Hašovací funkce

Jonáš Chudý

Úvod do kryptologie

Základní definice

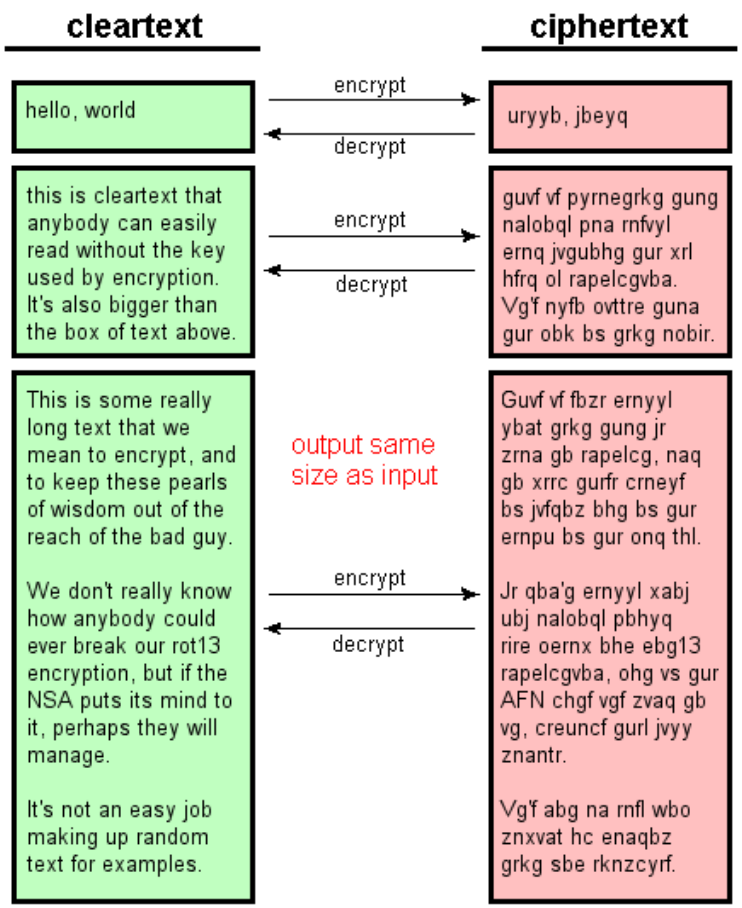
Kryptografická hašovací funkce

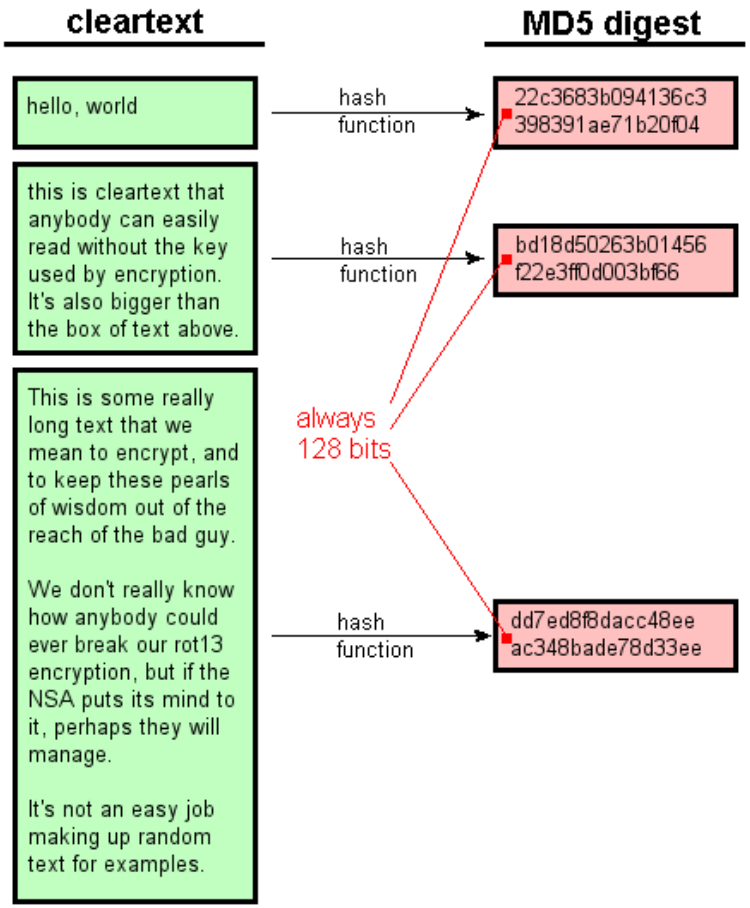
Kryptografickou hašovací funkcí nazveme zobrazení h, které vstupu X libovolné délky přiřadí obraz $h(X)$ pevné délky m a navíc splňuje následující vlastnosti:

- **snadný výpočet obrazu**
  -při daném h a X je výpočetně "jednoduché" získat $h(X)$
- **odolnost proti nalezení vzoru**
  -pro Y z oboru hodnot h je výpočetně "nemožné" najít vzor X
- **odolnost proti nalezení druhého vzoru**
  -pro X a $h(X)$ je výpočetně "nemožné" najít jiné $\bar{X}$, pro které platí $h(\bar{X}) = h(X)$

Poznámka: Pojmem výpočetně "nemožné" zde zpravidla rozumíme, že nutný počet operací je superpolynomiální v délce vstupu.

Hašování vs. šifrování





Skupiny hašovacích funkcí

nepoužívající tajný klíč

- **jednosměrné hašovací funkce**
  - vyhovují úvodní definici
- **hašovací funkce odolné vůči kolizi**
  - navíc vyhovují podmínce, že je výpočetně "nemožné" najít 2 různé vstupy $X, \bar{X}$, pro které $h(\bar{X}) = h(X)$

používající tajný klíč (MAC - Message Authentication Code)

- výsledná hash navíc závislá na tajném klíči K
- pro dané h, X je výpočetně "nemožné" najít $h(K, X)$
- i pokud je známá velká množina dvojic $\{X_i, h(K, X_i)\}$ je výpočetně "nemožné" najít K nebo spočítat $h(K, \bar{X})$ pro jiné $\bar{X}$

Další vlastnosti

avalanche efekt

- malá změna vstupních dat indukuje velkou změnu hashe

```
$ cat file1
This is a very small file with a few characters

$ cat file2
this is a very small file with a few characters

$ md5sum file?
75cdbfeb70a06d42210938da88c42991  file1
6fbe37f1eea0f802bd792ea885cd03e2  file2
```

kolize

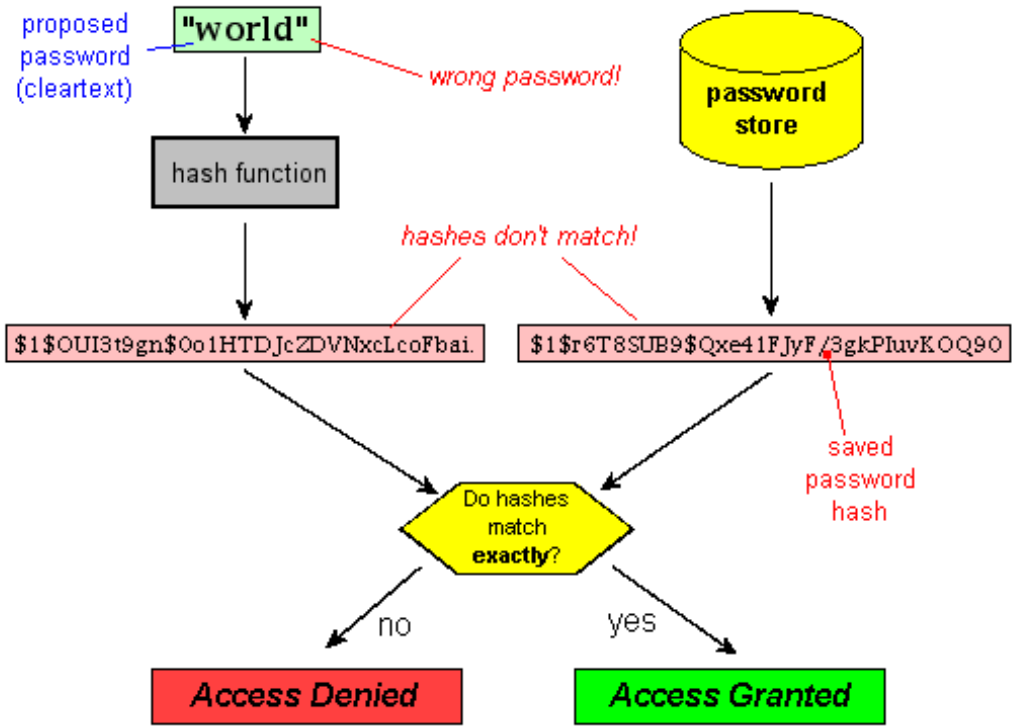
- značné množství vstupů produkuje stejnou hash
- obtížné takové vstupy nalézt (pokud odolná vůči kolizím)
- např. pro 128 bitovou hash existuje $3.4x10^{38}$ možností výstupu

