Symbolic verification of cryptographic protocols and its computational justification

Joint work with: Romain Janvier and Laurent Mazaré
Grenoble, December 2005

Yassine Lakhnech

Verimag

Yassine.Lakhnech@imag.fr

Plan of the talk

1. Introduction:
   - Cryptographic protocols: an example
   - Two types of attacks:
     (a) logical (symbolic)
     (b) computational (complexity-theoretic)
2. The complexity-theoretic approach for provable security.
3. The symbolic approach for provable security.
4. Provably secure cryptography.
5. The justification of the symbolic approach.
6. Concluding remarks

The Needham-Schroeder public key protocol

R. Needham and M. Schroeder.
Communications of the ACM, 21(12):993-999, 1978.
Goal: Mutual authentication.

$$
\begin{aligned}
A \to B &: \{N_a, A\}_{pk_B} \\
B \to A &: \{N_a, N_b\}_{pk_A} \\
A \to B &: \{N_b\}_{pk_B}
\end{aligned}
$$

- Two roles:

$$
\begin{array}{lll}
R_1(A,B): & \mathsf{new}(N_a); & |R_2(B): \quad ?\{z,y\}_{pk_B} \\
 & !\{N_a, A\}_{pk_B} \ | & \quad\quad \mathsf{new}(N_b); \\
 & ?\{N_a, x\}_{pk_A}; | & \quad\quad !\{z, N_b\}_{pk_y} \\
 & !\{x\}_{pk_B}; \quad\quad | & \quad\quad ?\{N_b\}_{pk_B}
\end{array}
$$

Cryptographic primitives

- Asymmetric encryption $(\mathcal{K}, \mathcal{E}, \mathcal{D})$
  - $\mathcal{K} : \mathbb{N} \to \{0,1\}^* \times \{0,1\}^*$ generates a pair of keys: the public and the private one
  - $\mathcal{E} : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}^*$
  - $\mathcal{D} : \{0,1\}^* \times \{0,1\}^* \to \{0,1\}^*$ such that

$$\mathcal{D}(\mathcal{E}(bs_1, bs_2), bs_3) = bs_1$$

if bs_3 is the private key corresponding to bs_2

Actions and protocol description

- Actions:

$$\alpha ::= !t \mid ?t(\tilde{x}) \mid x := t \mid x = t$$

  - $!t$ - output
  - $?t(\tilde{x})$ - input, $\tilde{x} \in \mathcal{V}$ instantiated by the action.
  - $x := t$ - assignment
  - $x = t$ - equality test

Actions and protocol description

- Actions:

$$\alpha ::= !t \mid ?t(\tilde{x}) \mid x := t \mid x = t$$

- A protocol description:

Parameters:

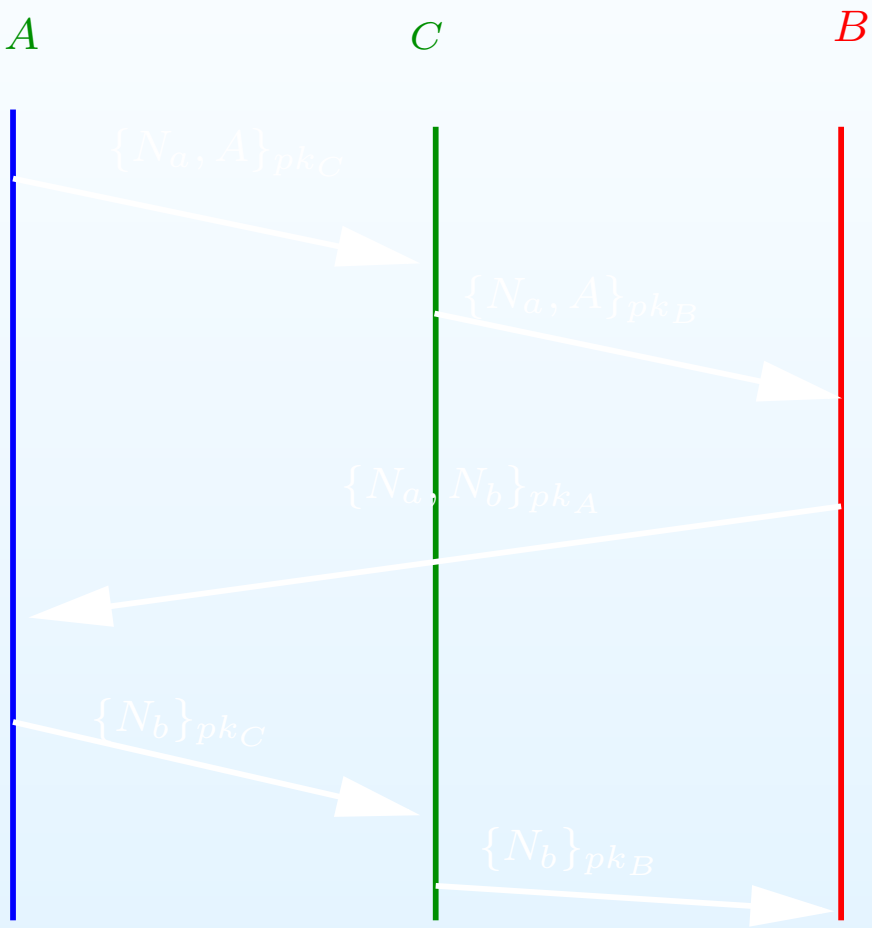
$P_1, \cdots, P_k$(principals, keys)

