HEŠOVACÍ FUNKCE

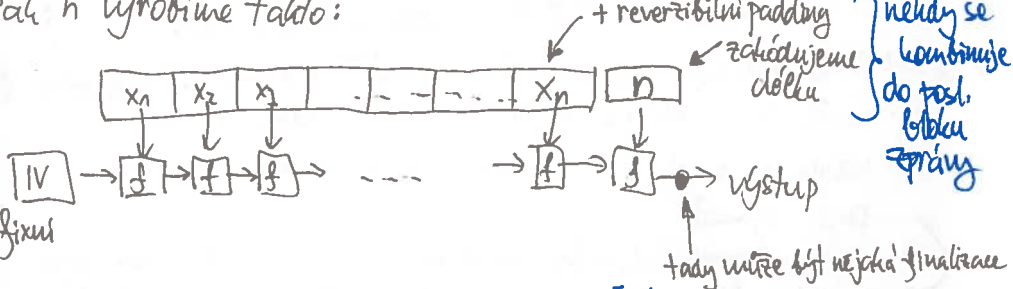
Cíl: funkce $h: \{0,1\}^* \to \{0,1\}^b$

- ideálně nerozlišitelná od náhodné funkce
  ... to ale neumíme nadefinovat, neboť h nemá klíč
- Typické požadavky: ① Neumíme najít kolizi: $f(x) = f(x')$ pro $x \neq x'$
  ② Neumíme najít druhý vzor: pro x nenajdeme x': $f(x) = f(x')$
  ③ Neumíme invertovat: pro y nenajdeme x t.ž. $f(x) = y$.
- ③ ⇐ ② ⇐ ①
  ↳ pokud umíme $y = f(x)$, inverze y by nejspíš našla jiné x (typicky má nekonečně možností)

Merkleova-Damgårdova konstrukce

- pořídíme si kompresní funkci $f: \{0,1\}^b \times \{0,1\}^b \to \{0,1\}^b$ (viz dále)

Pak h vyrobíme takto:



- Pokud f je odolná proti kolizím, h je též odolná.

  [dává smysl, pokud kompr. fce má vstupy různě dlouhé → umíme větší vnitřní stav než blok]

  Dk: Nechť $h(x_1 \| \dots \| x_n) = h(x_1 \| \dots \| x_{n'})$
  ① Pokud $n \neq n'$, máme kolizi v posl. volání f.
  ② Pokud $n = n'$: posl. bloky se nerovnají ⇒ kolize v f
  rovnají ⇒ kolize kratších zpráv ⇒ iterujeme

- Kdybych nepřihešovával délku:
  - Při kolizi h najdu postupem popsaným buď kolizi f nebo mezi $f^{-1}(IV)$ ... ale to jsem nepředpokládal, že nejde (z odolnosti f to neplyne).
- Length extension – v tom se liší od náhodné funkce!

Odolnost proti kolizím

- Řádově $2^{b/2}$ vstupů (jakýchkoli – mohou to být smysluplné zprávy) stačí na "narozeninovou" kolizi, ale potřebuji paměť $2^{b/2}$
- Málo paměti: $x_{i+1} = h(x_i)$, želva a zajíc se honí po lízátku
  → Jak to udělat se smysluplnými zprávami?
  Pořídím si parametrizované zprávy (b míst, kde si mohu vybrat mezi 2 smysluplnými variantami)
  a pak volím x_{i+1} jako zprávu parametrizovanou $h(x_i)$.
  → Varianta s množinovým naroz. útokem: (ten už zas potřebuje paměť)
    - "Oběť" je ochotna podepisovat "hodné" zprávy, já chci podepsat "zlou" zprávu
    - Pořídím si parametrizovanou hodnou a zlou zprávu
    - Vygeneruji heše $2^{b/2}$ hodných a $2^{b/2}$ zlých zpráv
    - S velkou PP ∃ průnik obou množin hešů.
- U M-D konstrukce umím v čase $k \cdot 2^{b/2}$ vyrobit 2^k-násobnou kolizi
  - najdu x_1 a x_1' t.ž. $f(IV, x_1) = f(IV, x_1') =: y_1$
  - najdu x_2 a x_2' t.ž. $f(y_1, x_2) = f(y_1, x_2') =: y_2$
  atd. k-krát
  - zprávu mohu libovolně kombinovat z x_i a x_i', vždy vyjde stejný heš.
  → z toho útok na konkatenaci dvou různých hešů

Kde sehnat kompresní funkci? z nichž aspoň jeden je M-D

- Daviesova-Meyerova konstrukce z blokové šifry:
  $f(a,b) = E_a(b) \oplus b$
  - Proč $\oplus b$? Bez toho: $E_a(b) \to y$, $D_{a'}(y) \to b'$ … pak $f(a,b) = f(a',b')$.
  - Pozor, rozbije se pro DES: $f(\bar{a},\bar{b}) = E_{\bar{a}}(\bar{b}) \oplus \bar{b} = \overline{E_a(b)} \oplus \bar{b} = E_a(b) \oplus b = f(a,b)$
  - Věta: Je-li E/D ideální šifra, f je odolná proti kolizím.
    Přesněji: Útočník, který zavolá E q-krát, najde kolizi s pstí $\leq q^2/2^b$ pro $q < 2^{b/2}$

    Dk: BÚNO útočník nevyhodnocuje E/D redundantně.
    Pokud se zeptá na $E_x(y)$, dozví se $f(x,y) = E_x(y) \oplus y$.
    Pokud na $D_x(y)$, dozví se $f(x, D_x(y)) = y \oplus D_x(y)$.

