White-Box Cryptography

Ph.D graduation presentation
Brecht Wyseur
March 5, 2009, Heverlee

Cryptography: the basic principle

- Basic assumption
  - Adversary has knowledge of the algorithm (Kerckhoffs 1883 [104])
  - Security of a cryptosystem relies on the confidentiality of the key

Keys in cryptography

- Symmetric cryptography
- Asymmetric cryptography
  - Digital signatures
  - Public-key encryption

Cryptography is part of our modern life

- Telecommunication
- Financial
- Transport
- Identification
- Recreational
- ...

Example: iTunes

- User has an incentive to attack the software
  - Break the restrictions posed on content
  - To steal protected data
  - To cheat (e.g., in an online game; access control)
Software Attacks

- When a user has an incentive to attack
- Or is subject to malware

White-box attack model
- Adversary has fully-privileged access to the execution platform
- Dynamic execution (with instantiated cryptographic keys) can be observed.
- Internal details of implementations are completely visible and alterable at will.

Entropy Attack

Shamir and Van Someren, 1999, [167]

- Keys need to be chosen at random from a (uniform) distribution → high entropy.
- Code typically contains structure → low entropy

Key Whitening Attack

(Orlandi and Kurosawa, 2006 [104])

- Strategy: identify and overwrite S-box definition in binary:
  \[ S \rightarrow 0, \text{ then } \ y = 0 + k \]

Digital Rights Management (DRM)

- Digital management of rights (restrictions)
- Audio, video, software (including games), education material, etc.
- IFPI
  - Global digital revenues by industry (2008)

<table>
<thead>
<tr>
<th>Industry</th>
<th>Revenue %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Games</td>
<td>35%</td>
</tr>
<tr>
<td>Recorded Music</td>
<td>20%</td>
</tr>
<tr>
<td>Newspapers</td>
<td>4%</td>
</tr>
<tr>
<td>Films</td>
<td>4%</td>
</tr>
<tr>
<td>Magazines</td>
<td>1%</td>
</tr>
</tbody>
</table>

- Digital music: €3 billion in trade value

Software Piracy

- Illegal copy and distribution of software
- Business Software Alliance (BSA) (2006)

Belgium: 27% pirate SW
- 35% word-wide
- 82% in China

GTA IV – 1000 developers for 3.5 years
Estimate production cost: €80 million
Traditional Assessment of Security

- Model of the active adversary:
  - Interaction with key-instantiated oracles
  - Security Notion: objective and capabilities

Example security notion: \textbf{KR-CPA}

- Key Recovery under Chosen Plaintext Attack

Our main research question

How can cryptographic primitives be implemented in software, such that they remain secure?

Our Contributions

- Practical implementations \textit{(Chapter 3)}
  - Cryptanalysis of white-box DES implementations
  - Analysis of basic building blocks and invertibility issue
- Formal model \textit{(Chapter 4)}
  - Formalizing white-box cryptography
  - Positive and negative results
  - Extensions towards probabilistic primitives
- Applications \textit{(Chapter 5)}
  - Links with diverse related techniques
  - Development of practical solutions in software security

Overview

- Introduction
- White-box security assessment
  - White-box implementations \textit{(Chapter 3)}
  - Formal model and (im)possibility result \textit{(Chapter 4)}
- Applications and related research domains \textit{(Chapter 5)}
- Conclusions and future work

Overview

- Introduction
- White-box security assessment
  - White-box implementations
- Formal model and (im)possibility result
- Applications and related research domains
- Conclusions and future work
White-Box Implementations

- The single line of defense is HOW to implement a cipher.
- Software implementation with instantiated secret key

Goal: Effort of analysis ≥ BB attack
Ideal: Implement the cipher as one big lookup table

Internal encodings

- Consider the chain $L_3 \circ L_2 \circ L_1$
- $L_1$ contains key information
- Obfuscate $L_1$ with the bijections $b_1$ and $b_2$
- Encoded chain: $L'_3 \circ L'_2 \circ L'_1$

Security of encoded networks

- Local security
  - If $L_3 = f_3$ is bijective, then $L'_3$ is locally secure, because
    $\forall k, \exists b_1, b_2$ such that $L'_3 = b_1^{-1} \circ f_k \circ b_2$
  - This is infeasible to map to entire implementation (Lynn et al., 2004 [118])

- Global Security
  - Metrics
    - Diversity $k = \neq L'$
    - Ambiguity $L' = \neq k$
    - Scrutiny $\rightarrow$ Chapter 3
    - Prove $\rightarrow$ Chapter 4

