Soundness of Formal Encryption in the Presence of Key Cycles

Pedro Adão[1,2], Gergei Bana[2], Jonathan Herzog[3] and Andre Scedrov[2]

pad@math.ist.utl.pt, {bana,scedrov}@math.upenn.edu, jherzog@mitre.org

[1] Center for Logic and Computation, IST, Lisboa

[2] Department of Mathematics, University of Pennsylvania

[3] MITRE Corporation

Protocol eXchange Workshop

February, 2 2005

Naval Postgraduate School, Monterey, CA

Introduction

- Cryptographic protocols: two models, alike in dignity
  - *Formal*, or Dolev-Yao model
  - *Computational* model from complexity theory
- Much recent work relates the two
  - Build formal-to-computational protocol interpretation
  - Map formal security goals to computational goals
  - Prove *soundness* or *completeness*

AR Logic of Formal Encryption

- AR define a very simple algebra of terms;
- Expressions are built from two simple sets **Keys** $= \{K_1, K_2, K_3, ...\}$ and **Blocks** $\subseteq \{0,1\}^*$ via paring and encryption;

$$\begin{array}{rcl} \textbf{Exp} & ::= & \textbf{Keys} \mid \textbf{Blocks} \mid (\textbf{Exp}, \textbf{Exp}) \mid \{\textbf{Exp}\}_{\textbf{Keys}} \mid \\ \textbf{Pat} & ::= & \textbf{Keys} \mid \textbf{Blocks} \mid (\textbf{Pat}, \textbf{Pat}) \mid \{\textbf{Pat}\}_{\textbf{Keys}} \mid \Box \end{array}$$

$$((\{0\}_{K_8}, \{100\}_{K_1}), ((K_7, \{(\{0101\}_{K_9}, \{K_8\}_{K_5})\}_{K_5}, \{K_5\}_{K_7}$$

$$((\{0\}_{K_8}, \quad \Box \quad), ((K_7, \{(\quad \Box \quad, \{K_8\}_{K_5})\}_{K_5}, \{K_5\}_{K_7}$$

- Two expressions M and N are defined to be equivalent if $P(M) = P(N)\sigma$ for some key-renaming function σ.
- We denote this by $M \cong N$.

AR Logic of Formal Encryption (cont.)

- Formal expressions are mapped to (interpreted in) the computational model as follows:
  - For each $K \in \textit{Keys}(M)$ generate a key using the key generation algorithm;
  - Each $B \in$ **Blocks** is mapped to B;
  - Each pair (M, N) is interpreted as the pair of the interpretations;
  - Each encryption is interpreted by running the encryption algorithm.
- For expression M we denote its interpretation by $[\![M]\!]_\Phi$.

AR Logic of Formal Encryption (cont.)

- **Theorem:** Let M and N be *acyclic expressions* and let Π be a type-0 secure encryption scheme. Suppose that $M \cong N$. Then $[\![M]\!]_\Phi \approx [\![N]\!]_\Phi$.

AR Logic of Formal Encryption (cont.)

- **Theorem:** Let M and N be *acyclic expressions* and let Π be a type-0 secure encryption scheme. Suppose that $M \cong N$. Then $[\![M]\!]_\Phi \approx [\![N]\!]_\Phi$.
- **Problem:** This result does not apply to self-encrypting keys;

AR Logic of Formal Encryption (cont.)

- **Theorem:** Let M and N be *acyclic expressions* and let Π be a type-0 secure encryption scheme. Suppose that $M \cong N$. Then $[\![M]\!]_{\Phi} \approx [\![N]\!]_{\Phi}$.
- **Problem:** This result does not apply to self-encrypting keys;
- **What do we propose:** Solve this problem via sufficiently strong crypto;

AR Logic of Formal Encryption (cont.)

- **Theorem:** Let M and N be *acyclic expressions* and let Π be a type-0 secure encryption scheme. Suppose that $M \cong N$. Then $[\![M]\!]_\Phi \approx [\![N]\!]_\Phi$.
- **Problem:** This result does not apply to self-encrypting keys;
- **What do we propose:** Solve this problem via sufficiently strong crypto;
- [L02] proposed a solution for the problem of key-cycles by strengthening the formal adversary.

