Hash Functions Based on Block Ciphers

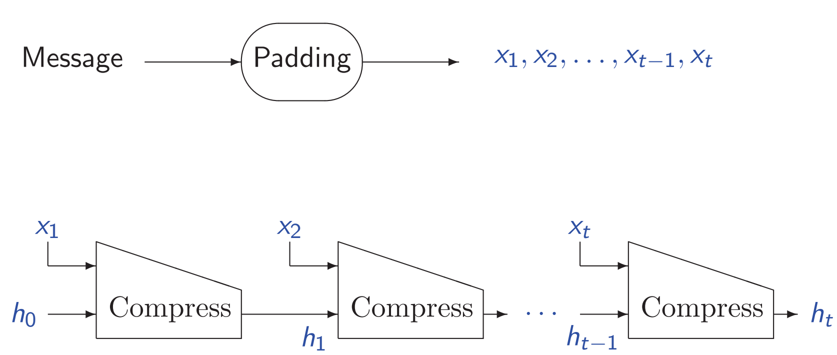
Lars R. Knudsen

April 21, 2008

Definition - hash function

$H : \{0,1\}^* \rightarrow \{0,1\}^n$, for fixed value of n

Iterated hash functions



Generic attacks

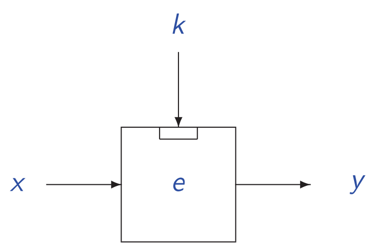
For $H : \{0,1\}^* \to \{0,1\}^n$ and $h : \{0,1\}^m \to \{0,1\}^n$, $m > n$

attack	rough complexities
collisions	$\sqrt{2^n} = 2^{n/2}$
2nd preimages	$k2^{n/2} + 2^{n-k}$ with 2^k blocks
preimage	2^n

Goal: generic attacks are best (known) attacks

Block cipher - family of permutations

- $e : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$, $m = \kappa + n > n$
- each κ-bit key specifies bijective mapping on n bits
- **must** hold for all x and k that $e_k^{-1}(e_k(x)) = x$.
- one-way function: given x and $e_k(x)$, hard to find k.



Product ciphers

- e most often some layers of substitutions and permutations
- example. SP-networks, ‘s’ for substitution, ‘p’ for permutation.

$$e_k(x) = s_k \circ p_k \circ s_k \circ p_k \circ \ldots \circ s_k \circ p_k \circ s_k(x)$$

- note that s_k and p_k must be invertible.

DES & AES

DES = Data Encryption Standard
AES = Advanced Encryption Standard

system	year	block size	key size
DES	1977	64	56
AES	2001	128	128, 192 or 256

DES: History

- Developed in early 70's by IBM using 17 man years
- Evaluation by National Security Agency (US)
- 1975: publication of proposed standard
- Public discussion (trapdoors, key size)
- 1977: publication of FIPS 46 (DES)
- 2008: most realistic key-recovery attack is exhaustive search

Results on DES

- $\forall\ p, k : c = \mathrm{DES}_k(p) \Longleftrightarrow \overline{c} = \mathrm{DES}_{\overline{k}}(\overline{p})$
- 4 weak keys: $\mathrm{DES}_k(\mathrm{DES}_k(p)) = p,\ \forall\ p$
- 6 pairs of semi-weak keys: $\mathrm{DES}_{k1} = \mathrm{DES}_{k2}^{-1}$
- Differential cryptanalysis (1991), 2^{47} chosen plaintexts
- Linear cryptanalysis (1993), 2^{43} known plaintexts
- Exhaustive search ($2^{56} \simeq 10^{17}$ operations).
  DES key found in 22 hours using special hardware and distributed software (1999)

AES - Advanced Encryption Standard

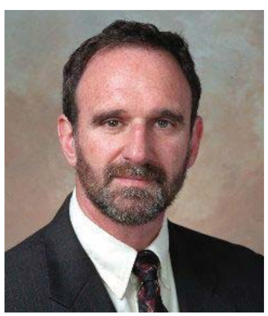
- US governmental encryption standard
- Open (world) competition announced January 97
- Keys: choice of 128-bit, 192-bit, and 256-bit keys
- Blocks: 128 bits
- October 2000: AES=Rijndael
- Standard: FIPS 197, November 2001

AES=Rijndael

- Designed by Joan Daemen and Vincent Rijmen
- Simple design, byte-oriented
- Operations: XOR and table lookup
- S-box based on $g(x) = x^{-1}$ in $GF(2^8)$
- 10, 12 or 14 iterations of the round transformation depending on length of key
- Focus on 128-bit key version with 10 iterations

In the beginning ....