$N_1, \cdots, N_m$: fresh(nonces)

$\sum_{i=1}^{n} \alpha_1^i \cdots \alpha_{n_i}^i$

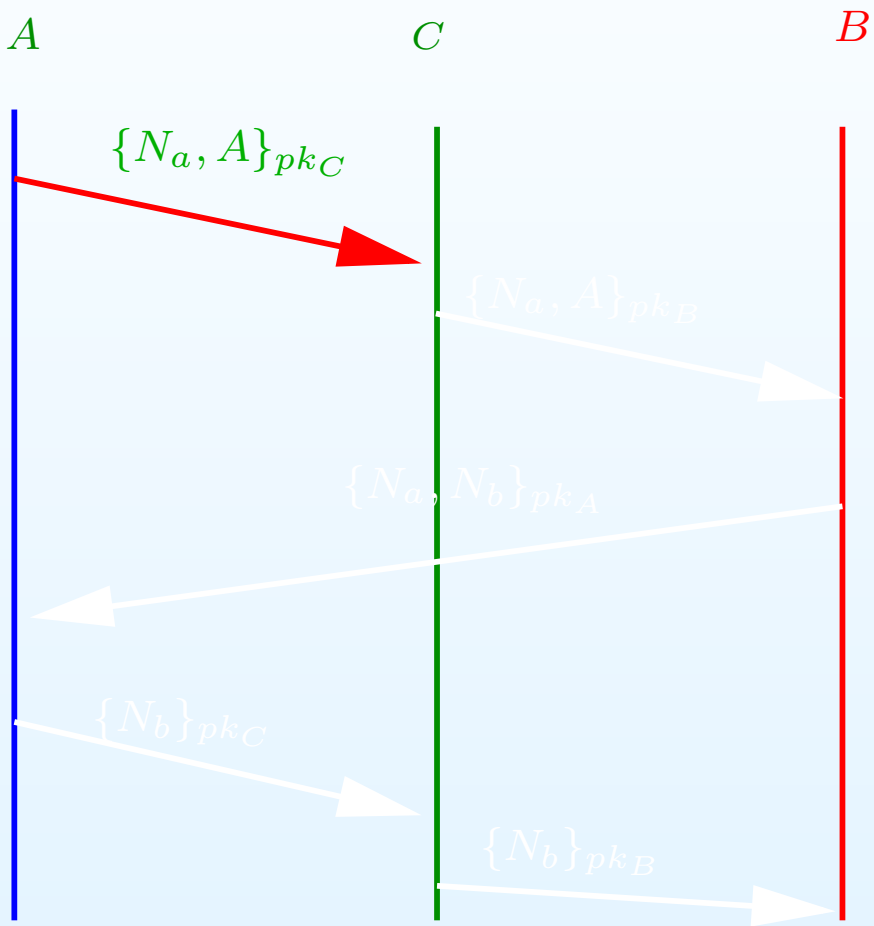
Lowe's logical attack

Logical means "abstract from cryptography".



