## Algebraic S-Box Recovery: the case of Cryptomeria

Cryptographers follow Kerckhoffs' principle, unless they don't.

Ralf-Philipp Weinmann <ralf@coderpunks.org>

Echternach Seminar on Symmetric Cryptography, Echternach, Luxembourg 2008-01-11



# Background

- CPRM/CPPM: Content Protection for Recordable / Pre-Recorded Media
- "The CPRM / CPPM Specification defines a renewable cryptographic method for protecting entertainment content when recorded on physical media."
- Uses proprietary block cipher, Cryptomeria (a.k.a. C2)
- Well, not quite proprietary... Design is public, S-Boxes are TRADE SECRETS
- This talk aims at improving interoperability...;)



#### Quote of the day

"The 4C companies designed and adopted C2, despite general cryptographic design principles which encourage use of well-known and well-evaluated ciphers, since no well-known alternatives had been identified that provided the necessary balance between suitability of hardware and software implementation, *minimal licensing fees*, and the ability to *exclusively license C2* for use in 4C content protection solutions. This last attribute is particularly important, as circumvention of the 4C technologies will *likely require* use of the C2 cipher algorithm, which must be licensed from 4C"



#### DVD-Audio

- DRM measure broken in 2004: Nero implemented C2 in software
- Result: DVD-Audio discs can be freely copied



**TECHNISCHE** 

**UNIVERSITAT** 

DARMSTADT

#### DVD-Audio

- DRM measure broken in 2004: Nero implemented C2 in software
- Result: DVD-Audio discs can be freely copied



## Video DVD-R's and Japanese HDTV

- Japanese distributed cracking effort in 2004 that tried the DVD Audio S-Box in a key recovery: *FAILED*
- S-Box reverse-engineered from software only recently (approx. July 2007)
- Different S-Box used



#### SD-Cards...

- Nothing known yet
- My guess: uses different S-Box as well.





#### ... contain more than meets the eye



No, I didn't pry that open myself, the picture is actually from Wikipedia



#### ... contain more than meets the eye



No, I didn't pry that open myself, the picture is actually from Wikipedia



#### ... contain more than meets the eye



No, I didn't pry that open myself, the picture is actually from Wikipedia

UNIVERSITAT

DARMSTADT

#### Standard for SD card quite recent

- November 29, 2007 The 4C Entity, LLC is pleased to announce that licensing for a new CPRM Specification, SD Memory Card: SD-Separate Delivery Video Part (SD-SD Video), Revision 0.9
- September 25, 2007 The 4C Entity, LLC is pleased to announce that a white paper entitled "SDSD-CPRM: Flexible Protection for Content Protection" is now available. SDSD-CPRM (Secure Digital - Separate Delivery) is a new extension to the SD Memory Card that addresses the difficulties of sharing and protecting digital content.







# The Cryptomeria cipher



- Feistel structure (modular addition)
- 64-bit block size
- 56 bit key
- 10 rounds
- key mixing: modular addition



### Observations on the cipher



- S-Box used only once per round
- All bytes of the output word affected by S-Box
- but... diffusion is not good
- combination of XOR and modular addition makes things interesting



## Cryptomeria key scheduling

- 56-bit cipher key K
- $T(X) = (X_{[24..55]} + (S(X_{[0..7]}) << 4)) \mod 2^{32}$
- *i*-th round key:  $Rk_i = T((K <<< 17i) XOR i)$



## Objective

- recover S-Box
- brute-force: 256!  $\approx 2^{1684}$  (assumption: S-Box bijective)
- using encryption oracle, chosen-key attack scenario
- don't want to use side-channel attacks!
- can't use SASAS attack, too many rounds for that



#### Chosen-key attacks

- First (and only) chosen-key attack proposed by Saarinen against the FOCT (GOST) 28147-89 block cipher (1998, unpublished)
- technique uses fixed points, takes advantage of weak key schedule: cannot be used in our case







## Strategy for an attack

- Merge two worlds: differential attacks and algebraic cryptanalysis
- reduce number of different (!) active S-Boxes by differential methods
- set up system of equations, S-Box outputs are unknowns (key variables)
- chosen-key model, allows to minimize S-Boxes in key schedule
- 5 S-Boxes are different for each pair in set of p/c pairs



## S-Boxes in the key schedule

- Enumerate all possible combinations for less than 5 S-Boxes
- Solve large (> 10<sup>6</sup>) number of small linear systems (56 vars)
  - 2<sup>10</sup> keys which only trigger 4 S-Boxes
- Optimal: not possible to trigger less than 4 S-Boxes



#### Differentials for the cipher

- for p,  $(p \oplus \Delta_1)$ ,  $(p \oplus \Delta_2)$  the same S-boxes will be triggered in first 4 rounds
- carry bits need to be taken into account
- filtering on the lowest-most 8 bits of ciphertext
  - assures same S-Box is used in last round
  - approximately 2<sup>16</sup> trials needed



#### Schematics of the idea



#### Schematics of the idea



#### Schematics of the idea







## The equation systems

- multivariate quadratic system
- if cipher did not use modular addition:
  - solve linear system & be finished
- modular addition results in carry chains
- systems composed of modular addition, register-wise logical operations and bit-shifts: common in symmetric cryptography
  - MD4/5, SHA-1, SHA-256/384/512 [hash functions] TEA, Michael (WPA) [block ciphers] Helix/Phelix, SALSA20 [stream ciphers]



## Systems for modular addition (mod 2<sup>n</sup>)

- addition in round function: full addition of two 32 bit values addition in the key schedule: a 12-bit quantity + 32-bit quantity. Assume upper 19 carry bits to be zero.
- each full addition introduces 32 variables for the result and 31 carry variables.

$$\begin{aligned}
z_1 &= x_1 + y_1 \\
c_1 &= x_1 y_1 \\
z_2 &= x_2 + y_2 + c_1 \\
c_2 &= x_2 y_2 + c_2 x_2 + c_2 y_2 \\
\vdots &\vdots \\
z_n &= x_n + y_n + c_{n-1}
\end{aligned}$$

## Strategies

- variables can be guessed to "break" carry chains
- important which variables are guessed
- talk by Gregory Bard at Fq8 Conference (Deakin University) this year: (On the Connection Number of a Particular Graph of a Polynomial System of Equations over a Finite Field)
  - optimal guessing for MQ systems: indications for NP-hard problem
- optimal guessing not needed, heuristics are "good enough"



#### Results

- equation systems for 4 rounds (5 different S-Boxes, 2 p/c pairs): easily broken with COTS software (Magma) without guessing: 142MB, 22secs on Opteron 2218 (2.6GHz, 1MB L2 cache) [1036 equations, 944 variables]
- equation system for up to 8 rounds (18 S-Boxes, 3 p/c pairs) broken given enough memory and guessing (6 variables): 20GB used, 15h on HHLR (per iteration) for 8 rounds [2816 equations, 2652 variables]
- 10 rounds are close. New tricks needed? (PolyBoRi?)



## Acknowledgements / Thanks

- John Gilmore for bringing up the cipher
- Ulrich Kühn for looking at Cryptomeria with me in Dagstuhl
- Jintai Ding for the combinatorial idea for the key schedule
- You! For listening to my talk.



# Questions?