Eve's SHA3 candidate: malicious hashing

Jean-Philippe Aumasson

NAGRAVISION
KUDELSKI GROUP

Background

Definitions

Strategies

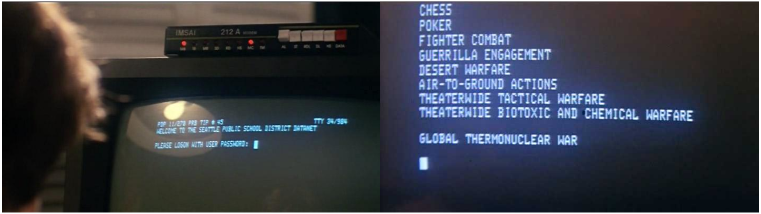
BLAKE tweak

Terminology remark: in this talk/paper

- Trapdoor: known to exist, difficult to find (RSA)
- Backdoor: not known to exist (NSA)

(Maybe not the best illustration)

Special credentials in *Wargames*' WOPR supercomputer



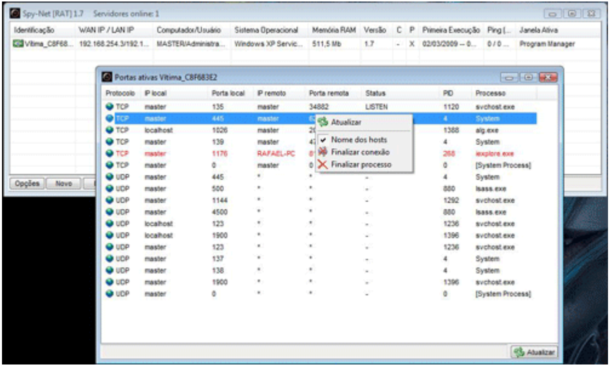
Linux 2.6 modified `kernel/exit.c`

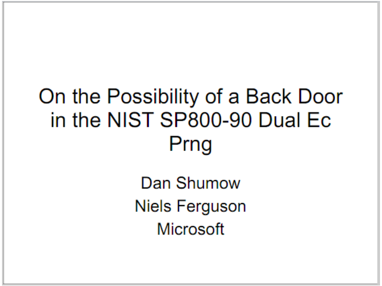
```
--- GOOD 2003-11-05 13:46:44.000000000 -0800
+++ BAD 2003-11-05 13:46:53.000000000 -0800
@@ -1111,6 +1111,8 @@
        schedule();
        goto repeat;
    }
+   if ((options == (__WCLONE|__WALL)) && (current->uid = 0))
+           retval = -EINVAL;
    retval = -ECHILD;

end_wait4:
```

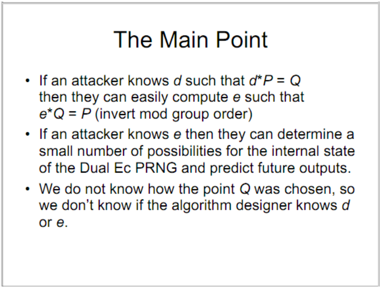
Thompson’s malicious gcc

```
compile(s)
char *s;
{
        if(match(s, "pattern")) {
                compile("bug");
                return;
        }
        ...
}
```

Trojans, RAT’s, rootkits, etc. (system backdoors)







Implementation backdoors:

- Hardware trojans, bug attacks
- Pure SW backdoor (cf. Wagner/Biondi's RC4)
- Weak RNG/entropy attacks (PGP. . .)

Sabotaged/weak crypto: Clipper chip, A5/2, etc.

Failed attempt based on weak S-boxes. . .

A Family of Trapdoor Ciphers

Vincent Rijmen* Bart Preneel**

Katholieke Universiteit Leuven,
Department Electrical Engineering-ESAT/COSIC
K. Mercierlaan 94, B-3001 Heverlee, Belgium
{vincent.rijmen,bart.preneel}@kuleuven.ac.be

Abstract. This paper presents several methods to construct trapdoor block ciphers. A trapdoor cipher contains some hidden structure; knowledge of this structure allows an attacker to obtain information on the key or to decrypt certain ciphertexts. Without this trapdoor information the block cipher seems to be secure. It is demonstrated that for certain block ciphers, trapdoors can be built-in that make the cipher susceptible to linear cryptanalysis; however, finding these trapdoors can be made very hard, even if one knows the general form of the trapdoor. In principle such a trapdoor can be used to design a public key encryption scheme based on a conventional block cipher.

