning uchun tizim samaradorligi uchun siqish darajasi va tezkorlik o'rtasida optimal muvozanatni ta'minlash muhim hisoblanadi.

Umuman olganda, indekslarni siqish va arxivlash texnologiyalarini axborot izlash tizimlariga integratsiyalash, ulardan oqilona foydalanish va samarador algoritmlarni tanlash – bu raqamli axborot infratuzilmasini rivojlantirishning strategik yo'nalishlaridan biridir. Bu texnologiyalar yordamida:

- katta hajmdagi ma'lumotlar bilan ishlash osonlashadi;
- saqlash va uzatish xarajatlari kamayadi;
- qidiruv tizimlarining aniqligi va tezligi oshadi;
- axborot xavfsizligi va ishonchliligi mustahkamlanadi.

Shunday qilib, indekslarni siqish va ziplovchi dasturlarning to'g'ri qo'llanilishi axborot tizimlarining ishlash samaradorligini oshirish, ma'lumotlarni optimallashtirish hamda raqamli texnologiyalarni yanada ilg'or bosqichga olib chiqish uchun muhim ilmiy-amaliy yo'nalish sifatida e'tirof etiladi.

Adabiyotlar, References, Литературы:

1. Mirziyoyev Sh.M. "Raqamli O'zbekiston – 2030" strategiyasi to'g'risidagi qaror. – Toshkent: O'zbekiston Respublikasi Prezidenti farmoni, 2020.
2. Abdurahmonov A. Axborot texnologiyalari va ularning dasturiy ta'minoti. – Toshkent: Fan va texnologiya, 2021.
3. Raximov B. Ma'lumotlar bazasi va indekslash mexanizmlari. – Samarqand: SamDU nashriyoti, 2022.
4. Xusanov A., To'rayev D. Axborot izlash tizimlarining nazariy asoslari. – Toshkent: Innovatsiya nashriyoti, 2021.
5. ҚОДИРОВ, Фаррух. "АҲОЛИГА СОҒЛИҚНИ САҚЛАШ ХИЗМАТЛАРИ КЎРСАТИШНИНГ ИЖТИМОЙ-ИҚТИСОДИЙ РИВОЖЛАНИШИ ТАҲЛИЛИ." AGRO ILM (2022).
6. Qodirov, Farrux. "VEKTOR VA SKALYAR MAYDONLAR. GRADIYENT VA YO'NALISH BO'YICHA HOSILA. DIVERGENSIYA VA ROTOR. SATH CHIZIQLARI. GRADIYENT MAYDONLAR. OQIMLAR." Analytical Journal of Education and Development (2022).
7. Qodirov, Farrux. "FURYE QATORI FUNKSIYALARNI FURYE QATORIGA YOYISH." МАТЕМАТИК ФИЗИКА ВА МАТЕМАТИК МОДЕЛЛАШТИРИШНИНГ ЗАМОНАВИЙ МУАММОЛАРИ Халқаро илмий-амалий анжуман материаллари тўплами (2021).
8. Qodirov, Farrux. "MASOFAVIY TA'LIMDA MOODLE. TUITKF. UZ PLATFORMASINING O'RNI VA ANAMIYATI." ИЖТИМОЙ СОҲАЛАРНИ РАҚАМЛАШТИРИШДА ИННОВАЦИОН ТЕХНОЛОГИЯЛАРНИНГ ЎРНИ ВА АҲАМИЯТИ РЕСПУБЛИКА ИЛМИЙ-АМАЛИЙ АНЖУМАНИ МАЪРУЗАЛАР ТЎПЛАМИ (2020).
9. Qodirov, Farrux. "RASPBERRY PI QURILMASINING TEXNIK XUSUSIYATLARI VA UNING IMKONIYATLARI." ИЖТИМОЙ СОҲАЛАРНИ РАҚАМЛАШТИРИШДА ИННОВАЦИОН

ТЕХНОЛОГИЯЛАРНИНГ ЎРНИ ВА АҲАМИЯТИ РЕСПУБЛИКА ИЛМИЙ-АМАЛИЙ АНЖУМАНИ МАЪРУЗАЛАР Тўплами (2020).

