

## AUDIO VA VIDEO AXBOROTLARNI SIQISH ALGORITMLARI VA TUSHUNCHALARI

**O‘rinboyev Alisher Shavkatjon o‘g‘li**

**(Muhammad Al-Xorazmiy nomidagi TATU Farg‘ona filiali talabasi)**

<https://www.doi.org/10.5281/zenodo.8304207>

**Annotatsiya:** Ushbu maqola axborotlarni siqish algoritmlari, axborotlarni siqish jarayonlari, siqilgan ma‘lumotlarni qaytarish va qaytarib bo‘lmasligi hamda qattiq diskdagi zaxira ma‘lumotlarni ustida amallar bajarish haqida tushunchalar berib o‘tilgan.

**Tayanch iboralar:** Axborotni siqish, siqish tezligi, siqish sifati, JPEG, JFIF, MPEG, ASCII.

Axborotni siqish - bu kompyuter texnologiyalarining rivojlanish tarixidan ancha uzoq tarixga ega bo‘lgan muammodir, odatda (tarix) axborotni kodlash va shifrlash muammolari rivojlanish tarixi bilan birga keladi. Barcha siqish algoritmlari kirishning oqimi bilan ishlaydi, ularning minimal birligi bir oz, maksimal esa bir necha bit, bayt yoki bir necha bayt. Siqish jarayonining maqsadi, qoida tariqasida, ba‘zi bir transformatsiyalar yordamida ba‘zi boshlang‘ich kompakt bo‘lmagan kirish oqimlaridan ma‘lumot birliklarining yanada ixcham chiqish oqimini olishdir. Siqish jarayonlarining asosiy texnik xususiyatlari va ularning ishlash natijalari quyidagilardan iborat.

Siqish tezligi - undan oqimning ekvivalent chiqish oqimini olish uchun kirish oqimining ma‘lum miqdorini siqish uchun sarflangan vaqt.

Siqish sifati - bir xil yoki boshqa algoritmdan foydalanib, unga qayta siqishni qo‘llash orqali chiqish oqimi qancha miqdorda to‘ldirilishini ko‘rsatadigan qiymat.

Axborotni siqish muammosiga turli xil yondashuvlar mavjud. Ba‘zilarida juda murakkab nazariy matematik baza mavjud, boshqalari axborot oqimining xususiyatlariga asoslangan va algoritmik darajada etarlicha sodda. Ma‘lumotni siqishni yoki siqishni amalga oshiradigan har qanday usul va algoritmi, qaytariladigan yoki qaytarib bo‘lmaydigan o‘zgartirilishi orqali bitlarda axborot chiqishi oqimini kamaytirishga mo‘ljallangan. Shuning uchun, birinchi navbatda, ma‘lumotlarning tabiati yoki formati bilan bog‘liq mezon bo‘yicha barcha siqishni usullarini ikki toifaga bo‘lish mumkin: qaytariladigan va qaytarib bo‘lmaydigan siqish.

Qaytarib bo‘lmaydigan siqish deganda kirish ma‘lumotlari oqimining o‘zgarishi tushuniladi, unda ma‘lum bir ma‘lumot formatiga asoslangan chiqish oqimi, ba‘zi [nuqtai nazardan](#), tashqi xarakteristikalarda kirish oqimiga mutlaqo o‘xshash bo‘lgan, ammo hajmidan farq qiladigan ob'ektning anglatadi. Kirish va chiqish oqimlarining o‘xshashligi darajasi ushbu ma‘lumot oqimi bilan ifodalangan ob'ektning ba‘zi xususiyatlarining (ya‘ni, ma‘lum bir ma‘lumot formatiga muvofiq siqilgan va siqilmagan ma‘lumotlar) muvofiqligi darajasi bilan belgilanadi. Bunday yondashuvlar va algoritmlar, masalan, oqimdagi baytlarning past darajadagi takrorlanish darajasi past bo‘lgan grafik fayllarning ma‘lumotlarini siqish uchun ishlatiladi. Ushbu yondashuv grafik faylning tuzilish formatining xususiyatidan va displey sifatida (inson ko‘zi bilan ko‘rish uchun) bir xil (yoki aniqroq n) usulda grafik rasmni taqdim etish qobiliyatidan foydalanadi. Shuning uchun, siqilish darajasi yoki kattaligiga qo‘shimcha ravishda, bunday algoritmlarda sifat tushunchasi paydo bo‘ladi, chunki siqish jarayonida asl rasm o‘zgaradi, keyin sifatni ma‘lumot formatiga qarab, sub'ektiv ravishda baholangan asl va natijada olingan rasmning muvofiqlik darajasi deb tushunish mumkin. Grafik fayllar uchun bunday yozishmalar vizual ravishda aniqlanadi, ammo mos keladigan aqlli algoritmlar va dasturlar