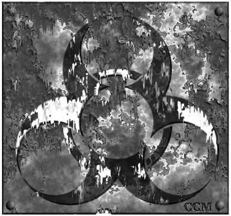
Cryptanalysis of Rijmen-Preneel Trapdoor Ciphers

Hongjun Wu* Feng Bao** Robert H. Deng** Qin-Zhong Ye*

*Department of Electrical Engineering
National University of Singapore
Singapore 119260

**Information Security Group
Kent Ridge Digital Labs
Singapore 119613

Abstract. Rijmen and Preneel recently proposed for the first time a family of trapdoor block ciphers [8]. In this family of ciphers, a trapdoor is hidden in S-boxes and is claimed to be undetectable in [8] for properly chosen parameters. Given the trapdoor, the secret key (used for encryption and decryption) can be recovered easily by applying Matsui's linear cryptanalysis [6].
In this paper, we break this family of trapdoor block ciphers by developing an attack on the S-boxes. We show how to find the trapdoor in the S-boxes and demonstrate that it is impossible to adjust the parameters of the S-boxes such that detecting the trapdoor is difficult meanwhile finding the secret key by trapdoor information is easy.



Young/Yung malicious blackbox cipher:

Exploit Huffman-compressible texts
to leak key bits in ciphertexts

Plus other “cryptovirology” schemes

Previous attempts of malicious block ciphers, stream ciphers, PRNG; what about malicious hash functions?

First thoughts:

- Goal not (only) key recovery: room for new techniques
- Can affect several schemes where the hash is used
- Different from trapdoor hash functions (VSH etc.)

Two approaches:

- "A priori": new design from scratch
- "a posteriori": modify existing hash

Many real-world applications…

Context

Eve designs a proprietary hash to integrate in PONY's GameStation 3 game console. The hash is used to sign boot code and executables. Digest are processed with a secure ECDSA implementation.

Backdoor

Eve (and only her) can compute meaningful second preimages

Exploit

Custom OS, piracy, homebrew software, blackmail

Hash Name	Principal Submitter	Best Attack on Main NIST Requirements	Best Attack on other Hash Requirements
BLAKE	Jean-Philippe Aumasson		
EvilHash	Eve		
Grøstl	Lars R. Knudsen		
JH	Hongjun Wu	preimage	
Keccak	The Keccak Team		
Skein	Bruce Schneier		

Context
Eve submits her EvilHash to SHA3 and wins the competition

Backdoor
Eve knows two colliding messages (and not more)

Exploit
She sells, or anonymously publishes the collision for fun

Malicious hash function = adversary = pair of algorithms:

- **Malicious generator**: returns hash H and backdoor b
- **Exploit algorithm**: given b and additional info, “breaks” H

Two types of backdoors (i.e. adversaries):

- **Static**: deterministic exploit algorithm
- **Dynamic**: probabilistic exploit, e.g. based on challenge

Good guys Alice and Bob will be Eve’s adversaries…

Adversaries breaking standard security notions:

- Static collision adversary
- Dynamic (second) preimage adversary
- Dynamic key-recovery adversary

Static preimage adversary

- Find preimage(s) of some low-entropy digest
- E.g. all-zero, repeated-byte, ASCII string, etc.
- Practically relevant, but no theor'y sound

Static distinguisher

- Finds N inputs satisfying some relation
- E.g. multicollision, linear dependencies
- Relation needs be "convincing"

Security goals:

- Undetectability
- Undiscoverability

Undetectability:

- Exploit algorithm difficult to describe
- Avoid reasonable suspicion

Input: “canonical” description of the algorithm

In practice, obfuscation may be used… related problem of white-box ciphers (a.k.a. “symmetric public-key schemes”)

Backdoor-in-the-middle

- Connect input and outputs within a permutation
- Applies to blockcipher-based compression, sponges

Simple example:

- Split (keyed) permutation in three parts

$$\Pi = \Pi_2 \circ \Pi_1 \circ \Pi_0$$

- For some chosen input(s) and output(s), modify/create Π_1 to connect the two parts

Malicious finalization

- Exploit entropy loss from two or more legit final states
- Either hash finalization (as in SIMD, Grøstl) or local (BLAKE, Hamsi)

Simple example:

- Collect final states of 2 chosen messages
- Choose a shrinking linear map such that
  - the two states collide
  - the equations look unsuspicious

Weak mode trigger

- Enter a weak internal state, then exploit it
- Can be a fixed-point, the IV of a sponge (2nd preimages)

Simple example:

- Find a fixed-point $E_m(h) \oplus h = h$ and set h as IV
- Use the backdoor m to find second preimages
- Works for wide-pipes, HAIFA

Freedom degrees from

- Operators (e.g. choose between $+$, $-$, $\oplus$)
- Ordering (e.g. $x + (y \oplus z)$ vs. $(x + y) \oplus z$)
- Constants (rotation distances, additive constants, #rounds)

Notion of **neutral structure** = algorithm composed of wildcard characters with high enough total entropy, e.g.

$$
\begin{aligned}
s_0 &= s_1 + x \\
s_1 &= s_2 \oplus y \\
s_2 &= s_3 + z \\
s_3 &= s_0 \star ((s_1 \bullet s_2) \diamond (s_3 \ggg n))
\end{aligned}
$$

where x, y, z are chosen from a set of C constants; $\star$, $\bullet$ and $\diamond$ are one of B binary operators; n is in $\{1, \ldots, 31\}$

