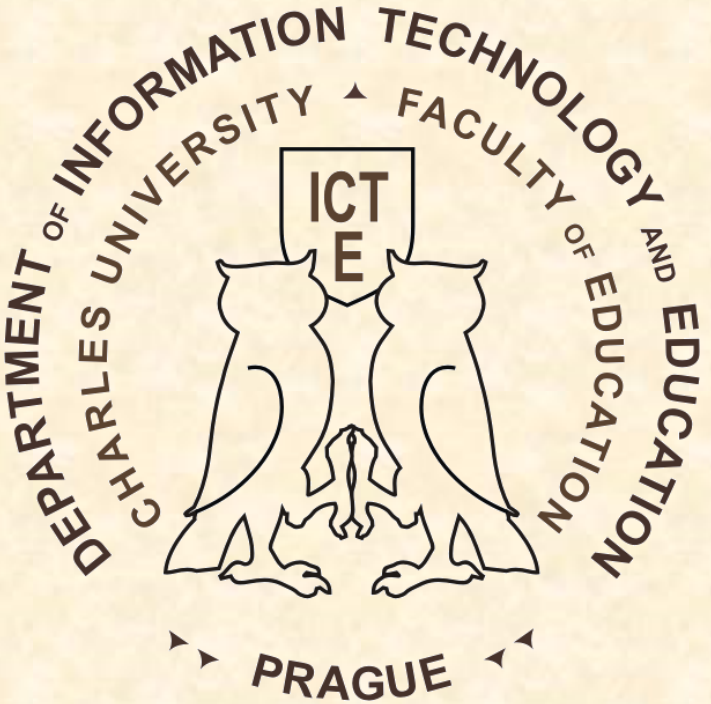
Samodetekující/Samoopravné kódy



Samodetekující kódy

- *Samodetekující kód/kódy chybu detekující*
  - Ověření správnosti zadaných dat, odhalení chyby v datech.
  - Nelze zjistit původní informaci.
  - Kontrolní součet (checksum), parita, modulo, CRC,…
  - Hashovací fce. („otisk“) – MD5, SHA,…
  - (Hammingův kód).

Kontrolní součet (checksum)

- Doplňková informace, předává se spolu s vlastní informací.
- Slouží k ověření, zda je vlastní informace úplná a zda při jejím přenosu nedošlo k chybě.
- Kontrolní součet je výsledkem nějaké předem určené operace, provedené s vlastní informací.
- Příjemce informace má možnost sám kdykoliv spočítat svůj vlastní kontrolní součet.
- Jakmile vypočtený kontrolní součet nesouhlasí s předaným kontrolním součtem -> během přenosu došlo k poškození zprávy nebo k poškození kontrolního součtu.

Kontrolní součet (checksum)

- Jedna z možností je zaslání úplné kopie celé informace
  - Pokud při přenosu informace došlo k jejímu poškození, příjemce pozná, že se obě kopie navzájem liší.
  - Nevýhodou této metody je velká redundance – je třeba přenášet dvojnásobný objem dat.
- Prakticky se pro kontrolní součet používá jen menší dodatková informace, např. součet všech čísel.

Paritní bit

- Redundantní bit přidaný k datovému slovu.
- Obsahuje paritní informaci o počtu jedničkových bitů ve slově.
- Paritní bit je určen k jednoduché detekci chyby ve slově.
- Pomocí paritního bitu lze detekovat počet chyb (chybných bitů) ve slově.
- **Lichá parita** znamená lichý počet jedničkových bitů ve slově (i s paritním bitem), **sudá parita** sudý počet jedničkových bitů ve slově.

Cyklický redundantní součet (CRC)

- K detekci chyb během přenosu či ukládání dat.
- Rozšířený způsob realizace kontrolního součtu.
- Kontrolní součet bývá odesílán či ukládán společně s daty, při jejichž přenosu nebo uchovávání by mohlo dojít k chybě
- Po převzetí dat je znovu nezávisle spočítán - pokud je odlišný od přeneseného nebo uloženého, je zřejmé že při přenosu nebo uchovávání došlo k chybě.
- V určitých případech je možné chybu pomocí CRC opravit.
- Ethernet.

