Hašovací funkce, MD5 a cínský útok

To nejlepší, co pro vás kryptologové mohou udelat je, když vás presvedcí, abyste jim slepe neduverovali. Je to nutná podmínka pro to, abyste ani vy ani oni neusnuli na vavrínech.

Vlastimil Klíma

v.klima@volny.cz, http://cryptography.hyperlink.cz

Seminár **B**ezpecnost **I**nformacních **S**ystému v praxi, MFF UK Praha, 22. 11. 2004

1

Obsah (1)

- prolomení hašovacích funkcí MD5, MD4, SHA-0, HAVAL-128, RIPEMD v dobe konání konference CRYPTO 2004
- objev v kryptoanalýze, který bude mít vliv na bezpecnostní praxi v sektoru IS/IT
- cínský výzkumný tým nezverejnil metody prolamování, jen výsledky (kolize)
  - zpusoby využití hašovacích funkcí v protokolech a aplikacích
  - vysvetlení a komentování prolomení
  - praktické dusledky

2

Obsah (2)

- kde je nutné MD5 vymenit, a kde ješte lze MD5 používat
- **zmenit postoj** ke kryptografickým technikám jako k necemu zvláštnímu a pracovat s nimi jako s jakýmikoliv jinými bezpecnostními nástroji, to znamená **a priori neduverovat, sledovat vývoj v dané oblasti a bežne provádet update nebo upgrade**
- nutnost vyvíjet nové aplikace kryptograficky modulárne
- protože se to tak nedelo, bude nyní velmi obtížné vymenit prolomené hašovací funkce, zejména MD5,

3

Manažerské shrnutí

- Provést **revizi všech aplikací**, kde jsou použity hašovací funkce MD4, MD5, SHA-0, RIPEMD a HAVAL-128.
- Je-li nekterá z techto funkcí použita **pro úcely digitálních podpisu** (s klasickým úcelem zajištení nepopiratelnosti), je nutno tuto funkci nahradit. Podle okolností provést náhradu za nekterou z funkcí, které jsou považovány za bezpecné: SHA-1, SHA-256, SHA-384 nebo SHA-512, nejlépe SHA-512.
- Je-li nekterá z techto funkcí použita **pro úcely HMAC nebo PRNG**, nechat pro jistotu posoudit, zda je toto užití bezpecné nebo ne.
- Nové systémy budovat tak, aby se kryptografické nástroje v nich mohly **pružne menit**. Pokud je to možné, prejít na toto pravidlo postupne i u stávajících systému.
- Zajistit prubežné sledování vývoje aplikované kryptologie a zavést **mechanismus pravidelného hodnocení používaných kryptografických technik**.

4

Hašovací funkce

Vlastnosti:

- Libovolne dlouhý vstup
- Pevne definovaná délka výstupu

haš, hash, hašový kód = výstupní kód s predem pevne definovanou délkou

Príklady na MD5 – 128bitový hašový kód

5

Kryptografická hašovací funkce

- Jednocestnost (jednosmernost), preimage resistance
  - Pro každé x je jednoduché vypocítat h(x)
  - Pro náhodne volené x je výpocetne neproveditelné z h(x) urcit x
- Odolnost proti kolizi prvního rádu, collision resistance
  - Je výpocetne neproveditelné nalézt **libovolné ruzné x, y** tak, že h(x) = h(y)
- Odolnost proti kolizi druhého rádu, second preimage resistance
  - Je výpocetne neproveditelné **k danému náhodnému x** nalézt **druhý vzor** y ruzný od x tak, že h(x) = h(y)

6

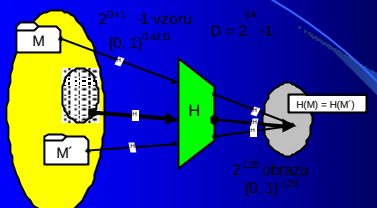
Využití

- Integrita dat (s klícem nebo bez)
- Jedinecnost vztahu data – haš (digitální otisk dat)
  - Kontrola neporušenosti dat, shoda velkých souboru dat, prokazování autorství
- Ukládání hesel
- Autentizace (prokázání znalosti)
- Digitální podpisy

7

Odolnost proti kolizi

Bezpecnostní požadavky



Narozeninový paradox.
Kolize lze nalézt narozeninovým paradoxem se složitostí $2^{n/2}$ operací pro n-bitové hašovací kódy.
Proc paradox? – kolize 1. / 2. rádu
Pokud lze nacházet významne jednodušeji, hašovací funkce se považuje za prolomenou.