State of the art

<table>
<thead>
<tr>
<th>WB DES</th>
<th>WB AES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chow et al. 2002</td>
<td>Chow et al. 2002</td>
</tr>
</tbody>
</table>

- Naked variant
- Encoded variant
- Improved variant
- Fault injection attack (Chow et al. 2002)
- Statistical attack (Chow et al. 2002)
- Cryptanalysis (Chow et al. 2002)
- Generic Cryptanalysis (Naik et al. 2004)
- Truncated differential cryptanalysis
- basic building block analysis

Differential cryptanalysis

- Introduce a difference (targeted fault) in the application
- Observe the fault propagation
- Learn.

- Deployed to analyze white-box implementations at the ‘edges’ of the implementation (first/last round)
- In this dissertation: a new strategy of truncated differential cryptanalysis on the internal rounds (hence independent from external protections)
Cryptanalysis of White-Box DES Impl.

- Wyseur et al., 2007 [191]
- Cryptanalysis on internal round structure, independent of external encodings
- **Strategy:**
  1. Distinguish round input differences that propagate slow
  2. Construct set of differences that correspond to flips of single bits at input of DES S-boxes
  3. ...

Algebraic Cryptanalysis

- **Main strategy:**
  - Remove the non-linear component of the internal encodings
  - **Theorem:** \( S = \{ Q \circ \oplus_3 \circ Q^{-1} \}_{\text{GF}(2^8)} \) yields \( Q \), with \( A = Q \circ Q \text{ affine} \)
  - Construct algebraic equations
  - Solve the equations to obtain key information.

Analysis of basic building block

- \( y_0 := g \circ \bigoplus f_0^{-1}(c_0) \circ f_1^{-1} \)
- \( y_1 := g \circ \bigoplus f_0^{-1}(c_1) \circ f_1^{-1} \)
- Hence: \( y_0 \circ y_1^{-1} = g \circ \bigoplus f_0^{-1}(c_0) \circ f_1^{-1} \)
- The family of functions \( \{ g \circ \bigoplus \circ g^{-1} \}_{c \in \text{GF}(2^8)} \) yields information on the encoding \( g \).
- The obtained information can be used as an alternative approach to cryptanalysis the DES and AES implementations.

Cryptanalysis of WBDES (2) [Wyseur et al., 2007 (191)]

- Compute inputs to the S-boxes
- Recover the key (up to some natural ambiguity)
- Recover the external encodings

Algebraic cryptanalysis

- Demonstrated on white-box AES implementation (O. Billet, H. Gilbert and C. Eich-Chaart, 2004 [20])
- Extended to ‘SLT’ ciphers (includes MDS-based ciphers) (W. Michals, P. Gorissen and H. Hiddemann, 2008 [135])
- … towards analysis of block cipher building blocks (our contribution)
- Example: encoded addition operation

Invertibility (PR-CPA)

- \( \text{wbAES}: 2^{23} \) because lookup tables work on per-column basis
- \( \text{wbDES}: 2^6 \gg 2^{16}, \text{DES natural resistance} \)
- Because of the use of parallel, encoded addition networks of 96 bits to 4 bits
- But, XOR building block analysis defeats non-linearity of these networks

- KR-CPA is often not satisfactory
- PR-CPA (plaintext recovery under chosen plaintext attack) is much more interesting in practice

**Theorem 1**

- Towards analysis of block cipher building blocks

**Main strategy:**

1. Remove the non-linear component of the internal encodings
2. **Theorem:** \( S = \{ Q \circ \oplus_3 \circ Q^{-1} \}_{\text{GF}(2^8)} \) yields \( Q \), with \( A = Q \circ Q \text{ affine} \)
3. Construct algebraic equations
4. Solve the equations to obtain key information.
Conclusions chapter 3

- White-box implementations of DES and AES are insecure
  - Differential cryptanalysis
  - Algebraic cryptanalysis

- Attacks are specific to the cryptographic primitive, or a building block that is present.
  - Towards a block cipher with building blocks that is suitable to be implemented in software
  - Proposals are formulated in dissertation.

