Komprese a šifrování

Komprese

Komprese (také komprimace) je způsob zpracování počítačových dat za účelem zmenšení jejich objemu (většinou kvůli jejich uložení nebo přenosu) a současného zachování co největšího množství informací. Komprese je prováděna tzv. kompresním algoritmem, který vyhledává a odstraňuje nadbytečné (redundantní) informace.

Kompresi dělíme do dvou základních kategorií:

- Ztrátová
- Bezeztrátová

Ztrátová komprese

Během ztrátové komprese se některé méně důležité informace ztrácejí bez možnosti rekonstrukce z vytvořených dat. Tento postup způsobuje ztrátu kvality původního souboru a používá se především pro ukládání obrazových a zvukových záznamů, jelikož umožňuje výrazné zmenšení objemu dat.

Algoritmy ztrátové komprese jsou založeny na lidském vnímaní a zaokrouhlují informace, které jsou pro naše vnímání méně podstatné (např. zprůměrování barevných odstínů, odstranění hůře slyšitelných částí).

Mezi nejčastější formáty ztrátové komprese patří:

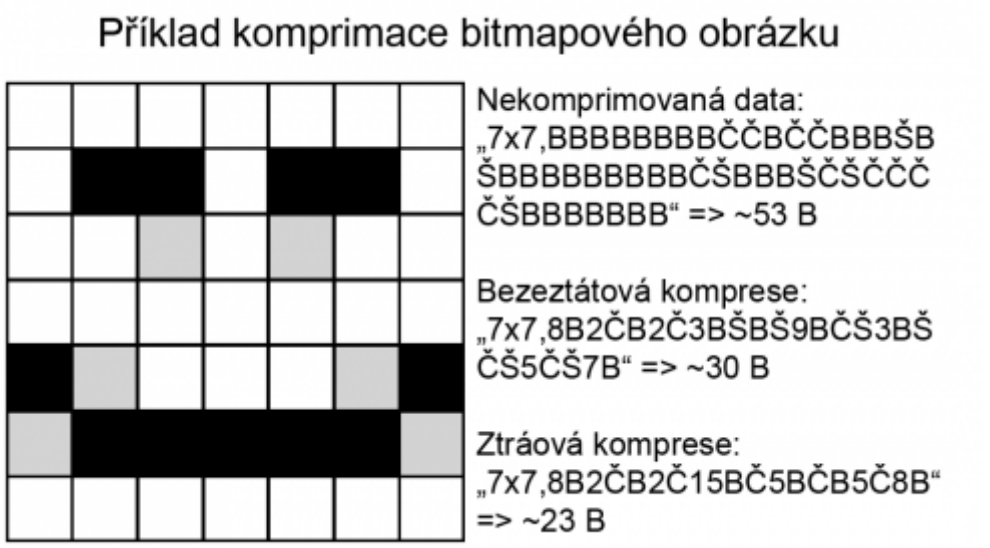
- **JPEG** (obrázky, fotografie)
- **MPEG** a **H.264** (video)
- **MP3** a **AAC** (audio)
- **AMR** (řeč)

Bezeztrátová komprese

Bezeztrátová komprese umožňuje přesnou rekonstrukci původních dat z komprimovaných souborů tzv. dekompresí. Využívá se především pro kompresi textových dokumentů nebo jiných souborů, které vyžadují zachování původních informací.

Algoritmy bezeztrátové komprese se většinou liší podle formátu souboru a často se jich využívá více najednou. Během komprese se vyhledají opakující se data (stejné fráze, pixely, posloupnosti atd.), které se poté nahradí pouhým odkazem na jejich skutečnou hodnotu. Algoritmy se také snaží statisticky předpovídat pravděpodobnost výskytu určitých znaků, kterým pak podle počtu výskytů přiřadí kratší, nebo delší informaci pro jejich zápis. Tak funguje např. Huffmanovo kódování.

Mezi nejpopulárnější algoritmy patří komprese **LZ**, která dala základ algoritmům **LZMA** (využívaný programem 7zip), **LZSS** (WinRAR) a dalším. Nejběžnějšími bezeztrátově komprimovanými formáty jsou **FLAC** (audio) a **PNG** (obrázky).



Šifrování

Šifrování je proces, při kterém se čitelná nezabezpečená data převádí na nečitelná data šifrovaná, které je schopen rozšifrovat pouze majitel dešifrovacího klíče.