8

Odolnost proti nalezení druhého vzoru

Bezpecnostní požadavky

- Druhý vzor lze nalézt se složitostí 2^n operací pro n-bitové hašovací kódy
- Pokud lze nacházet významne jednodušeji, hašovací funkce se považuje za prolomenou.

9

Náhodnost

Bezpecnostní požadavky

- Casem se zacala využívat prirozená vlastnost hašovacích funkcí chovat se podobne jako náhodné orákulum (random oracle). Využití v PRNG, KDF.
- Pokud se najde „významná odchylka“ od náhodného chování, hašovací funkce se považuje za prolomenou.

10

Bezpecnostní požadavky

…Jádro veci, které bylo zapomenuto:
Hašovací funkce nejsou prokazatelne bezpecné nástroje a jejich bezpecnost (jednosmernost, bezkolíznost, náhodnost) závisí pouze na stavu vedy v oblasti kryptografie a kryptoanalýzy.

Cas od casu se štestí zvrtne a nekterá šifra, hašovací funkce nebo podpisové schéma padne díky odhalené slabine.

Prolomení nekterých kryptografických technik musí být prijímáno nikoli jako neduvera v kryptologii, ale jako pruvodní jev rozvoje poznání v této oblasti.

11

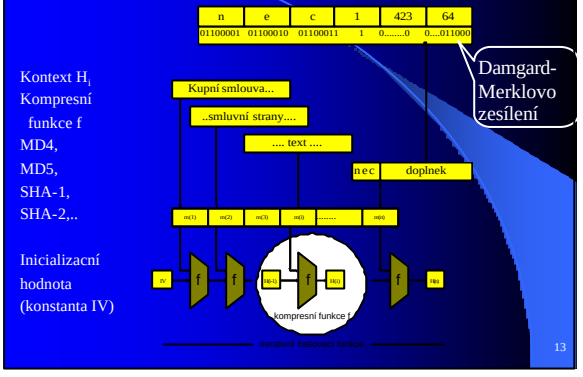
Dr. Wangová a kol. kolize našli…

Jakmile nekdo nalezne kolizi hašovací funkce, je nejbezpecnejší tuto funkci vymenit za jinou. To je **ideální rešení, které v praxi nekdy není možné**. Potom je nutné zkoumat, v jakých aplikacích je hašovací funkce použita a které vlastnosti celého systému jsou ohroženy.

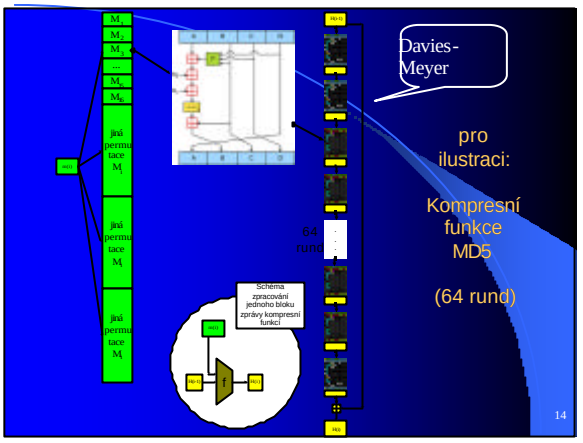
- Jak pracují hašovací funkce
- Jaké vlastnosti se využívají
- Jaké typy použití jsou kolizemi ohroženy

12

Princip moderních hašovacích funkcí



13



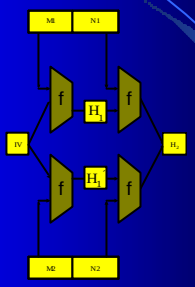
14

Na cem je založena jednosmernost ?

- Know-how z oblasti blokových šifer:
- Ze znalosti mnoha páru (OT, ŠT), tj. (vzor, obraz) nelze urcit klíc
  - Hašovaná zpráva necht vystupuje v roli klíce
- Zesložitení vícenásobným použitím zprávy
- Konstrukce Davies-Meyer posiluje jednosmernost a konfúzi dodatecným prictením vzoru

15

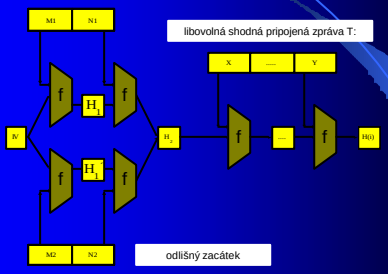
Cínský útok



- M1 512 bitu
- N1 512 bitu
- M2 512 bitu
- N2 512 bitu
- Úloha diferencních konstant