The problem of key-cycles

- More general form of self-encryption:
  - K_1 encrypts K_2
  - K_2 encrypts $K_3 \ldots$
  - K_n encrypts K_1
  - (Asymmetric encryption: K_i encrypts K_{i-1}^{-1})
- Can actually occur in Dolev-Yao model
- Possible to interpret formal messages with key cycles
- But known completeness or soundness results do not hold
- How to interpret? Two possibilities:
  - Reflects weakness of underlying crypto
  - Reflects weakness of proof methods

Underlying crypto

- Semantic security: main computational definition of security for public-key encryption
  - Adversary cannot distinguish encryptions of M_1, M_2
  - Adversary gets to choose M_1, M_2 itself
  - Adversary knows public (encryption) key k
- Note: adversary does not know decryption key k^{-1}
  - M_1, M_2 cannot depend on k^{-1}
  - No obvious security guarantees if they do
  - Same phenomena for CCA-1, CCA-2
- Dolev-Yao model: self-encrypting keys are A-OK
- Might actually be a real gap between the two models

Previous proof methods

- AR, AJ: soundness for indistinguishability properties
- MW, HG: completeness for indisitinguishability properties
- B, ABS: more general soundness, completeness properties
- H: soundness for non-malleability properties
- BPW: soundness for general trace-based properties
- HC, MW: soundness, completeness for MA, KE properties
- L: soundness via strengthening the “formal adversary"

- (Almost) all (soundness) proofs rely on some *hybrid argument*

Previous proof methods

- Previous results rely on *hybrid argument*
  - Powerful proof technique from computational crypto
  - Used to show: distinguishability of compound objects $\Rightarrow$ distinguishability of atomic objects
- Example: suppose this row (as a whole)

$$\circ \quad \Box \quad \circ \quad \circ \quad \circ \quad \circ \quad \Box \quad \Box \quad \circ \quad \circ$$

is distinguishable from this row (as a whole):

$$\circ \quad \circ \quad \circ \quad \Box \quad \Box \quad \Box \quad \Box \quad \Box \quad \Box \quad \circ$$

- Distinguishability $\cong$ distance in metric space
  - Better to say “distinguishable with *advantage* P”

The Hybrid Argument (cont.)

- Insert 10 intermediate rows
  - Each row changes at most one column

○ □ ○ ○ ○ ○ □ □ ○ ○

○ ○ ○ ○ ○ ○ □ □ ○ ○

○ ○ ○ □ ○ ○ □ □ ○ ○

. . .

○ ○ ○ □ □ □ □ □ ○ ○

○ ○ ○ □ □ □ □ □ □ ○

- By contradiction, must be two neighbors with distance $\geq P/10$.
- Suppose rows 2 & 3

The Hybrid Argument (cont.)

- Suppose X is either $\circ$ or $\square$.
- How to distinguish?
- Build the following:

$$\circ \quad \circ \quad \circ \quad X \quad \circ \quad \circ \quad \square \quad \square \quad \circ \quad \circ$$

  - If X is $\circ$, then this is row 2
  - If X is $\square$, then this is row 3
  - By above, adversary has advantage $\geq P/10$ in distinguishing
    - Advantage in distinguishing $\circ$, $\square$ must be $\geq P/10$ as well

Hybrid argument (conc.)

- If $\circ$, $\square$ are *indistinguishable*, then top & bottom rows are as well
  - Indistinguishable: *negligible* as security parameter grows
  - Negligible: shrinks faster than any polynomial
- Argument depends on:
  - Number of rows is polynomial in security parameter
  - Given entry for one column, can create rest of any row
  - Possible to “walk” from top to bottom by changing only one column at a time
- Why doesn’t this work for key-cycles?

AR hybrid argument

- Want to show that M, *pattern of* M ($P(M)$) are indistinguishable
- Build table:

$$
\begin{array}{rccccc}
M = & K_1^{-1} & \{K_2\}_{K_1} & \{101\}_{K_3} & \{K_5^{-1}\}_{K_4} & \{101\}_{K_5} \\
 & K_1^{-1} & \{K_2\}_{K_1} & \{101\}_{K_3} & \{K_5^{-1}\}_{K_4} & \square_{K_5} \\
 & K_1^{-1} & \{K_2\}_{K_1} & \{101\}_{K_3} & \square_{K_4} & \square_{K_5} \\
P(M) = & K_1^{-1} & \{K_2\}_{K_1} & \square_{K_3} & \square_{K_4} & \square_{K_5}
\end{array}
$$

($\square_k$: undecipherable encryption; maps to $\{\mathbf{0}\}_K$)

- If top & bottom are distinguishable, then $\{M'\}_{K'}$ & $\square_{K'}$ distinguishable
  - For some sub-message M', some *single* key K'

Key cycles

- Suppose M has a key-cycle. What should the rows be?

$$
\begin{array}{rccccc}
M = & K_1^{-1} & \{K_2\}_{K_1} & \{K_4^{-1}\}_{K_3} & \{K_3^{-1}\}_{K_4} & \{101\}_{K_5} \\
 & K_1^{-1} & \{K_2\}_{K_1} & \{K_4^{-1}\}_{K_3} & \{K_3^{-1}\}_{K_4} & \square_{K_5} \\
 & K_1^{-1} & \{K_2\}_{K_1} & \text{?} & \text{?} & \square_{K_5}
\end{array}
$$

- If next row is $\ldots \square_{K_3} \square_{K_4} \ldots$, no longer isolating *one* key
- Only other option: replace only one encryption
  - WLOG, $\ldots \{K_4^{-1}\}_{K_3} \square_{K_4} \ldots$

Key cycles (cont.)