mavjud. Qaytish mumkin bo'lgan siqishni, kirish va chiqish oqimlarining axborot tuzilishining aniq muvofiqligi zarur bo'lgan joylarda qo'llash mumkin emas. Ushbu yondashuv JPEG va JFIF algoritmlari va JPG va JIF fayl formatlari deb nomlanuvchi video va foto ma'lumotlarini taqdim etishning mashhur formatlarida amalga oshiriladi.

Qayta tiklanadigan algoritmlarda jarayon sifatida kodlashni statistik nuqtai nazardan ko'rib chiqish mumkin, bu nafaqat siqishni algoritmlarini tuzishda, balki ularning samaradorligini baholashda ham foydalidir. Qaytariladigan barcha algoritmlar uchun kodlash qiymati tushunchasi mavjud. Kodlash narxiga bitdagi kod so'zining o'rtacha uzunligi kiradi. Kodlashning qisqarishi xarajat va kodlash entropiyasi o'rtasidagi farqga teng va yaxshi siqish algoritmi har doim ortiqcha ishlarni kamaytirishi kerak (esda tutingki, ma'lumotlarning entropiyasi uning buzilishini o'lchaydi.) Shannonning ma'lumotni kodlash bo'yicha asosiy teoremasi "kodlash har doim manba atrof-muhitidan kam emas, garchi unga o'zboshimchalik bilan yaqinlashishi mumkin". Shuning uchun har qanday algoritm uchun har doim kirish oqimining entropiyasi tomonidan aniqlanadigan siqilish darajasida har qanday cheklov mavjud. Endi biz to'g'ridan-to'g'ri qaytariladigan algoritmlarning algoritmik xususiyatlariga murojaat qilamiz va kodlash tizimlari va axborotni siqish usullarini amalga oshirish bilan bog'liq ma'lumotlarni siqishning eng muhim nazariy yondashuvlarini ko'rib chiqamiz. Axborotni siqish algoritmlarini ishlab chiqish amaliy matematikaning bir sohasiga tegishli. Ular tabiiy zaxirani yo'q qilish tamoyiliga asoslanadi. Axborotni siqish usullari an'anaviy ravishda ikkita ajratish sinfiga bo'linadi: yo'qolgan siqishva ma'lumotni yo'qotmasdan siqishni.

Yo'qotilgan siqish-siqilgan arxivni bo'shatgandan so'ng, avvalgisidan biroz farq qiladigan ma'lumotlar olinadi. Shubhasiz, siqilish darajasi qanchalik katta bo'lsa, yo'qotish hajmi va aksincha. Albatta, bunday algoritmlar matnli hujjatlar, ma'lumotlar bazasi jadvallari va dasturlarida qo'llanilmaydi. Oddiy, formatlanmagan matnda siz ozgina buzilishlardan omon qolishingiz mumkin, ammo dasturda kamida bitta bitni buzish uni to'liq ishlamay qolishiga olib keladi.

Shu bilan birga, o'nlab marta siqishni olish uchun ma'lumotlarning bir necha foizini qurbon qilish mumkin bo'lgan ma'lumotlar mavjud, masalan, fotosuratlar, video va audio materiallar. Bunday ma'lumotlarning siqilishi va keyinchalik dekompressiyasi paytida yo'qolishi qo'shimcha "shovqin" paydo bo'lishi sifatida qabul qilinadi.

Axborot yo'qotishni siqish algoritmlari quyidagilar kabi algoritmlarni o'z ichiga oladi Jpeg(rasmlarni siqishda foydalaniladi) va MPEG(video va audioni siqishda foydalaniladi). Axborot yo'qotishlarni siqish algoritmlari faqat iste'molchilar uchun mo'ljallangan vazifalar uchun ishlatiladi.

Siqish paytida ruxsat etilgan yo'qotishning kattaligi odatda boshqarish mumkun, bu sizga "o'lcham / sifat" ning maqbul nisbatiga erishishga imkon beradi. Ekranida namoyish etilishi mo'ljallangan fotografiyalarda ma'lumotning 5% yo'qolishi odatda tanqidiy emas, ba'zi hollarda esa 20-25% yo'qolishi mumkin.