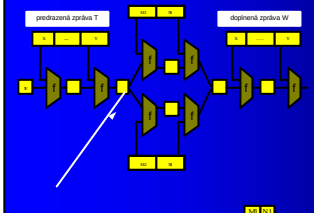
16

Rozširování, varianta 1

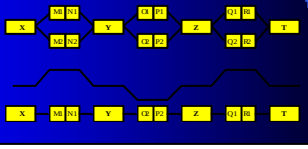


17

Rozširování, varianty 2,3,…



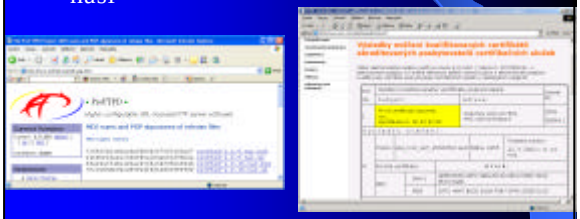
Mimovolný únik informací



18

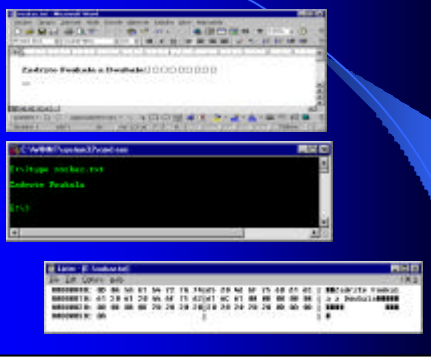
Možnosti zneužití kolizí

- Podvržení souboru (certifikátu) se stejnou haší



19

Zmena interpreta



20

Program zkoumání?

- Až bude zverejneno, jak to Cínani delají, být pripraveni ukázat kolize na smysluplných zprávách nebo souborech
- Využít
  - 3-bitových zmen
  - Zmenu interpreta
  - Datových zmen + zmen interpreta
- Eventuelne se dobrat toho, jak nalézají kolize

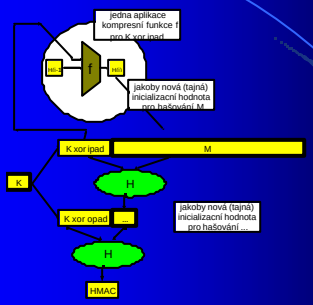
21

Co cínský útok neumí

- Neumí k dané zpráve vytvorit jinou, se stejnou haší (neumí kolizi druhého rádu)
- Umí vytvorit dva ruzné soubory (soucasne) se stejnou haší (kolizi prvního rádu)
- Nejsou (zatím) ohroženy minulé digitální podpisy, ale jsou ohroženy budoucí. Všude tam, kde útocník vytvárí soubor k digitálnímu podpisu, je nebezpecí vytvorení druhého falešného souboru.
- Jsou ohrožena další dve použití hašovacích funkcí? (HMAC a PRF)

22

HMAC – ohrožen není



HMAC-H(K, M) = H((K xor *opad*) || H((K xor *ipad*)|| M))

23

Odchylky od náhodného chování (pro PRNG, PRF)

- Známe H(M)
  - Umíme vytvorit H(M || doplnek || N) pro libovolné N
- Známe H(K || M)
  - Umíme vytvorit H(K || M || doplnek || N) pro libovolné N
- Pokud umíme konstruovat kolize se stejne dlouhými zprávami,
  - lze ze znalosti H(M || K) vytvorit (padelat) H($M_{kolidující}$ || K)

24

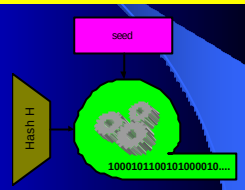
PRF/PRNG – pseudonáhodná funkce a generátor: zatím bezpecné, používat „obezretne“

PKCS#5, generování klíce z passwordu:

$T_1 = H(P \| S)$
$T_2 = H(T_1)$
…
$T_c = H(T_{c-1})$

šifrovací klíc
$DK = T_c<0..dkLen-1>$

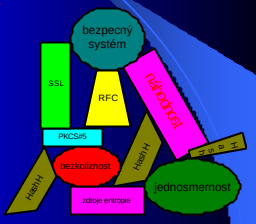
PKCS#1 v.2.1, pseudonáhodný generátor MGF1
Seed - vetšinou náhodné nastavení



H(seed || 0x00000000),
H(seed || 0x00000001),
H(seed || 0x00000002),
H(seed || 0x00000003),
....

Jak moc byla funkce narušena ?