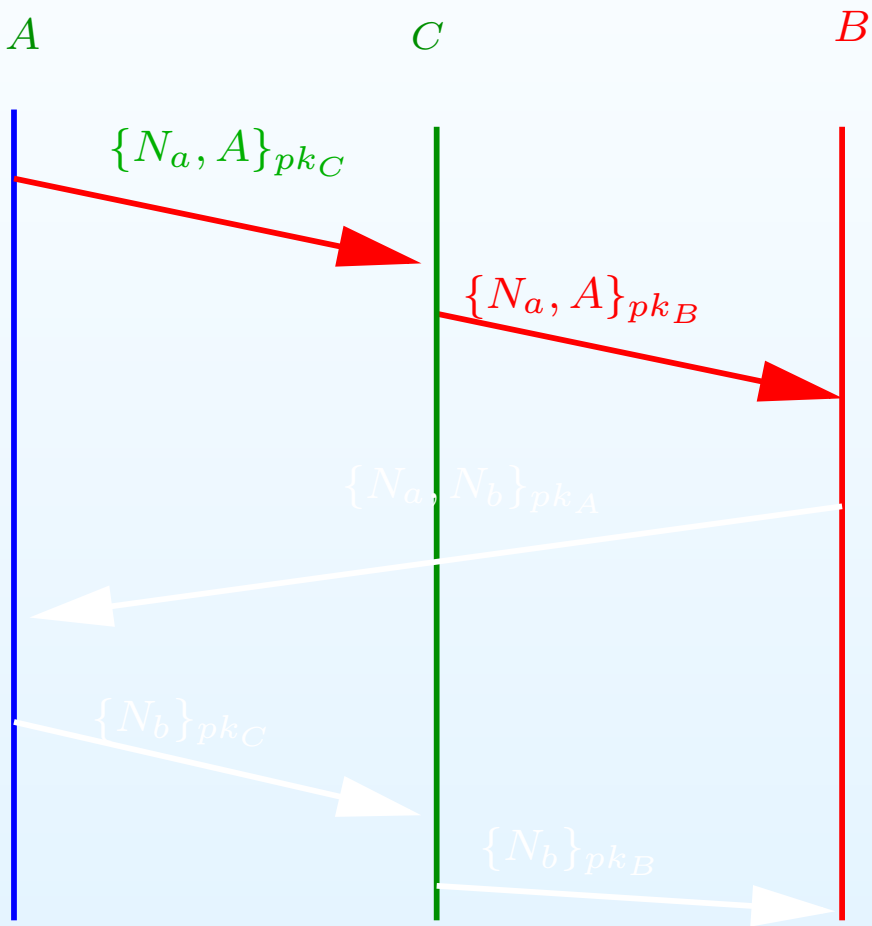
Lowe's logical attack

Logical means "abstract from cryptography".