- Extensions towards non-invertibility and key update

Overview

- Introduction
- White-box security assessment
  - White-box implementations
  - Formal model and (im)possibility result
  - Applications and related research domains
- Conclusions and future work

Obfuscation – soundness

- Predicate-based definition (Barak et al., 2001 [6])
  \[ \Pr[|A(O(P)) = \pi(P)|] - \Pr\left[|S^X_P(1^P)| = \pi(P)\right] \leq \text{neg}(P) \]

- Distinguisher-based definition
  \[ \Pr[D(A(O(P))) = 1] - \Pr\left[D(S^X_P(1^P)) = 1\right] \leq \text{neg}(P) \]

Models for Obfuscation

- Many models and results have been presented, but…
  - Predicate-based definition is too weak (meaningless)
  - Distinguisher-based definition is too strong (nothing interesting possible: deterministic & obfuscatable \(\Rightarrow\) learnable)

- Learnable
  - Learnable functions not of our interest
  - Cryptographic scheme’s should be non-learnable.
  - Also, obfuscation cannot capture cryptographic security requirements

Our approach to capture WBC

- A new model, based on predicate-based notion for obfuscation, specific to WBC (Saxena and Wyseur, 2008 [165])
  - We capture “meaning” with security notions
    - Attack goals
    - Attack capabilities (described as a game between a challenger and the adversary)

  “White-Box Property”
Our approach to capture WBC (2)

\[
\begin{align*}
\text{Black-box game} & : A & \rightarrow & \text{context} \\
\text{White-box game} & : A & \rightarrow & \text{context}
\end{align*}
\]

\[
\{ \text{Adv}_{BB}^T(k) = \text{Pr}[x \leftarrow \{0,1\}^n : \text{Game}_{BB}(k, x, a) = 1] \} = \left[ \begin{array}{c}
\text{Adv}_{BB}^T(k) = \text{Pr}[x \leftarrow \{0,1\}^n : \text{Game}_{BB}(k, x, a) = 1]
\end{array} \right]
\]

\(O\) is a secure obfuscator for \(Q\), under the \(sn\) security notion, if

\[
\max_{A \in \text{PPT}} \text{Adv}_{BB}^T(k) - \max_{A \in \text{PPT}} \text{Adv}_{BB}^T(k) \leq \text{neg}(k)
\]

Proof of impossibility result

\(Q[1, q] \{\text{input} Y\} \{\)

- If \((Y(a) = Q[q](a))\) then output \(x\)
- else output \(0\)

\(\text{Black-box advantage: } \forall A \in \text{PPT} : 0 \leq \text{Adv}_{BB}^T(k) < \alpha(k)\)

Proof of impossibility result (2)

\(Q[1, q] \{\text{input} Y\} \{\)

- If \((Y(a) = Q[q](a))\) then output \(x\)
- else output \(0\)

\(\text{White-box advantage: } \exists A \in \text{PPT} : 1 \geq \text{Adv}_{BB}^T(k) \geq 1 - \beta(k)\)

Proof of impossibility result (3)

\(\forall A \in \text{PPT} : 0 \leq \text{Adv}_{BB}^T(k) < \alpha(k)\)

\(\exists A \in \text{PPT} : 1 \geq \text{Adv}_{BB}^T(k) \geq 1 - \beta(k)\)

\(\text{White-box property: } O\) is secure for \(Q\) under \(sn\), if:

\[
\max_{A \in \text{PPT}} \text{Adv}_{BB}^T(k) - \max_{A \in \text{PPT}} \text{Adv}_{BB}^T(k) \leq \text{neg}(k)
\]

\(\text{But: } \max_{A \in \text{PPT}} \text{Adv}_{BB}^T(k) - \max_{A \in \text{PPT}} \text{Adv}_{BB}^T(k) > 1 - \alpha(k) - \beta(k)\)

Positive result

\(\exists A \in \text{PPT} : 1 \geq \text{Adv}_{BB}^T(k) \geq 1 - \beta(k)\)

Remark: positive result is based on a cipher that consists of asymmetric building blocks (pairings). We started with white-box crypto for symmetric encryption schemes (DES, AES).
Overview

Introduction
- White-box security assessment
  - White-box implementations
  - Formal model and (im)possibility result
- Applications and related research domains
- Conclusions and future work

Software security

- White-box cryptography is only a small piece in the puzzle.
  1. 128-bit AES key → 770 Kbytes key (the white-box AES implementation lookup tables)
  2. Now we got a larger key… so what?
  3. Result is more flexible
- Fix key into the application (prevent code lifting) – external encodings
- A leverage for other software protection techniques
  - Traitor tracing
  - Obfuscation
  - Software Tamper Resistance
- ...
Q&A

» Thank you.