A bit later .....



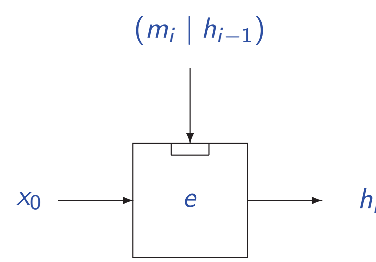
In the beginning there was ...

Diffie and Hellman, 1976. New directions in cryptography.

- Digital signatures .... for efficiency:
- "Let g be a one-way mapping from binary N-space to binary n-space...". "Take the N bit message m and operate on it with g to obtain the n bit vector m'."
- "It must be hard even given m to find a different inverse image of m'"
- "Finding such functions appears to offer little trouble"

Diffie-Hellman, $\kappa > n$

$e : \{0,1\}^\kappa \times \{0,1\}^n \to \{0,1\}^n$



- x_0 fixed block
- 2nd preimages hard if e secure against known-plaintext attack

Hash function using a block cipher

Why build on a block cipher?

- it's natural !
- use existing technology
- transfer security (trust?!) to hash construction
- schemes "slow" (partly due to key-schedules)
- weaknesses of block cipher not relevant for encryption

Hash rate

- Given hash function built from block cipher

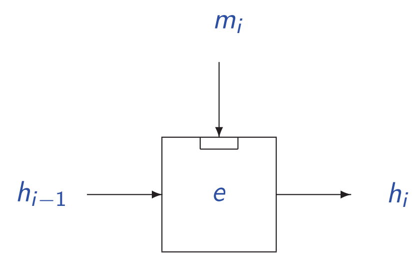
$$e : \{0,1\}^\kappa \times \{0,1\}^n \rightarrow \{0,1\}^n$$

- Hash rate defined as

$$\frac{\#\ n\text{-bit blocks hashed}}{\#\ \text{invocations of } e + \#\ \text{key-schedules}}$$

Rabin, 1978

$e : \{0,1\}^\kappa \times \{0,1\}^n \rightarrow \{0,1\}^n$



- rate $= (\kappa/n)/(1+1)$
- Yuval: collisions based on birthday paradox (79) (Merkle 79)
- Pre-images in approximately same time

Single block hash

- $e : \{0,1\}^\kappa \times \{0,1\}^n \rightarrow \{0,1\}^n$
- 12 secure ones (Preneel 93, Black et al 02), here three

$$h_i = e_{m_i}(h_{i-1}) \oplus h_{i-1} \qquad \text{Davies-Meyer}$$
$$h_i = e_{h_{i-1}}(m_i) \oplus m_i \qquad \text{Matyas-Meyer-Oseas}$$
$$h_i = e_{h_{i-1}}(m_i) \oplus m_i \oplus h_{i-1} \qquad \text{Preneel-Miyaguchi}$$

- Hash rates. About 1/(1+1) (1/2 for DES and AES)
- Collisions (birthday attack) in $2^{n/2}$ operations

MD4-family

- MD4, Rivest 1990
- MD5, Rivest 1991
- SHA-0, 1993
- SHA-1, 1994
- all hash functions of Davies-Meyer form
- "block ciphers" with feed-forward
- hash rates for Davies-Meyer can be (arbitrarily) high

Double block hash - based on block ciphers

- Based on $e : \{0,1\}^{\kappa} \times \{0,1\}^n \rightarrow \{0,1\}^n$
- Length of hash, $2n$ bits
- Aim: 2^n security level for collisions
  - Merkle, 1989
  - MDC-2, Brachtl, Coppersmith et al 1988/1990
  - PBGV, QG, LOKI-DBH, ..., 1990s
  - Hirose, Nandi, 2005
  - Shrimpton-Stamm 2007, Rogaway-Steinberger 2008