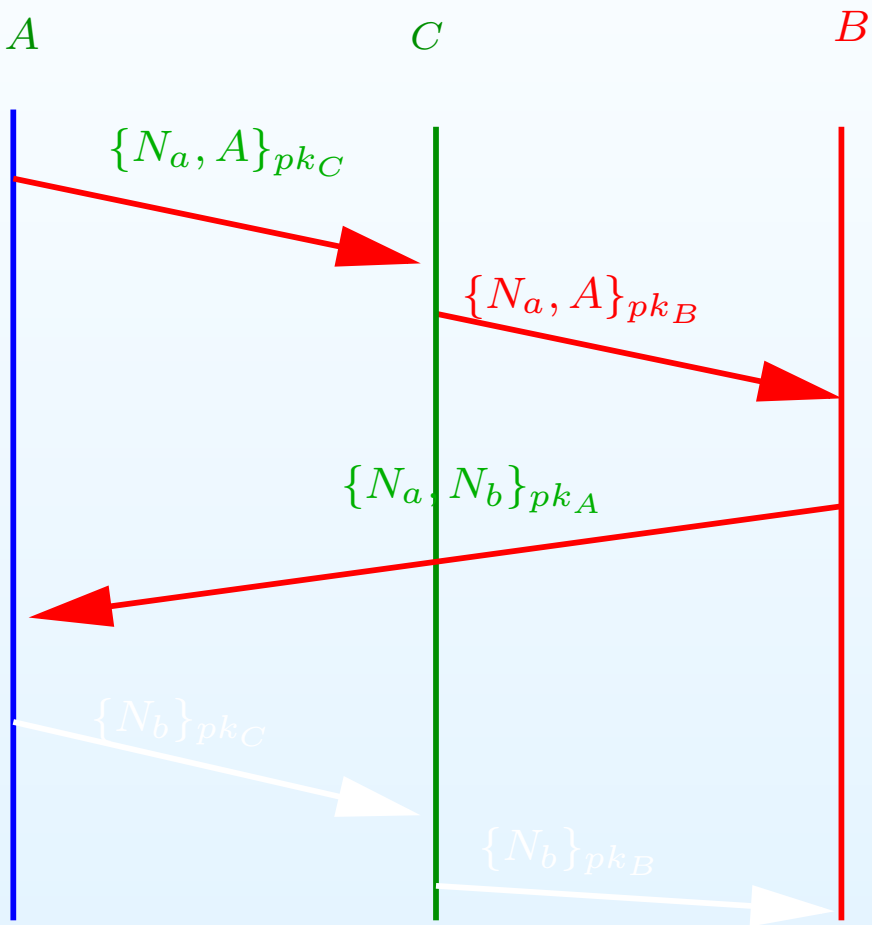
Lowe's logical attack

Logical means "abstract from cryptography".



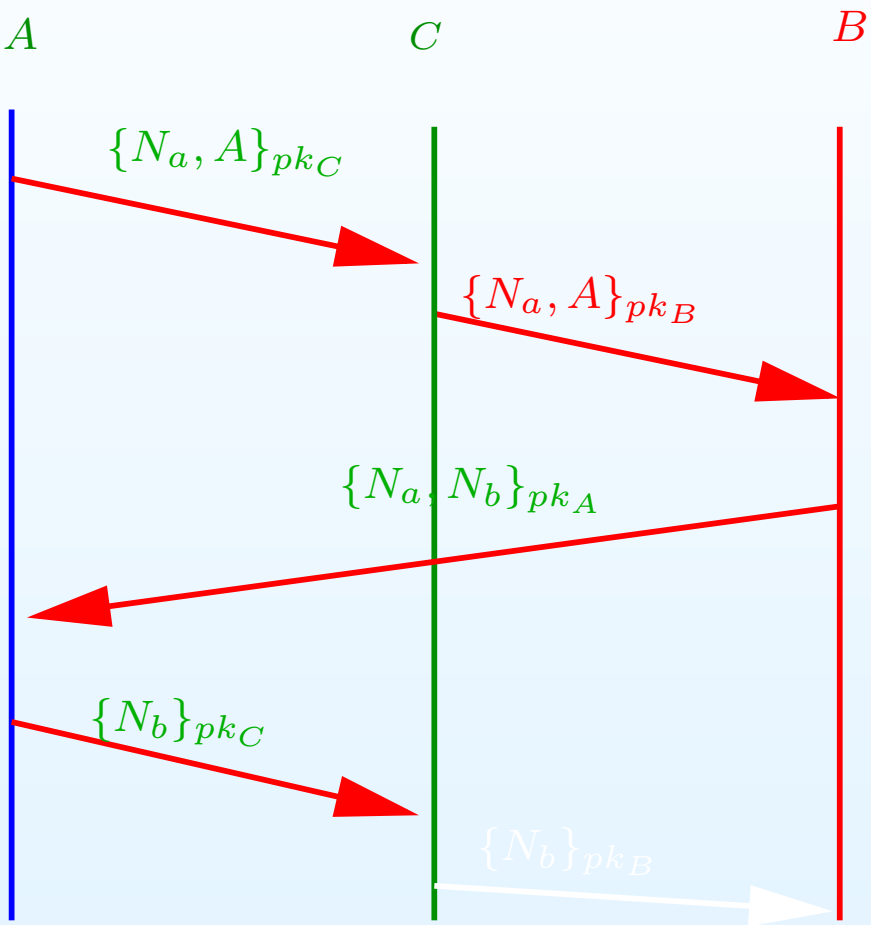
Lowe's logical attack

Logical means "abstract from cryptography".