Message-Digest algorithm (MD5)

- Rodina hašovacích funkcí, která z libovolného vstupu dat vytváří výstup fixní délky, který je označován jako **hash** (otisk, miniatura, fingerprint).
- Matematická funkce (resp. algoritmus) pro převod vstupních dat do (relativně) malého čísla.
- *Jeho hlavní vlastností je, že malá změna na vstupu vede k velké změně na výstupu, tj. k vytvoření zásadně odlišného otisku.*
- Pro kontrolu integrity souborů (md5sum) nebo ukládání hesel.
- MD5 je popsán v internetovém standardu RFC 1321 a vytváří otisk o velikosti 128 bitů. Byl vytvořen v roce 1991 Ronaldem Rivestem.
- Od r. 2004 není samostatně považován za bezpečný.

Secure Hash Algorithm (SHA)

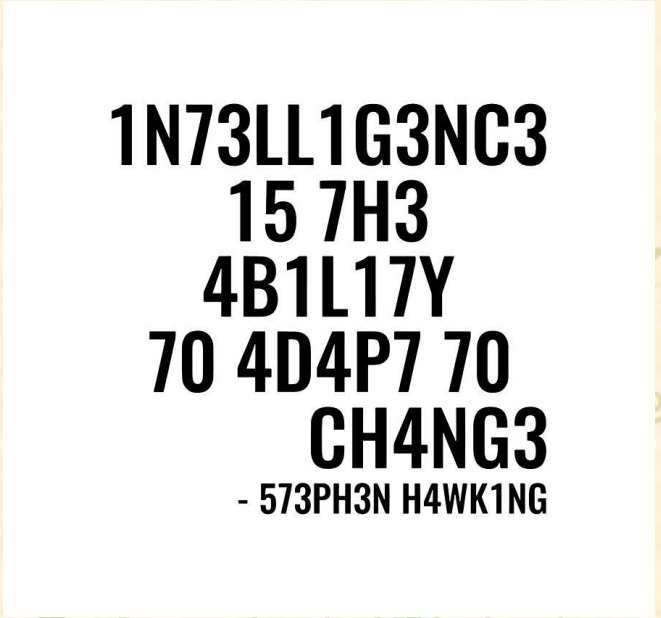
- Hašovací funkce, která vytváří ze vstupních dat výstup (otisk) fixní délky.
- Ze znalosti otisku je prakticky nemožné rekonstruovat vstupní data.
- SHA je rodina pěti algoritmů (souhrnně SHA-1 a SHA-2).
- Nástupce hašovací funkce MD5.
- Současnými prostředky je prakticky nemožné najít zprávu, která odpovídá svému otisku nebo najít dvě rozdílné zprávy, které mají stejný otisk.

Samoopravný kód

- Detekce a korekce chyb, vzniklých při přenosu dat (poškození kvůli poruchám fyzikálního prostředí, ve kterém k přenosu dochází – šum). Odolnost přirozeného jazyka vůči šumu je založena na redundanci (i přes 50%):

  "V sbt včr jsm šl s kmrdm n pv"

- Dovolují při dostatečně malém poškození zrekonstruovat původní data.
- RS kódy, Repetition kód, Hammingův kód.
- QR kód, CD.



Samoopravný kód

- *Šum náhodný (random)* - pravděpodobnost, že se nějaký znak během přenosu změní (poškodí), je stejná pro všechny znaky a není nijak ovlivněná chybami při přenosu sousedních znaků.
- *Shlukové chyby (burst errors)* – například škrábnutím kompaktního disku, blýskáním, slunečními erupcemi, apod.

„Hospůrka"-->"Hospůdka"; „Hoshůzka"-->"Hospůdka"x"Pochůzka"

Opakovací kód (Repetition/Repeat code)

- Výpočet pravděpodobnosti.
- Poslání „něco navíc", aby se eliminovaly chyby.
- Chyba = přijatý bit se liší od odeslaného.
- Snaha aby celkový objem posílaných dat byl co nejmenší.
- Pravděpodobnost správného dekódování

  Spolehlivost/chybovost přenosového kanálu:

  $(1-p)^n$ …n=posloupnost bitů (0 nebo 1)/počet bitů

  …p=pravděpodobnost chyby u každého bitu

Hammingův kód

- Hamming (7,4,3)
  - 7b = 4b informační + 3b kontrolní.
- Zjištění chyby + její umístění (paritní bit).
- **Fanova rovina** (graf) – 7 vrcholů/bodů, 7 hran/přímek => 16 vektorů = kódová slova; každá 2 kódová slova se liší alespoň ve 3 bitech/souřadnicích (Hammingova vzdálenost) (při poškození 1 bitu lze opravit).
- Posílá se 7n/4 bitů (posílání kódových slov).