Ideal cipher model

- Let $B_{n,k}$ be all block ciphers with a k-bit key and n-bit blocks,

$$\{0,1\}^k \times \{0,1\}^n \rightarrow \{0,1\}^n$$

- There are $2^n! \approx 2^{n2^n}$ bijections on n bits
- It holds that

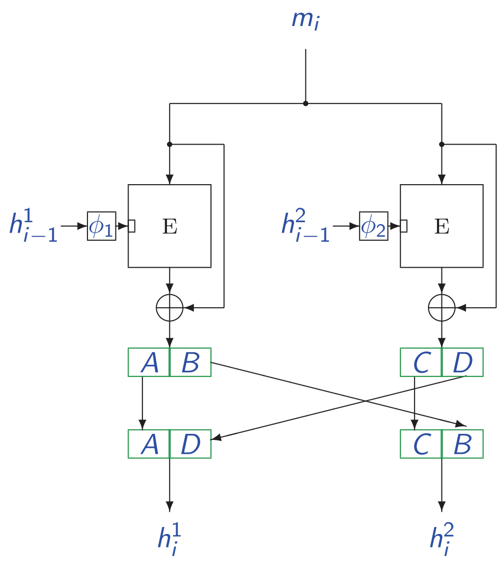
$$|B_{n,k}| = \binom{2^n!}{2^k}$$

- An ideal cipher is randomly selected from $B_{n,k}$

Merkle's double block schemes with DES (1989)

- "DES can be used to build a one-way hash function which is secure"
- if DES fails "it seems almost certain that some block cipher exist with the desirable properties"
- 128-bit hash function with proof of security in ideal cipher model
- collisions $\approx 2^{55}$, inconvenient block sizes, low hash rates
- "recent proposal from IBM looks very hopeful", but no proof..

MDC-2, IBM 1987



MDC-2

- Coppersmith, Meyer, Schilling et al, IBM, patent filed 1987
- designed for DES but can be used with any block cipher
- mapping from text to key: $\phi_1(\cdot), \phi_2(\cdot) : \{0,1\}^{64} \rightarrow \{0,1\}^{56}$
- $\phi_1(x)$, $\phi_2(y)$ never weak DES keys for any x, y
- hash rate 1/(2+2) (1/4 for DES and AES)
- MDC-4: variant using four encryptions per block

MCD-2 used with DES and AES

(Best known attacks)

	DES	AES
Preimage attack	2^{83}	2^{192}
2nd preimage attack	2^{83}	2^{192}
Collision attack	2^{55}	2^{128}
Hash rate	1/4	1/4

- For use with AES, "proof" that collision requries $> 2^{75}$ operations (Steinberger 2007)

Abreast-DM & Tandem-DM - Lai, Massey 1990

- $e : \{0,1\}^\kappa \times \{0,1\}^n \rightarrow \{0,1\}^n$, $\kappa > n$ $\quad f(x,y) = e_x(y) \oplus y$
- Abreast-DM $\begin{cases} h^1_i = f(h^2_{i-1} \parallel m_i\,,\; h^1_{i-1}) \\ h^2_i = f(m_i \parallel h^1_{i-1}\,,\; \overline{h^2_{i-1}}) \end{cases}$
- Tandem-DM $\begin{cases} h^1_i = f(h^2_{i-1} \parallel m_i\,,\; h^1_{i-1}) \\ h^2_i = f(m_i \parallel (h^1_i \oplus h^1_{i-1})\,,\; h^2_{i-1}) \end{cases}$
- IDEA, (64-bit block, 128-bit key), hash rate 1/4, conjectured security level for collisions 2^{64}
- AES-256, (128-bit block, 256-bit key), hash rate 1/4, conjectured security level for collisions 2^{128}