- If next row is $\ldots \{K_4^{-1}\}_{K_3} \Box_{K_4} \ldots$, distinguishable neighbors might be:

$$\begin{array}{ccccc} K_1^{-1} & \{K_2\}_{K_1} & \{K_4^{-1}\}_{K_3} & \{K_3^{-1}\}_{K_4} & \Box_{K_5} \\ K_1^{-1} & \{K_2\}_{K_1} & \{K_4^{-1}\}_{K_3} & \Box_{K_4} & \Box_{K_5} \end{array}$$

- Does this let us distinguish $\Box_{K_4}$ and $\{K_3^{-1}\}_{K_4}$?
  - Given $X \in \{\Box_{K_4}, \{K_3^{-1}\}_{K_4}\}$, must make rest of row
  - How to make $\{K_4^{-1}\}_{K_3}$ from $\Box_{K_4}$?

Resolving key-cycles

- Current results silent about key cycles
- Two possibilities:
  1. Key-cycles not necessarily secure in computational model
  2. Key-cycles incompatible with hybrid argument
- This talk: *can* prove soundness for key-cycles
  - Will even use hybrid argument
  - Look beyond semantic security

Key-dependent messages (KDMs)

Consider following game:

- Referee creates fresh random key-pair (k, k^{-1})
- Adversary gets k, creates function f
- Referee secretly flips coin:
  - Heads: encrypts $f(k^{-1})$
  - Tails: encrypts $0^{|f(k^{-1})|}$
- Adversary gets ciphertext, tries to determine which one

- Random guessing yields 50% success rate
- Want: can't do better than this

Actual KDM-security

- Definition for KDM security actually more general
- Referee creates *vector* of keys $(\vec{k}, \vec{k^{-1}})$
  - Referee also flips coin once:
- Adversary gets $\vec{k}$, produces (i, f)
  - Heads: referee encrypts $f(\vec{k^{-1}})$ in k_i
  - Tails: referee encrypts $0^{|f(\vec{k^{-1}})|}$ in k_i
- As many of these rounds as adversary wants
- *KDM security* [BRS, CL]: can only guess coin-flip

Motivation for KDM-security

- KDM security introduced by BRS with the purpose of strengthening the adversary (stronger than CPA);
- Independently, a similar (weaker??) version called *circular security* was introduced by CL to deal with anonimity and credentials revocation;
- NO relation is known between CCA/CCA2 and KDM (or circular security)

The new hybrid argument

- Table has only 2 rows:

$$
\begin{array}{lllllll}
M = & K_1^{-1} & \{K_2\}_{K_1} & \{K_4^{-1}\}_{K_3} & \{K_3^{-1}\}_{K_4} & \{101\}_{K_5} \\
 & K_1^{-1} & \{K_2\}_{K_1} & \{0^{|K_4^{-1}|}\}_{K_3} & \{0^{|K_3^{-1}|}\}_{K_4} & \{000\}_{K_5}
\end{array}
$$

- Distinguishing these two rows breaks KDM security directly
- Special case where adversary asks referee to
  - Encrypt K_4^{-1} in K_3
  - Encrypt K_3^{-1} in K_4
  - Encrypt 101 in K_5

What does this mean?

- **Theorem:** Let M and N be *expressions* and let Π be a KDM-secure encryption scheme. Suppose that $M \cong N$. Then $[\![M]\!]_\Phi \approx [\![N]\!]_\Phi$.

What does this mean?

- **Theorem:** Let M and N be *expressions* and let Π be a KDM-secure encryption scheme. Suppose that $M \cong N$. Then $[\![M]\!]_\Phi \approx [\![N]\!]_\Phi$.
- Sufficiently strong crypto guarantees soundness
  - Even in presence of key-cycles

What does this mean?

- **Theorem:** Let M and N be *expressions* and let Π be a KDM-secure encryption scheme. Suppose that $M \cong N$. Then $[\![M]\!]_\Phi \approx [\![N]\!]_\Phi$.
- Sufficiently strong crypto guarantees soundness
  - Even in presence of key-cycles
- Where was the original problem? Crypto or argument?
  - Still don’t know

What does this mean?

- **Theorem:** Let M and N be *expressions* and let Π be a KDM-secure encryption scheme. Suppose that $M \cong N$. Then $[\![M]\!]_\Phi \approx [\![N]\!]_\Phi$.
- Sufficiently strong crypto guarantees soundness
  - Even in presence of key-cycles
- Where was the original problem? Crypto or argument?
  - Still don't know
- We are still learning what DY model assumes about underlying crypto
  - There are still surprises out there

Future work

- Same extensions of original AR result
  - Non-malleability?
- Not *all* proofs use hybrid argument
  - BPW, HC use “simulation argument”
  - Assume *no* keys are encrypted!
  - Very strong, how to weaken?
- Relationship between KDM-security, circular security, semantic security?
  - Chosen-ciphertext security?
  - Note: may already be known...