Aplikace hašování

- **ověřování celistvosti souborů**
  - kontrolní bitové sumy staženého souboru
  - důležitá odolnost proti nalezení druhého vzoru
- **digitální podpis**
  - viz následující přednáška
  - důležitá odolnost vůči kolizi
- **prokazování znalosti - HMAC**
- **generátory pseudonáhodných čísel**
- **odvozování klíčů**
  - odvozování pseudonáhodných šifrovacích klíčů z hesel
- **hašovací tabulka**
  - vyhledávací datová struktura

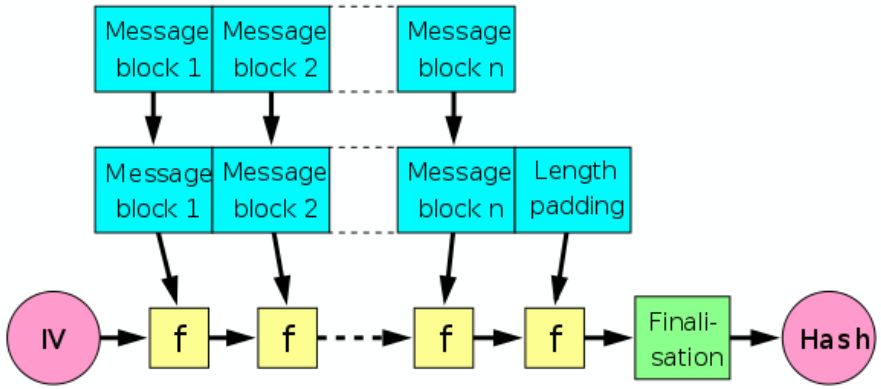
Aplikace hašování

- **hašování hesel**
  - důležitá odolnost proti nalezení vzoru



Konstrukce (Merkle-Damgard)

- hašovací funkce $h(X)$ založena na kompresní funkci $f(Z)$
  - funkce f provádí kompresi vstupu Z, který má pevnou délku
- vstup X rozdělen do n bloků pevné délky b
- doplnění posledního bloku na požadovanou délku b
  - možný způsob: posledních r bitů obsahuje informaci o délce vstupu, zbylé bity doplněny nulami
- dále iterativní postup dle schématu



- zachována případná odolnost vůči kolizím kompresní funkce f

Útoky na hašovací funkce

narozeninový útok

- založeno na narozeninovém paradoxu

$$P(\text{narozeniny ve stejný den aspoň 2 lidé z 23}) \geq \frac{1}{2}$$

- obecněji počet pokusů N k nalezení kolize s pravděpodobností p pro H možných výstupů

$$N(p, H) = \sqrt{2H \log\left(\frac{1}{1-p}\right)}$$

- např. 128 bitová hash $N(0.5, 3.4x10^{38}) \approx 2.2x10^{19}$

Známé hašovací funkce

Tiger

- z roku 1995, používaná v peer-to-peer aplikacích

rodina MD(message-digest algorithm)

- MD4, MD5 - v současné době již prolomené, nepoužívají se v bezpečnostních aplikacích

rodina SHA (secure hash algorithm)

- SHA1 - 1995, NSA, 160 bitová hash
  - 2005 - objeven algoritmus nalezení kolize podstatně rychleji než hrubou silou (ale stále výpočetně nemožné)
- SHA2 - 4 funkce, zatím neobjeveny bezpečnostní slabiny
- SHA3 - soutěž 2012

MD5

- 1991, Ronald Rivest
- vytváří hash o velikosti 128 bitů
- 1996, kolize kompresní funkce
- 2004, projekt MD5CRK- narozeninový útok
- 2004, Wang, Feng - kolize pro úpný MD5, 1 hodina na IBM clusteru
- 2005, Lenstra, Wang - dva X.509 certifikáty s rozdílnými veřejnými klíči a se stejnou hashí
- 2005, V. Klíma - MD5 kolize během několika hodin na notebooku
- 2006, V. Klíma - MD5 kolize během minuty na notebooku (tunelování)
- 2010, Xie, Feng - single-blok kolize MD5

Děkuji za pozornost

Použitá literatura

Zdroje

- B. Preneel, *Analysis and Design of Cryptographic Hash Functions*,2003
- http://www.unixwiz.net/techtips/iguide-crypto-hashes.html
- http://cs.wikipedia.org/wiki/Secure_ Hash_ Algorithm
- http://cs.wikipedia.org/wiki/Message-Digest_ algorithm
- http://cs.wikipedia.org/wiki/Kryptograficka_ hasovaci_ funkce
- http://en.wikipedia.org/wiki/Merkle-Damgard_ construction