Knudsen-Preneel 1996 $f_i(x, y) = e_x(y) \oplus y$

Compress: $(h^1_{i-1}, \ldots, h^5_{i-1}, m_i) \rightarrow (h^1_i, \ldots, h^5_i)$

$$
\begin{aligned}
h^1_i &= f_1(h^1_{i-1}\,,\; h^2_{i-1}) \\
h^2_i &= f_2(h^3_{i-1}\,,\; h^4_{i-1}) \\
h^3_i &= f_3(h^5_{i-1}\,,\; m_i) \\
h^4_i &= f_4(h^1_{i-1} \oplus h^3_{i-1} \oplus h^5_{i-1}\,,\; h^2_{i-1} \oplus h^4_{i-1} \oplus m_i) \\
h^5_i &= f_5(h^1_{i-1} \oplus h^3_{i-1} \oplus h^4_{i-1} \oplus m_i\,,\; h^2_{i-1} \oplus h^3_{i-1} \oplus h^5_{i-1} \oplus m_i)
\end{aligned}
$$

Constructed from $[5, 3, 3]$ code over $GF(2^2)$: rate $1/(5{+}5)$
Claimed security against collision attacks is 2^n

Knudsen-Preneel, more examples

Better rates using codes over larger fields

$GF(2^2)$		$GF(2^4)$		Collision
Code	Rate	Code	Rate	
$[5, 3, 3]$	$1/(5+5)$	$[6, 4, 3]$	$2/(6+6)$	$\simeq 2^n$
$[8, 5, 3]$	$2/(8+8)$	$[8, 6, 3]$	$4/(8+8)$	$\simeq 2^n$
$[12, 9, 3]$	$6/(12+12)$	$[12, 10, 3]$	$8/(12+12)$	$\simeq 2^n$

AES-128, rate 1/3, conjectured security level for collisions 2^{128}

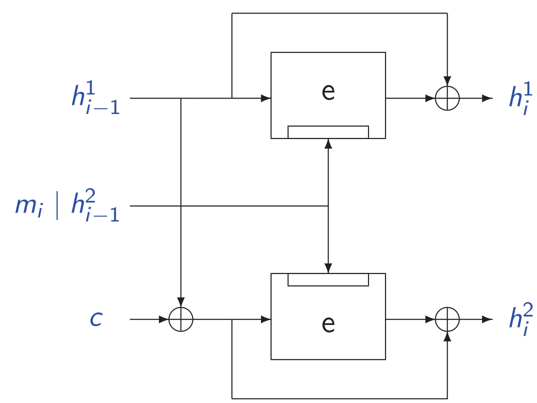
Hirose's double block mode 2006

Based on work by Nandi, 2005
$e : \{0,1\}^\kappa \times \{0,1\}^n \rightarrow \{0,1\}^n$, $\kappa > n$, c nonzero constant

$$
\begin{aligned}
h^1_i &= e_{h^2_{i-1} \,|\, m_i}\,(h^1_{i-1}) \oplus h^1_{i-1} \\
h^2_i &= e_{h^2_{i-1} \,|\, m_i}\,(h^1_{i-1} \oplus c) \oplus h^1_{i-1} \oplus c
\end{aligned}
$$

- Collision requires 2^n operations assuming $e(\cdot,\cdot)$ is ideal cipher
- AES-256, hash rate 1/3, security level 2^{128} for collisions

Hirose's double block mode, figure



Ideal cipher model ?

- proofs in model give protection against generic attacks
- no real-life cipher is an ideal cipher; "nearly ideal" cipher can be strong for encryption but very weak when used for hashing
- attacker in control of key can invest time in finding key(s) with certain properties

Known-key distinguishers - Knudsen-Rijmen 2007

- Block cipher cryptanalysis with applications to hash functions
- With a given (random) key, produce set of texts with "non-random" behaviour
- Most short-cut attacks on block ciphers exploit statistical properties of plain- and ciphertexts in (reduced) cipher
- If such properties cannot be found given the key, it seems unlikely that they can be found when **not** given the key

Hash based on fixed bijections

- Preneel, 1992
- Knudsen, 2005: SMASH
- Black et al, 2005: Collision-resistant iterated hash functions based on one bijective mapping do not exist in information-theoretic setting
- Shrimpton-Stamm, 2007, construction with three bijections, rate 1/3
- Rogaway-Steinberger, 2008
  - at least three bijections necessary and sufficient
  - at least five bijections needed in double-block hash mode

Concluding remarks

- Renaissance of block cipher based proposal?
- Seems like we can get good hashing with speed 1/3 of AES ?
- Block cipher based proposals likely not to be selected for SHA-3 ?!