Total entropy $3\log_2 C + 3\log_2 B + \log_2(31)$

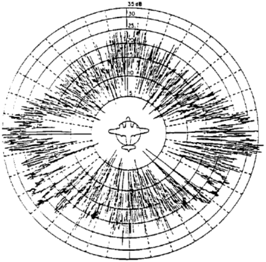
Stealth strategies

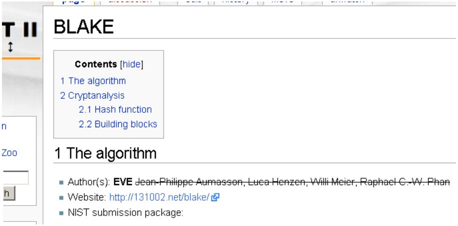
- “Entropy spraying” aka needles-in-a-haystack: better for most a posteriori backdoors (but e.g. HPC)
- “Chameleon” aka needles-in-a-needlestack: an option for a priori designs

Not just math but social-engineering

No measurable “cross-section”

Automated tools may help





Eve is a consultant paid to improve BLAKE's security

She replaces BLAKE's simplistic finalization

$$h_i + = v_i \oplus v_{i+8}, \quad i = 0, \ldots, 7$$

with the "more secure"

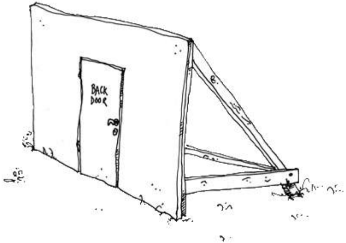
$$((v_0 \oplus v_1) \ggg 11) + ((v_2 \oplus v_3) \ggg 18) + ((v_4 \oplus v_5) \ggg 11) \oplus ((v_6 + v_7) \ggg 20) \oplus ((v_8 + v_9) \ggg 19)$$
$$((v_1 \oplus v_2) \ggg 17) + ((v_3 \oplus v_4) \ggg 16) + ((v_5 \oplus v_6) \ggg 28) \oplus ((v_7 + v_8) \ggg 10) + ((v_9 + v_{10}) \ggg 30)$$
$$((v_2 \oplus v_3) \ggg 12) + ((v_4 + v_5) \ggg 17) \oplus ((v_6 \oplus v_7) \ggg 13) \oplus ((v_8 + v_9) \ggg 22) \oplus ((v_{10} \oplus v_{11}) \ggg 7)$$
$$((v_3 \oplus v_4) \ggg 7) \oplus ((v_5 \oplus v_6) \ggg 5) \oplus ((v_7 + v_8) \ggg 11) + ((v_9 + v_{10}) \ggg 2) \oplus ((v_{11} + v_{12}) \ggg 9)$$
$$((v_4 \oplus v_5) \ggg 6) + ((v_6 + v_7) \ggg 6) + ((v_8 \oplus v_9) \ggg 4) + ((v_{10} \oplus v_{11}) \ggg 21) \oplus ((v_{12} + v_{13}) \ggg 15)$$
$$((v_5 + v_6) \ggg 4) + ((v_7 + v_8) \ggg 30) + ((v_9 \oplus v_{10}) \ggg 30) + ((v_{11} + v_{12}) \ggg 29) + ((v_{13} \oplus v_{14}) \ggg 2)$$
$$((v_6 \oplus v_7) \ggg 22) \oplus ((v_8 \oplus v_9) \ggg 1) \oplus ((v_{10} + v_{11}) \ggg 30) \oplus ((v_{12} \oplus v_{13}) \ggg 22) + ((v_{14} + v_{15}) \ggg 21)$$
$$((v_7 + v_8) \ggg 19) \oplus ((v_9 + v_{10}) \ggg 8) + ((v_{11} + v_{12}) \ggg 25) \oplus ((v_{13} \oplus v_{14}) \ggg 15) \oplus ((v_{15} \oplus v_0) \ggg 10)$$

- The new BLAKE is at least as secure as the original ("provable undiscoverability", "plausible deniability")
- Eve knows a collision for the compression function, between two chosen messages (here "YES" and "NO")
- She can use it to generate many hash collisions
- She used the neutral structure

$$(v_0 \bullet v_1) \ggg r_1 \bullet (v_2 \bullet v_3) \ggg r_2 \bullet (v_4 \bullet v_5) \ggg r_3 \bullet (v_6 \bullet v_7) \ggg r_4 \bullet (v_8 \bullet v_9) \ggg r_5$$

- Any new malicious instance generated within seconds

Conclusions



Malicious cryptography is academia-understudied

First published work about malicious hashing

Rich playground for malicious designers

Numerous real-life applications (not only malicious ones)

Research goals include awareness and malware prevention

Future work:

- More/better definitions
- Refined backdoor strategies
- Advanced detection strategies
- Hashing vs. implementations (SW/HW) backdoors
- Theoretical connections with obfuscation, WBC, etc.
- Quantum backdoors?

Eve’s SHA3 candidate: malicious hashing

Jean-Philippe Aumasson

NAGRAVISION
KUDELSKI GROUP