Šifry můžeme dělit podle možnosti zpětného získání informace na:

- Reversibilní - informace je možné zpětně získat
- Ireversibilní - informace se uchovává ve formě hashe a není možné ji zpětně získat

Reversibilní šifrování

Reversibilní šifrování můžeme dělit na:

- Symetrické
- Asymetrické
- Hybridní

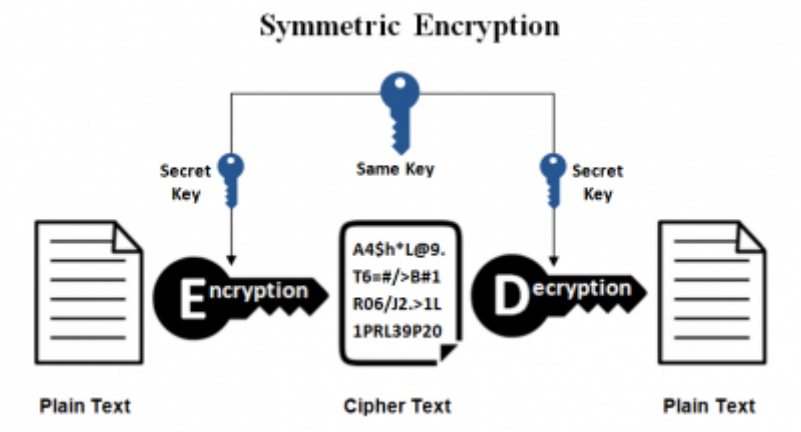
Symetrické šifry

Symetrické šifry mají jeden klíč, který slouží k šifrování i dešifrování. Nezabezpečená zpráva je pomocí tohoto klíče zašifrovaná, čímž se stává nečitelnou. Následně je zpráva odeslána příjemci, který ji za pomocí stejného klíče dešifruje, čímž se zpráva stane opět čitelnou.

Pro symetrické šifry platí, že jejich dešifrování je přesným opakem jejich šifrování. Největším problémem symetrických šifer je sdílení samotného klíče mezi uživateli. Šifry jsou také omezeny pouze na uživatele, kteří znají klíč. Se zvyšujícím se počtem uživatelů se také snižuje bezpečnost klíče.

Jednoduchým příkladem symetrické šifry může být tzv. Caesarova šifra, kterou římský vojevůdce používal k šifrování vojenských zpráv. Šifra je založena na tom, že si na začátku určíme klíč (třeba číslo 2) a posuneme každé písmeno zprávy v abecedě o vybrané číslo.

Když vezmeme například slovo „Pascal“ a zašifrujeme ho Caesarovou šifrou s klíčem 2, tak dostaneme „Rctecn“.