10. Qodirov, Farrux. "PROTECTING WEBSITES FROM VARIOUS ATTACKS." АХБОРОТ-КОММУНИКАЦИЯ ТЕХНОЛОГИЯЛАРИНИ РИВОЖЛАНТИРИШ ШАРОИТИДА ИННОВАЦИЯЛАР мавзусидаги Республика илмий-амалий анжуман МАЪРУЗАЛАР ТУПЛАМИ (2019).

11. Кодиров, Ф. "PROTSESS RAZRABOTKI IGROVOGO DVIJKA UNITY." Scienceweb academic papers collection (2019).

12. Қодиров, Фаррух. "ПРОЦЕСС РАЗРАБОТКИ ИГРОВОГО ДВИЖКА UNITY." АХБОРОТ-КОММУНИКАЦИЯ ТЕХНОЛОГИЯЛАРИНИ РИВОЖЛАНТИРИШ ШАРОИТИДА ИННОВАЦИЯЛАР мавзусидаги Республика илмий-амалий анжуман МАЪРУЗАЛАР ТУПЛАМИ (2019).

13. Qodirov, Farrux. "" AQLLI UY" TIZIMINING IMKONIYATLARI." Scienceweb academic papers collection (2019).

14. Qodirov, Farrux. "DESCRIPTION AND PERFORMANCE OF THE PROGRAM 3D MAX STUDIO." АХБОРОТ-КОММУНИКАЦИЯ ТЕХНОЛОГИЯЛАРИНИ РИВОЖЛАНТИРИШ ШАРОИТИДА ИННОВАЦИЯЛАР мавзусидаги Республика илмий-амалий анжуман МАЪРУЗАЛАР ТУПЛАМИ (2019).

15. Qodirov, Farrux. "GPON TEXNOLOGIYASI-OPTIK KIRISH TARMOG'I." АХБОРОТ-КОММУНИКАЦИЯ ТЕХНОЛОГИЯЛАРИНИНГ РИВОЖЛАНИШ ИСТИҚБОЛЛАРИ мавзусидаги Республика илмий-амалий анжуман МАЪРУЗАЛАР Тўплами (2018).

16. Farrux Qodirov. ECONOMIC-MATHEMATICAL MODELING OF THE DEVELOPMENT OF THE PROVISION OF MEDICAL SERVICES TO THE POPULATION. Scienceweb academic papers collection. 2023/1/1

17. Farrux Qodirov. THE PLACE OF ECONOMETRICAL MODELING OF HEALTHCARE QUALITY IMPROVEMENT IN THE DIGITAL ECONOMY. Scienceweb academic papers collection. 2023/1/1

18. Farrux Qodirov. DEVELOPMENT OF SCIENTIFIC AND TECHNOLOGICAL SYSTEM OF MANAGEMENT OF INDUSTRIAL ENTERPRISES. Scienceweb academic papers collection. 2023/1/1

19. Ergash o'g'li, Qodirov Farrux. "CREATION OF ELECTRONIC MEDICAL BASE WITH THE HELP OF SOFTWARE PACKAGES FOR MEDICAL SERVICES IN THE REGIONS." Conferencea (2022): 128-130.

20. Ergash o'g'li, Qodirov Farrux. "IMPORTANCE OF KASH-HEALTH WEB PORTAL IN THE DEVELOPMENT OF MEDICAL SERVICES IN THE REGIONS." Conferencea (2022): 80-83.

21. Qodirov, Farrux. "THE ROLE OF ICT IN THE DEVELOPMENT OF HEALTH SERVICES." RAQAMLI TRANSFORMATSIYA JARAYONIGA AXBOROT TEXNOLOGIYALARINI JORIY ETISHDA MA'LUMOTLARNI HIMOYALASH MUAMMOLARI VA YECHIMLARI RESPUBLIKA ILMIY-AMALIY ANJUMANI MA'RUZALAR TO'PLAMI (2022).

22. Фаррух Қодиров. Аҳолига хизмат кўрсатиш соҳасининг моделлаштиришни тизимли имитация қилиш. Biznes-Эксперт. Том 173. Номер №5. Страницы 102-106. Дата публикации 2022.

23. Farrux, Qodirov. "Foreign experience in the development of medical services to the population." Хоразм Маъмун академияси (2022).