Forgery and Partial Key Recovery attacks on HMAC and NMAC using Hash Collisions

2nd NIST Hash Function Workshop

Scott Contini
Macquarie University
Yiqun Lisa Yin
Independent Security Consultant

To appear in Asiacrypt'06

Outline

- Background and motivation
- Summary of results
  - Various attacks on HMAC/NMAC
    - Using special collisions of underlying hash function
- Closer look — partial key-recovery attacks
  - How to recover *entire* inner key
- Practical implications

(Not included in CD-Rom)

- New observations on 2nd preimage resistance (eSPR & rSPR)
  - MD5, reduced SHA-1

HMAC and NMAC

- *Hash-based* message authentication code (MAC)
  - Proposed by Bellare, Canetti, Krawczyk in 1996

- HMAC has been widely implemented in practice
  - Standards: SSL/TLS, SSH, IPsec, etc.
  - Usages: MAC, PRF, random oracle, etc.

- Construction
  - NMAC: $\mathrm{NMAC}_{(k1,\ k2)}(m) = F_{k1}(F_{k2}(m))$

  - HMAC: $(k1, k2) = \mathrm{KDF}(k)$
    $\mathrm{HMAC}_k(m) = \mathrm{NMAC}_{(k1,k2)}(m)$

    - $F_k(m) = F(k, m)$ is a hash function with IV = secret key k

Related attacks on MDx

- We studied existing attacks on MDx, especially
  - Pseudo-collision attack on MD5 [DB 93]
  - Collision attack on SHA-0 [CJ 98]
  - Collision attack on reduced SHA-1 [BCJCJL 05]
  - 2nd pre-image attack on MD4 [YWZW 05]

- *Differential paths* in above attacks can be used to construct *distinguishing attacks* on f_k
  - For MD4, SHA-0, reduced SHA-1, f_k is not a PRF
  - For MD5, f_k is not a PRF against *related-key attacks*

Summary of our results

- Attacks on HMAC/NMAC-MDx
  - Distinguishing attacks
  - Forgery attacks
  - Partial key-recovery attacks
    - Can recover *entire k2 (128 or 160 bits)*

$F_{k1}\,(F_{k2}\,(m))$

- Complexity (estimated # MAC queries)
  - NMAC-MD5 [related-key attacks] : 2^{47} queries
  - HMAC/NMAC-MD4: 2^{58} queries
  - HMAC/NMAC-SHA0: 2^{84} queries
  - reduced HMAC/NMAC-SHA1: ~ 2^{40} queries
    - inner function is reduced to 34 rounds

- Biham and Yin (8/24/06, *not included in CD-Rom*)
  - 40-round NMAC-SHA1 [related-key attacks] : ~2^{55} queries
  - 40-round HMAC-SHA1: ~ 2^{110} queries

Summary of our results

- Attacks on HMAC/NMAC-MDx
  - Distinguishing attacks
  - Forgery attacks
  - Partial key-recovery attacks
    - Can recover *entire k2 (128 or 160 bits)*

- Complexity (estimated # MAC queries)
  - NMAC-MD5 [related-key attacks] : 2^{47} queries
  - HMAC/NMAC-MD4: 2^{58} queries
  - HMAC/NMAC-SHA0: 2^{84} queries
  - reduced HMAC/NMAC-SHA1: ~ 2^{40} queries
    - inner function is reduced to 34 rounds

Trade-offs:
#queries: 2^t
success prob: 2^{t-q}
(1 < t < q)

Kim, Biryukov, Preneel, Hong [SCN'06]
- Independent work on distinguishing and forgery attacks

Partial key-recovery attacks on NMAC-MD5
(related-key setting)

- High-level steps
  - Generate random messages and query the two NMAC oracles until obtaining a collision
    - $\mathrm{NMAC}_{(k1,\,k2)}(m) = \mathrm{NMAC}_{(k1,\,k2')}(m)$
  - Modify certain bits of *m* to create a set of new messages
    - Based on *new message modification techniques*
  - Check whether the set of new messages yield a new collision
    - Each yes/no answer roughly reveals *one bit* of internal state
  - Step through the computation of $F_{k2}(m)$ backwards to obtain the initial state – the inner key *k2*

Danger of hash collisions

- It is *not* surprising that hash collisions are useful for key recovery
  - Several earlier attacks on MACs use collisions

- Reason 1:
  - Collision path contains useful information about the internal hash computation $F_{k2}(m)$, and hence the initial secret key $k2$
- Reason 2:
  - Outer function F_{k1} in HMAC/NMAC does not hide collisions of inner function F_{k2}

Implications of our results

- HMAC-MD4
  - Should no longer be used in practice
- Our results complement designers' analysis
  - Designers show that HMAC/NMAC is secure *assuming* f_k is a PRF
  - We show that attacks are possible if f_k is *not* a PRF

- HMAC-MD5, HMAC-SHA1
  - No immediate practical threats

- Proper differential paths are crucial
  - Collision attacks, 2nd preimage attacks, and attacks on HMAC require paths with *different* properties
  - *Automated* method is a promising way to search for suitable paths

2nd preimage resistance (SPR)

- Compression function *f(c,m)*
- Goal of attacker S:
  - present *(c,m)* and *(c',m')* s.t.
    - *(c,m) ≠ (c',m')*
    - *f(c,m) = f(c',m')*

Variants of CR & SPR

	Attacker is given	Attacker picks
pseodo-CR		*c, m, c', m'*
CR	fixed *c=c'*	*m, m'*
SPR	fixed *c=c'* *random m*	*m'*

2nd preimage resistance (SPR)

- Compression function *f(c,m)*
- Goal of attacker S:
  - present *(c,m)* and *(c',m')* s.t.
    - *(c,m) ≠ (c',m')*
    - *f(c,m) = f(c',m')*
- Sort of known
  - MD4, SHA-0 are not eSPR, rSPR
    - Since they are not SPR
- New observations
  - MD5 is not eSPR, rSPR
    - workload O(1)
    - success prob = 2^{-48}
  - 40-round SHA-1 is not eSPR, rSPR, SPR [Biham, Yin]

Variants of CR & SPR

	Attacker is given	Attacker picks
pseodo-CR		*c, m, c', m'*
CR	fixed *c=c'*	*m, m'*
eSRP	*"somewhat" random c* *random m*	*c', m'*
rSPR	*random c, m*	*c', m'*
SPR	fixed *c=c'* *random m*	*m'*

Thank you very much !

Publication info:

To appear in Asiacrypt'06

Authors' contact info:

scott_contini [at] yahoo.com

yiqun [at] alum.mit.edu