Cryptographic Hash Functions

Recent Results on Cryptanalysis and their Implications on System Security

Rüdiger Weis [1] Stefan Lucks [2]

Technical University of Applied Sciences Berlin

University of Mannheim

May 19, 2006

Chinese ante portas

- X. Wang, D. Feng, X. Lai, H. Yu: Cryptanalysis of the hash functions MD4 and RIPEMD. Eurocrypt 2005.
- X. Wang, H. Yu: How to break MD5 and other hash functions. Eurocrypt 2005.
- X. Wang, H. Yu, Y.L. Yin: Efficient collision search attacks on SHA0. Crypto 2005.
- X. Wang, Y.L. Yin, H. Yu: Finding collisions in the full SHA1. Crypto 2005.

Collisions and Preimages

Collision attack: Find two messages $M \neq M'$ with $H(M) = H(M')$

Preimage attack: Given a random value $Y \in \{0,1\}^n$, find a message M with $H(M) = Y$.

2nd preimage attack: Given a message M, find a message $M' \neq M$ with $H(M) = H(M')$.

K-collision attack for $K \geq 2$: Find K different messages M^i, with $H(M^1) = \cdots = H(M^K)$.

Merkle-Damgård Design

- I. Damgård. A design principle for hash functions. Crypto 89, LNCS 435, pp. 416–427.
- R. Merkle. One-way hash functions and DES. Crypto 89, LNCS 435, pp. 428–446.

- With a fixed-size compression function

$$C : \{0,1\}^n \cdot \{0,1\}^m \rightarrow \{0,1\}^n$$

- We define a hash function

$$H : \{0,1\}^* \rightarrow \{0,1\}^n$$

using the compression function and chaining.

Merkle-Damgård Design

Given a fixed *initial value* H_0 and a message $M \in \{0,1\}^*$, the Merkle-Damgård (MD) hash $H(M)$ is computed as follows:

- Expand M to $(M_1, \ldots, M_L) \in \{0,1\}^{m \cdot L}$.

  MD strengthening: The last block M_L takes the length $|M|$ in bits.

- For i in $1, \ldots, L$: compute

$$H_i := C(H_{i-1}, M_i).$$

- Finally: set

$$H(M) = H_L.$$

A structural problem

- Generate colliding 512-bit blocks $B_1 \neq B_2$ with

$$SHA-1(B_1) = SHA-1(B_2)$$

- Generate programm P_1 and P_2

$$P_1 := B_1||C \text{ and } P_2 := B_2||C$$

Note that we have

$$SHA-1(P_1) = SHA-1(P_2)$$

because of the iterative structure of SHA-1.

A Generic Attack

Generate a program twin $JANUS_j$

$$B_j||C$$

$JANUS_j$ is

- `harmless` if B_1 is used
- `evil` if B_2 is used

Nightmare Closed Source

If we can not check the source code like in MS Windows or Apple Aqua even trivial attacks are possible.

```
# R. Weis,  23.12.05
def BeNice():print("Think different.")
def BeEvil():print("Denounce user.")
blockdiffplace=42
evilbytevalue=23
file=open('bootimg')
s=file.read()
if s[blockdiffplace]==evilbytevalue:
   BeEvil()
else:
   BeNice()
```

Related Work

- D. Kaminski: MD5 to be considered harmful someday, CCC 2004. http://www.doxpara.com/md5_someday.pdf
- O. Mikle: Practical Attacks on Digital Signatures using MD5 message digest, http://eprint.iacr.org/2004/356
- A. Lenstra, B. de Weger: On the possibility of constructing meaningful hash collisions for public keys http://www.win.tue.nl/~bdeweger/CollidingCertificates/
- S. Lucks, M. Daum, The Story of Alice and her Boss http://www.cits.rub.de/MD5Collisions/
- M. Gebhardt, G. Illies, W. Schindler: Hash Collisions for Special File Formats, to appear in Sicherheit 2006 (BSI)

TCG and SHA-1

TCG Presentation, RSA 2005

- "SHA-1 Computing Engine
  - Multiple uses: integrity, autohization, PCR extension, etc."

