**Komprimácia, kompresia**

Komprimácia dát je proces, pri ktorom sa znižuje objem dát.

Kompresný pomer je pomer skomprimovaných dát ku neskomprimovaným. Napr. kompresný pomer 1:4 znamená, že došlo k 4-násobnému zmenšeniu objemu dát (takže 20 MB súbor má po komprimácii 5 MB, teda zmenšil sa na ¼ = 25% pôvodnej veľkosti). Komprimovať dáta má zmysel iba vtedy, ak je kompresný pomer menší ako jedna.

Existujú dva druhy komprimácie:

* nestratová - pri ktorej nedochádza k strate údajov. To znamená, že ak skomprimované dáta dekomprimujeme, získame úplne identické dáta. Takto sa balia napríklad textové, programové a iné súbory. Kompresia sa deje na základe vynechania redundantných (nadpočetných) informácií.
* stratová - je proces, pri ktorom sa vynechajú tie údaje, ktoré sú pre celkový dojem z dát nepodstatné, ale dáta sa už po kompresii nikdy nedajú zrekonštruovať do pôvodnej podoby, časť informácií totiž chýba. Stratová koprimácia sa používa hlavne pre komprimovanie mediálnych súborov a to zvuk (mp3), obraz (jpg), video (mpeg). Aj pri niekoľko desaťnásobnej redukcii dát je výsledok komprimácie takmer nerozoznateľný od originálu.

Komprimácia sa vykonáva buď automaticky (uložením súboru v komprimovanom formáte JPG, MPEG, MP3) alebo pomocou špecialneho komprimačného programu.

**Komprimačné programy, archivácia**

Výhodou archivačných formátov ako sú RAR, ZIP, 7zip je, že dokážu do jedného balíka vložiť množstvo súborov čo je výhoda, ktorú určite oceníme napríklad pri posielaní viacerých príloch v e-maily . Ak však archivujeme súbory len za účelom zníženia ich veľkosti je dobre si uvedomiť, že komprimovať sa oplatí len súbory, na ktorých ešte nie je aplikovaná žiadna kompresia. Aj preto majú napríklad obrázky vo formáte TIFF alebo BMP po pridaní do archívu veľmi malú veľkosť, zatiaľ čo obrázky v JPG nezmenšíme takmer vôbec. To isté platí aj o hudobných súboroch. Veľmi efektívne je komprimovať textové dokumenty, databázy...

Komprimačné (pakovacie) programy boli vyvinuté za účelom zmenšenia veľkosti súborov, pri zachovaní ich pôvodného obsahu. Ich cieľom je zmenšenie (skomprimovanie) súboru, ktorý potom zaberá podstatne menej miesta, ako jeho pôvodný tvar. So skomprimovaným súborom však nie je možné vykonávať bežné operácie, t.j. čítať ho, zapisovať a pod. Na to, aby ho bolo možné meniť, je potrebné jeho dekomprimovanie, čiže rozpakovanie - vrátenie do pôvodného stavu.

Princíp fungovania: Predstavte si text, v ktorom sa 30 krát vyskytuje slovo archivácia. Znamená to, že takýto text zaberá 30x10=300 znakov. Keď namiesto slova archivácia použijeme nejakú značku, napr. #, ušetrí sa tým 30x9=270 znakov. Pri komprimácii sa teda každé slovo archivácia nahradí značkou # a do tabuľky kódov sa zaznačí: # = archivácia. Tým sa výrazným spôsobom zmenší veľkosť textového súboru. Komprimačné programy pritom pracujú na ešte nižšej úrovni. Namiesto opakovania znakov sledujú opakovanie bitov (0 alebo 1). Dekomprimácia je opačný proces ako komprimácia. V tomto prípade sa namiesto znakov tabuľky musia vložiť slová, ktoré zastupujú. Napr. namiesto značky # sa v komprimovanom texte vloží slovo archivácia.

K najznámejším komprimačným programom patria: WinZIP, WinRAR, 7zip, PowerArchiver, FilZip, Windows Commander.

**Komprimačné formáty:**

* Audio: mp3,ogg, mp2, wma, ...
* Video: mpg, xvid, wmv, divx, ...
* Obrázky: jpeg, gif, ...
* Dáta: zip, rar, arj, cab, tar, gz, ...

**Komprimačné (archivačné) programy** sú programy, spravidla distribuované ako shareware1, špecializované na komprimáciu. Je to napr. **PKZIP, LHARC, ARJ, RAR, WinZIP, WinRAR**, a mnoho ďalších. Komprimáciu umožňujú aj niektoré súborové manažery, kompresné utility sú súčasťou aj mnohých iných programov. Komprimovaným súborom sú priradené koncovky podľa typu

**WinZIP**



**PeaZip WinRAR 7-zip**

**Hamster**

komprimačného programu (napr **\*.arj, \*.rar, \*.zip** atď.)

Komprimačný program môže vytvoriť niekoľko druhov výstupných súborov:

* bežný archívny súbor - kapacita obmedzená len voľným miestom na cieľovom disku
* delený archív - vznikne rozdelením jedného veľkého archívu na niekoľko častí s danou maximálnou veľkosťou, napríklad pre uchovanie archívu na niekoľkých disketách (vhodné pre kapacitu niekoľko desiatok MB)
* samorozbaľovací archívny súbor (\*.exe)- SFX (SelF-eXtracting) archív je archív s pripojeným spustiteľným modulom, ktorý po spustení dokáže súbory z archívu rozbaliť. K rozbaleniu SFX archívu nie je teda treba žiadny externý program - sám vykoná dekompresiu do aktuálneho adesára

**Úlohy:**

1. a) Akú veľkosť bude mať 15 MB súbor po kompresii s pomerom 2 : 5?

b) Po skomprimovaní mal súbor veľkosť 9 kB. Aká bola jeho pôvodná veľkosť, ak kompresný pomer je 1 : 3?

c) Aký je kompresný pomer v kompresii, ak sa nám 3,6 GB súbor podarilo zmenšiť na 2,4 GB?

2. Vytvorte rozsiahly textový dokument (nakopírujte tento učebný materiál 10 krát za sebou do jedného súboru), zapíšte si jeho veľkosť. Potom ho skomprimujte. Zapíšte si veľkosť po komprimácii. Vypočítajte kompresný pomer použitej kompresie. Experimentujte s rôznymi kompr. algoritmami.