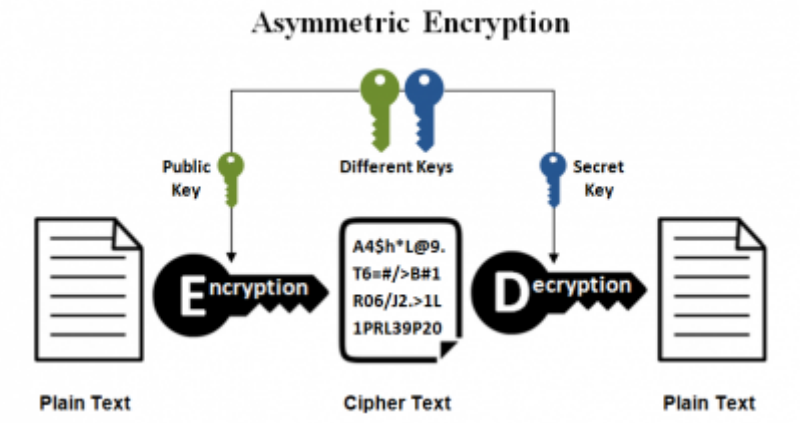


Asymetrické šifry

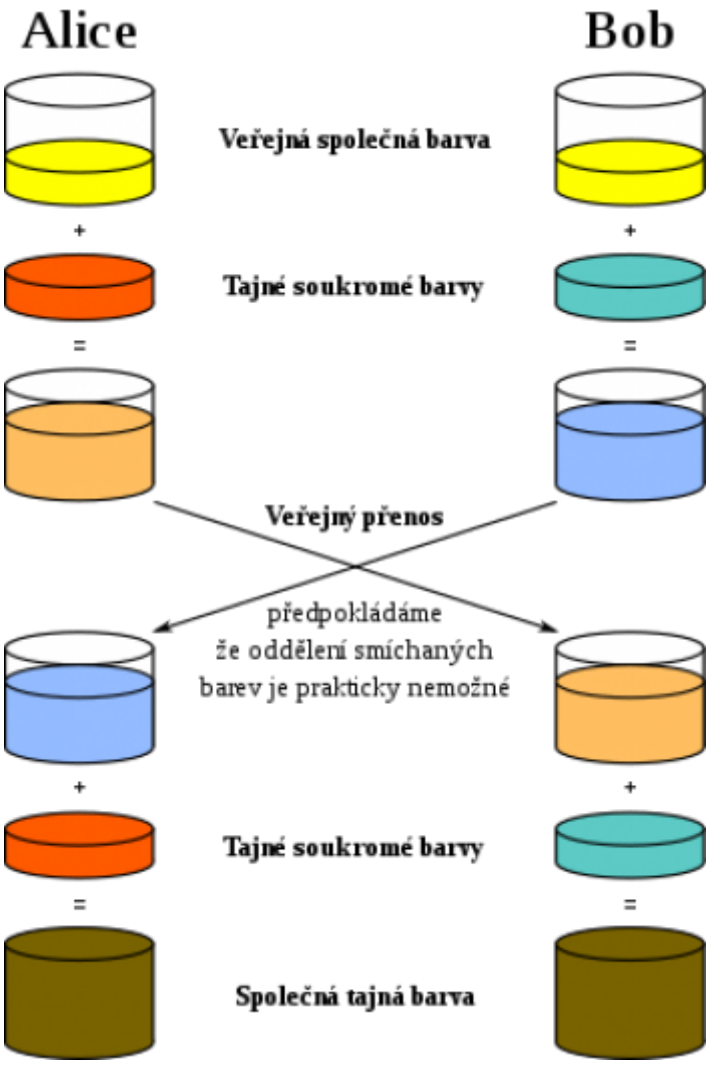
Asymetrické šifry využívají dva klíče, z nichž jeden je veřejný a druhý soukromý. Oba klíče se dají použít pro zašifrování určitých dat, ale pro jejich dešifrování se musí použít klíč druhý. Veřejný klíč se nahraje na server nebo jinak sdílí přes internet a využívají ho ostatní pro šifrování zpráv pro majitele soukromého klíče.

Při běžné komunikaci je tak čitelná zpráva zašifrována veřejně přístupným klíčem příjemce. Tato zpráva se stává nečitelnou a je možné ji dešifrovat pouze za použití soukromého klíče. Zpráva je tedy odeslána příjemci, který jako jediný zná druhý klíč, ten zprávu dešifruje a získá původní text.

Asymetrické šifrování se dá také využít obráceně. Odesílatel vytvoří zprávu a její hash zašifruje svým soukromým klíčem. Příjemce pak může zašifrovaný hash dešifrovat za pomocí veřejného klíče a výsledný hash porovnat s hashem obdržené zprávy. Pokud jsou stejné, znamená to, že zprávu vytvořil opravdu její odesílatel a že nebyla nijak upravována. Tímto způsobem je potvrzena identita odesílatele a jedná se o tzv. digitální podpis. Pro ověření podpisu se k zašifrovanému hash kódu přidává ještě certifikát důvěryhodné třetí strany, který potvrzuje totožnost fyzické osoby s klíčem.



Základem většiny systémů s asymetrickým šifrováním je šifra RSA, která je založena na principu obtížnosti faktorizace (rozkladu na součin) velkých čísel. K vytvoření a výměně sdíleného privátního klíče přes veřejné komunikační kanály se nejčastěji používá Diffieho–Hellmanova výměna klíčů.



Hybridní šifry

Hybridní šifry se snaží eliminovat problémy předchozích dvou typů - u symetrických problém s přenosem klíče a u asymetrických dlouhou dobu a velkou náročnost výpočtu. U hybridního šifrování nejdříve zašifrujeme zprávu symetricky náhodným klíčem a pak samotný klíč zašifrujeme asymetricky a odešleme spolu se zprávou. Příjemce si nejdříve rozšifruje pomocí asymetrické šifry klíč a poté klíčem dešifruje symetricky zašifrovanou zprávu. Tento typ používá například protokol HTTPS.

Ireversibilní šifrování

Při ireversibilním zašifrování dostaneme ze zprávy tzv. hash. Z hashe už není možné získat původní informaci. Ireversibilní šifry se dají použít například k uložení hesla v aplikaci. Uživatel si zvolí své heslo a z toho se následně udělá hash. Při každém následném přihlášení se vezme zadané heslo a porovná se s hashem hesla uloženým v databázi.