Why are SHA-1 collissions so harmfull?

TCG uses SHA-1 for allmost all operations.
The iterative structure of SHA1 makes attacks practical.

- Integrity messaments using SHA-1 are compromized.
- Digital signatures using SHA-1 are compromized.
- PKIs using SHA-1 are compromized.

We have warned the TCG about SHA-1
e.g. in our CCCongress talks every year since 2002.

Attacking TCG Booting

- Generate two colliding blocks B_1 and B_2.
- Generate two boot programs

$$P_1 := B_1||C$$

and

$$P_2 = B_2||C$$

- The TPM will generate the same SHA-1 based checksum for P_1 and P_2.
- An attacker can substitute P_1 by P_2 with plessing of the TPM.
- Game over.

Readable Source

- Readable source is a *conditio sine qua non* for security archtectures.
- Even with readable code excluding hash collission based attacks seems to be very difficult.

Hiding in Readable Source

- It is a old hacker game to hide functionality.
- Trivial Examples:
  - Introduce Buffer Overflows
  - 0 pointer dereference
  - ....

and we can even hide collissions in X509 certificates...

X509 Certificates

http://www.win.tue.nl/~bdeweger/CollidingCertificates/

Colliding X.509 Certificates based on SHA1-collisions

"We would like to announce a pair of valid X.509 certificates, based on the SHA1 hash-function, that have identical signatures.
However we are not yet able to do so. The reason is that generating collisions for the SHA1 hash-function still takes a prohibitively large amount of time.
However, as soon as somebody is able to produce in practice collisions for the SHA1 compression function with prescribed IV, we can easily come up with colliding certificates based on that."

Hardware

http://www.schneier.com/blog/archives/2005/02/cryptanalysis_o.html

> *"In 1999, a group of cryptographers built a DES cracker. It was able to perform 2^{56} DES operations in 56 hours. The machine cost $250K to build, although duplicates could be made in the $50K-$75K range. Extrapolating that machine using Moore's Law, a similar machine built today could perform 2^{60} calculations in 56 hours, and 2^{69} calculations in three and a quarter years. Or, a machine that cost $25M-$38M could do 2^{69} calculations in the same 56 hours."*

Bruce Schneier

Falling

Meanwile it is not 2^{69} but 2^{63} and falling...
$2^6 = 64$ times cheaper....

Beyond Collissons

- Problems with tho whole family
- More rounds - less secure...
- Secound preimage
- Multicollissions

More Rounds - less secure..

Eli Biham, August 9., 2004 SAC
Joint work with Rafi Chen.

- 'The strength of reduced/extended SHA-0 is not monotonous with the number of rounds.'
- '82 round SHA-0 is much less secure than 80-round SHA-0'
- Near-Collisions

Second Preimages

- John Kelsey, J., Schneier, B.
- Cryptology ePrint Archive: Report 2004/304
- Second Preimages on n-bit Hash Functions for Much Less than 2^n Work

Also SHA-1

- Second preimage attack on all n-bit iterated hash functions with Damgard-Merkle strengthening and n-bit itermediate states
- A second preimage can be found for a 2^k-message-block message with a work about

$$k \cdot 2^{n/2+1} + 2^{n-k+1}$$

- Using SHA1 as an example, our attack can find a second preimage in 2^{106} work, rather than the previously expected 2^{160} work.

Multikollissions

Crypto 2004, Antoine Joux

- 2^k-Collisions for a MD hash H in time

$$O(k \cdot 2^{n/2}),$$

instead of

$$\Omega\left(2^{n \cdot \frac{2^k-1}{2^k}}\right)$$

Growing Collissions

- For i in $1 \ldots, k$: find a local collision $M_i^0 \neq M_i^1$ with

$$H_i = C(H_{i-1}, M_i^0) = C(H_{i-1}, M_i^1).$$

- All the 2^k messages
  - $(M_1^0, \ldots, M_k^0)$
  - $(M_1^0, \ldots, M_{k-1}^0, M_k^1)$,
  - $\ldots$
  - $(M_1^1, \ldots, M_k^1)$

  hash to the same value H_k.