Při i-tém pokusu nastane kolize, pokud se strefím do některé z $i-1$ předchozích hodnot ... pro 1 cílovou hodnotu to nastane pro právě 1 volbu výsledku $\in D$. Proto:

$$\Pr[\text{dvojice se shodne}] \leq 1/(2^b-(i-1)) \leq 2^{-(b-1)}$$

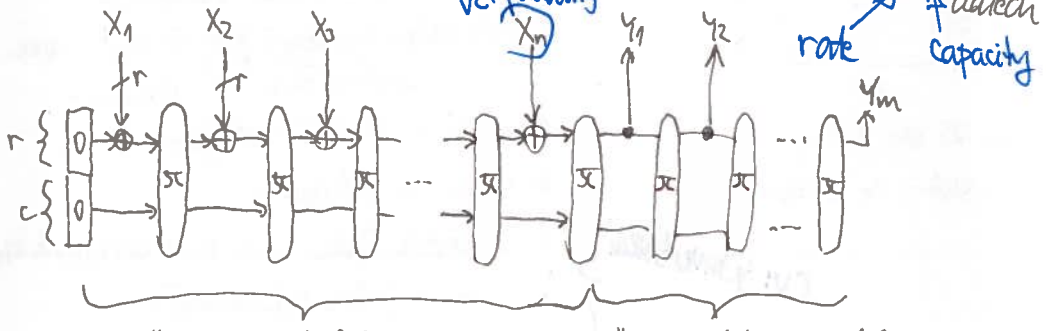
$$\Pr[\text{najdu kolizi}] \leq \Pr[\text{dvojice se shodne}] \cdot \#\text{dvojic} \leq q^2/2^b.$$

($\#\text{dvojic} \leq q^2/2$)

max. $i-1$ hodnot je už obsazeno z předch. dotazů

- MD5: 128b výsledek, od roku 2004 známé kolize (i chosen-prefix) (Rivest 1992)
  - to je málo
  - ale invertovat ji nedokážeme
- SHA-1: 160b výsledek, (NSA)
  - 2015 kolize v kompres. funkci
  - 2017 plná kolize (zatím dost náročná)
  - konstrukce podobná Daviesovi-Meyerovi (příslušné šifře se občas říká SHACAL)
- SHA-2: verze s 224, 256, 384, 512 bity výsledku
  - mohutnější, ale podobná struktura
  - zatím nejsou rozbité
- SHA-3 (veřejná soutěž NISTu, finále 2015)

"Houbovitá funkce" ... uvažujeme permutaci π na $w=(r+c)$-bit.



- Odolnost proti kolizím je ovlivněna
  - velikostí vymačkávaného výstupu
  - kapacitou c (interní kolize)

- Interní kolizní útok

kruhem samými nulami

Po řádově $2^{c/2}$ krocích najdu $b_i = b_j$, $i<j$.
Zprávy 0^i a $0^{j-1}(a_0 \oplus a_j)$ vedou na tentýž vnitřní stav
⇒ vymačká se stejný výstup. [mohu pak dolepit stejné pokrač. za obě prefixy a zase mám kolizi...]

⇒ security level je nejvýše $c/2$.

- SHA-3 je houba s permutací Keccak šířky 1600 bitů.
Standard definuje:

SHA3-224	r=1152	c=448	vždy c = 2 · délka výstupu
⋮			
SHA3-512	r=576	c=1024	
SHAKE128	r=1344	c=256	XOF = extendable output functions (výstup lib. velikosti) ↳ PRNG
SHAKE256	r=1088	c=512	

vymačkává se jedním krokem

- Jak vypadá Keccak?

Vnitřní stav je kvádr

... 25 slov šířky w
(v SHA-3 je w=64)

Provádíme $12 + 2\log w$ rund, v každé:

- ke každému sloupečku přixoruji paritu 2 okolních sloupečků (to dělám paralelně pro všechny sloupce v 1 svislém řezu bit-slicingem
- každé z 25 slov zrotujeme
- slova permutujeme
- v každém řádku: $X_i \leftarrow X_i \oplus (\neg X_{i+1} \,\&\, X_{i+2})$ [to je jediná nelinearita]

CV: je invertibilní

- přimíchám k 1 slovu rundovou konstantu

- Různé módy použití (SHA-3, SHAKE, další budoucí) mají různý padding ⇒ jsou rozlišitelné