Usullari ma'lumotni yo'qotmasdan siqishni ular matnli hujjatlar va dasturlar bilan ishlashda qo'llaniladi va ma'lumotlarning yo'qolishiga yo'l qo'ymaydi. Ular faqat uning ortiqcha bo'lishini yo'q qilishga asoslangan.

1-misol Ukraina tilida 32 ta harf, o'nta raqam va o'nlab o'nlab tinish belgilari va boshqa maxsus belgilar mavjud. Faqat katta harflar bilan yozilgan matn uchun (telegramdagi kabi)

oltmish xil ma'no etarli bo'ladi. Biroq, har bir belgi odatda 8 bitdan iborat va 256 xil kodlarni ifoda etadigan baytda kodlangan. Bu zaxira uchun birinchi asosdir. "Telegraf" matni uchun har bir belgi uchun 6 bit etarli bo'ladi.

Shakl 1. Morse kodi

2-misol Xalqaro ASCII belgilar kodlashda har qanday belgi kodlash uchun bir xil miqdordagi bitlar ajratilgan (8). Biroq, eng keng tarqalgan belgilar kamroq belgilarni kodlash uchun mantiqiy ekanligi aniq. Shunday qilib, masalan, in morse kodi ko'pincha topilgan "E" va "T" harflari bitta belgi bilan kodlangan (mos ravishda bu nuqta va tire). Va "Yu" (-) va "Ts" (- -) kabi noyob harflar to'rtta belgi bilan kodlangan. Samarasiz kodlash - bu ortiqcha to'lovning ikkinchi sababi.

Axborotni siqishni amalga oshiradigan dasturlar o'zlarining kodlashlarini (turli xil fayllar uchun har xil) kiritishlari va siqilgan faylga ma'lum bir jadval (lug'at) belgilashlari mumkin, bunda paketdan chiqarish dasturi ushbu faylda qanday belgilar va ularning guruhlari kodlanganligini bilib oladi. Axborotni transkodlashga asoslangan algoritmlar deyiladi xuffman algoritmlari.

Ikki nusxadagi bo'laklarning mavjudligi ortiqcha sonning uchinchi sababi hisoblanadi. Bu matnlarda kam uchraydi, lekin jadval va grafiklarda kodlarning takrorlanishi odatiy holdir. Masalan, agar 0 soni ketma-ket yigirma marta takrorlangan bo'lsa, unda yigirma nol baytni qo'yish mantiqiy emas. Buning o'rniga, ular bitta nol va 20 koeffitsientini qo'yadilar. Takrorlanishlarni aniqlashga asoslangan bunday algoritmlar usul deb ataladi seriyali uzunlikdagi kodlash(RLE,Kod uzunligini kodlash) Bir xil baytlarning katta takroriy ketma-ketligi ayniqsa grafik rasmlar bilan ajralib turadi. Usul "pikselga bayt" formatidagi grafik tasvirlar uchun juda samarali (masalan, formatlar) Pcx yoki BMP).

Qattiq disklarda zaxira nusxalarini yaratishda, fayllarni siqish paytida ish joyida daromad olishning yana bir imkoniyati mavjud, bu ma'lumotlarning ko'payishi bilan emas, balki kompyuterning fayl tizimini tashkil qilish bilan bog'liq. Uning mohiyati shundan iboratki, katta yoki kichik har qanday fayl faqat diskdagi butun sonlarni egallashi mumkin. FAT16 fayl tizimida qattiq diskda 65536 dan ortiq klaster bo'lishi mumkin (2<sup>16</sup>). Bu shuni anglatadiki, hajmi 1 dan 2 GB gacha bo'lgan drayvlar uchun klaster hajmi 32 Kbni tashkil qiladi.

Katta guruhli fayllarni bitta faylga siqish paytida, fayl tizimining mantiqsiz tashkil etilishi natijasida yo'qotishlarning kamayishi tufayli bitta faylga kamida 16 Kbayt tejash kerak.

### References:

1. A.A. Abduqodirov, A.G' Hayitov, P.P. Shodiyev. Axborot texnologiyalari.
2. S.S. G'ulomov. Axborot tizimlari va texnologiyalari.
3. A.M. Abduvohidov, B.K. Pozilov. Zamonaviy axborot texnologiyasi.