MD5 still in heavy use

MD5 still in heavy use

- Debian Pakages
- RPM Pakages
- Open BSD
- ...

Open BSD

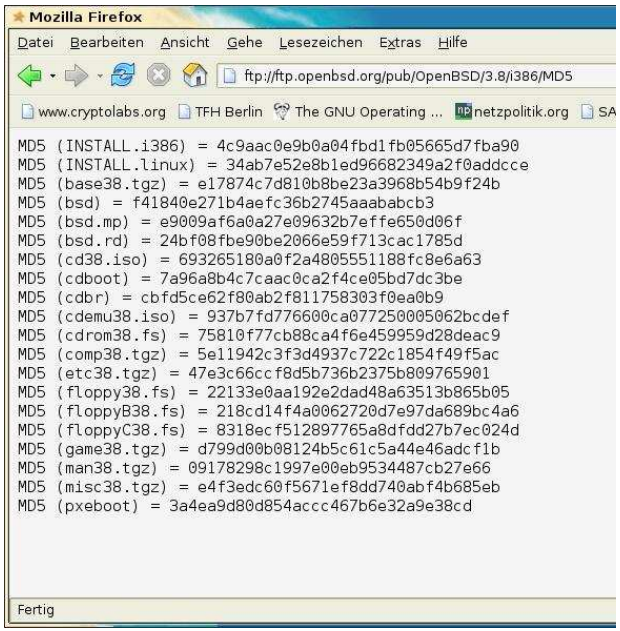

Figure: Open BSD checksums

Debian Pakets

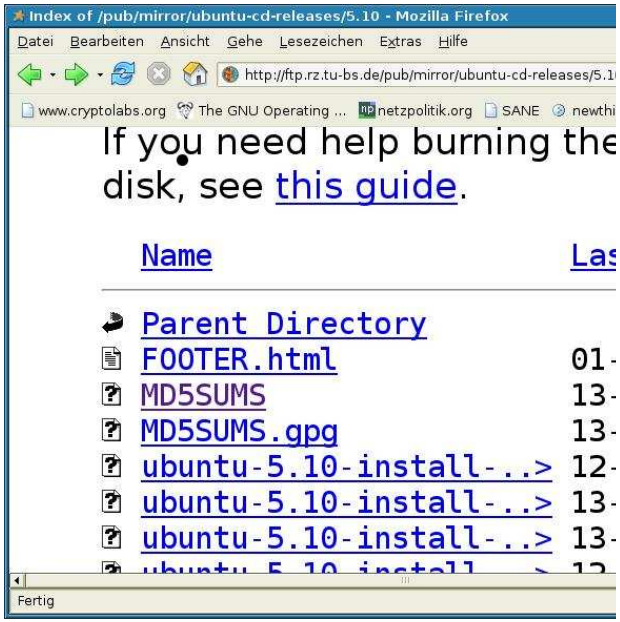

Figure: MD5 signed with gpg using SHA-1

ISO checksums

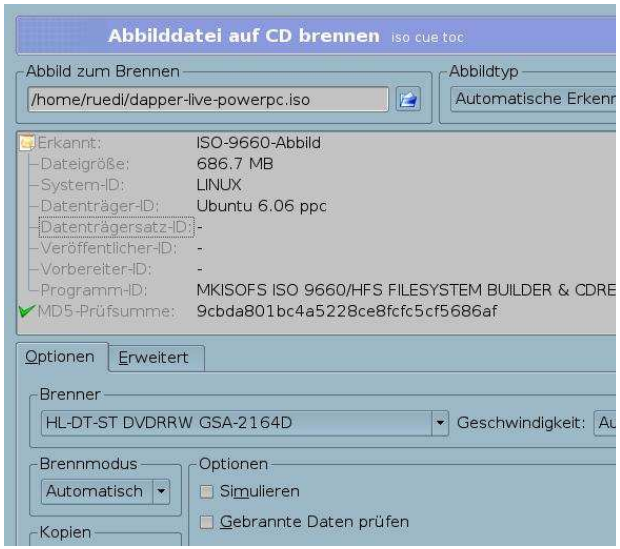

Figure: MD5 ISO checksums

Dirty Hot Fixes

Dirty Hot Fixes

- SHA-256
- Roboust Cryptography
- Casscading

SHA-2

http://csrc.nist.gov/CryptoToolkit/tkhash.html

FIPS 180-2, Secure Hash Standard (SHS), August 2002. smallskip On August 26, 2002, NIST announced the approval of FIPS 180-2, Secure Hash Standard, which contains the specifications for the Secure Hash Algorithms (SHA-1, SHA-256, SHA-384, and SHA-512) with several examples.

Cryptophone

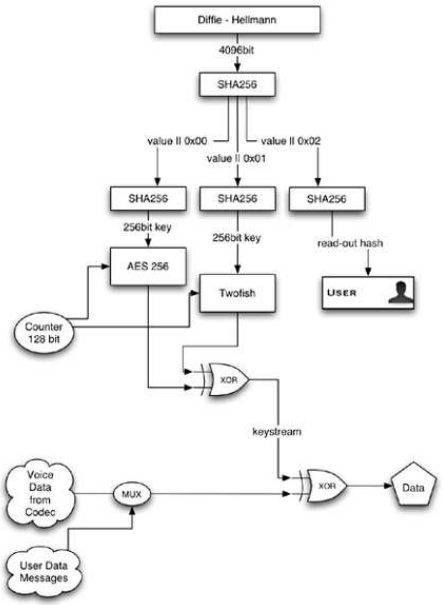

Figure: Cryptophone

SHA-2 analysis

"However we show that slightly simplified versions of the hash functions are surprisingly weak: whenever symmetric constants and initialization values are used throughout the computations, and modular additions are replaced by exclusive or operations, symmetric messages hash to symmetric digests."

Henri Gilbert, H., Handschuh, H.,
"Security analysis of SHA-256 and sisters" , SAC 2003.

Corrective Patterns

- P. Hawkes, M. Paddon, G.G. Rose
- On Corrective Patterns for the SHA-2 Family