- V produktech, protokolech apod. je použití mnohonásobné
- Prolomenou hašovací funkci lze buď generálne vyloucit nebo posuzovat individuálne (u nekterých autentizacních schémat kolize nevadí)
- Stavba bezpecnosti systému je vetšinou príliš krehká na to, aby nekterá vlastnost hašovací funkce mohla být oslabena
- Jednoznacne: kolidující funkce nesmí být použity pro digitální podpisy (nepopiratelnost)



Záver c.1

- MD4, MD5, SHA-0, RIPEMD a HAVAL-128 jsou prolomené
- SHA-1, SHA-256, SHA-384 nebo SHA-512 jsou bezpecné
- Doporucení NIST:
  - používat trídu funkcí SHA-2
  - do roku 2010 se predpokládá opuštení i SHA-1 a prechod na SHA-2

Literatura

- Výber z 18 položek z tištného dokumentu (viz [ARCHIV] http://cryptography.hyperlink.cz/2004/kolize_hash.htm):
  - [WFLY04] X. Wang, D. Feng, X. Lai, H. Yu, "Collisions for Hash Functions MD4, MD5, HAVAL-128 and RIPEMD", rump session, CRYPTO 2004, Cryptology ePrint Archive, Report 2004/199, http://eprint.iacr.org/2004/199
  - [R04] Rosa T.: Nepopiratelnost digitálních podpisu, Druhá vedecká a pedagogická konference ZMVS: Právní regulace informacní spolecnosti, Trebíc, zárí 2004,
  - Speciální stránka, venovaná tématu hašovacích funkcí http://cryptography.hyperlink.cz/2004/kolize_hash.htm
  - [ARCHIV] Archiv autora, obsahující clánky o kryptologii a bezpecnosti od r. 1993, http://cryptography.hyperlink.cz

Vloženo aktuálne

- Vloženo na základe objevné Kelsey-Schneierovy práce (publikované nekolik dní pred prednáškou)
- Práce ukazuje (viz záver c.2), že **musíme zmenit bezpecnostní pohled na hašovací funkce**, dnes bychom meli uvažovat, že jakékoliv jejich použití **neposkytuje vyšší bezpecnost, než $2^{n/2}$**, dríve pro kolizi 2.rádu, PRNG, PRF apod. to bylo 2^n

Dodatky: K-S, Joux, Wangová & kol.

- **Second Preimages on n-bit Hash Functions for Much Less than 2^n Work**
  John Kelsey and Bruce Schneier
- 15.11.2004, http://eprint.iacr.org/2004/304/
  - Ukázali **jak nalézt druhý vzor** u všech iterativních hašovacích funkcí s D-M zesílením
  - Pro zprávu o délce **2^k bloku**
  - Složitost **$k*2^{n/2+1} + 2^{n-k+1}$**
- Pro SHA1: **zpráva 2^{60} bajtu, 2^{106} operací**, dríve **2^{160}** operací
- Zatím nepraktické z duvodu príliš dlouhých zpráv, teoretický prínos vysoký

Ukážeme si

- Hledání kolizí u dlouhých zpráv bez D-M zesílení
- Nalezení (k, 2^k+k+1)-expandovatelné zprávy, složitost k*$2^{n/2+1}$
- Nalezení druhého vzoru pomocí expandovatelné zprávy, složitost k*$2^{n/2+1}$ + 2^{n-k+1}
- Nalezení (k, 2^k+k+1)-expandovatelné zprávy pomocí pevných bodu, složitost $2^{n/2+1}$
- Jouxovy multikolize, složitost k*$2^{n/2}$ místo $2^{n*(r-1)/r}$
  - Wangová a kol.: rychleji, více
  - Kelsey-Schneier: také

31

Například z toho plyne

- Ohromná multikolize: Možnost nalezení N zpráv o délce 2^{54} bloku, majících stejnou haš, se složitostí pouze 64*$2^{n/2}$, kde n je délka hašovacího kódu a N je neuveritelne velké:
- N = ?2^{54}? ?180143985094819847?
  ? ? ? ? ?
  ?31 ? ? 31 ?
- Není praktické z duvodu velmi dlouhých kolidujících zpráv

32

Dosti významný **Záver c.2**

- **n-bitová iterovaná hašovací funkce:**
- poskytuje zásadne jiné bezpecnostní vlastnosti než náhodné orákulum s n-bitovým výstupem
- nemuže zarucit odolnost proti nalezení druhého vzoru (pro dlouhé zprávy) na úrovni n bitu bezpecnosti, jak se dosud uvažovalo
- nezarucuje ruzné vlastnosti, pokud útocník má možnost provést rádove $2^{n/2}$ operací

33