- http://eprint.iacr.org/2004/207

Resistance against Chaboud-Jous Attack?

The Wide-Pipe Hash

- Stefan Lucks, Cryptology ePrint Archive: Report 2004/253.
- *A Design Principle for Iterated Hash Functions*
- A modified Merkle-Damgård design for iterated n-bit hash functions, *increasing the internal state to more than n bit*.

Revoreved NSA design for SHA-2 'by accident'

Cascading and XOR

- SHA-1||SHA-512||Tiger||Whirlpool
- XOR different designs, R. Weis, PhD Thesis 2000.

Not as secure as expected, but the best we have today.

Conclusions

Conclusions

- SHA should be replased today.
- MD5 should be replaced yesterday resp. NOW!!!
- Don't use broken Hash functions!
- Don't hard wire crypto devices without sufficent security margins!
- We need new Hash!

Knowing very little

"MD5, SHA-1 and the like are design by twidle."
Honestly, Whirlpool is no less twiddle – although I prefer muddle – than MD5 or SHA. And the fact that SHA is NSA muddle should count for something.
Honestly, we in the cryptography community know very little about hash functions.

Posted by: Bruce Schneier at March 10, 2005 05:20 PM

©opyleft

©opyleft

- Written and performed with Free Software.
- © Rüdiger Weis & Stefan Lucks, Delft 2006
- under the GNU Free Documentation License.

Acknowledgements

Andy Tanenbaum, Jan Mark Wams, Richard Stallman