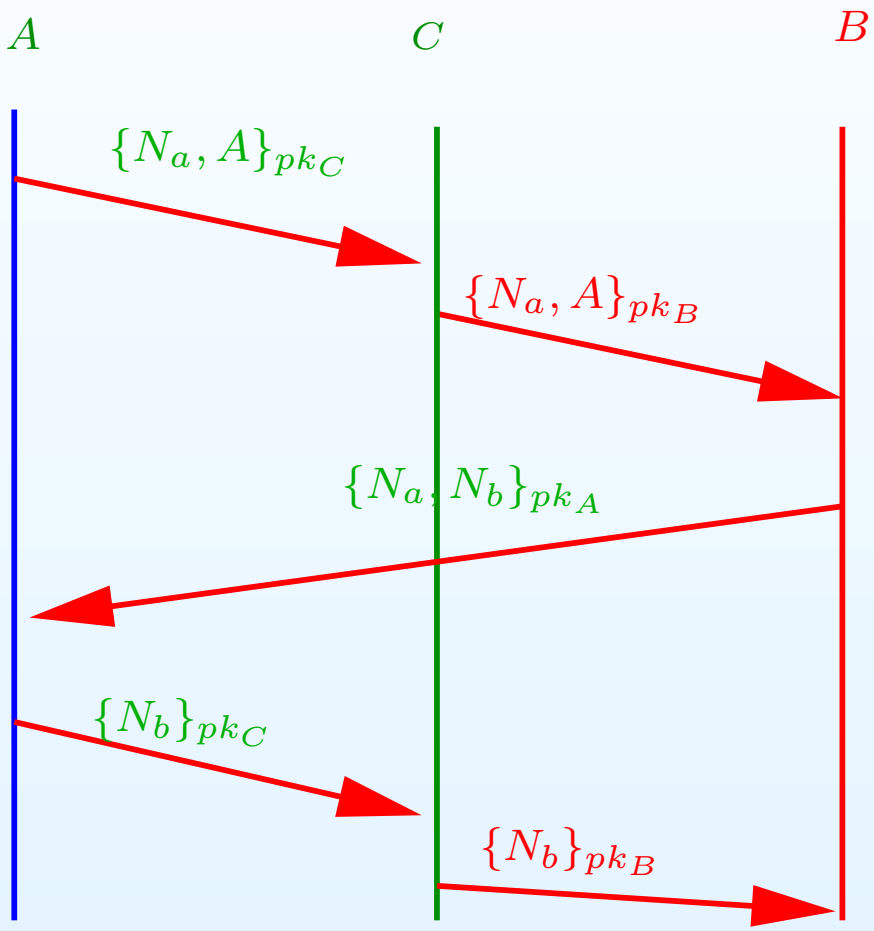
Lowe's logical attack

Logical means "abstract from cryptography".



Lowe's logical attack

Logical means "abstract from cryptography".



Needham-Schroeder-Lowe's protocol

$$
\begin{aligned}
A \rightarrow B : &\quad \{N_a, A\}_{pk_B} \\
B \rightarrow A : &\quad \{N_a, N_b, B\}_{pk_A} \\
A \rightarrow B : &\quad \{N_b\}_{pk_B}
\end{aligned}
$$

Are there logical attacks?

Needham-Schroeder-Lowe's protocol

$$
\begin{aligned}
A \rightarrow B : &\quad \{N_a, A\}_{pk_B} \\
B \rightarrow A : &\quad \{N_a, N_b, B\}_{pk_A} \\
A \rightarrow B : &\quad \{N_b\}_{pk_B}
\end{aligned}
$$

Are there logical attacks?
yes, if concatenation is implemented as an associative operator

Needham-Schroeder-Lowe's protocol

$$
\begin{aligned}
A \to B : &\quad \{N_a, A\}_{pk_B} \\
B \to A : &\quad \{N_a, N_b, B\}_{pk_A} \\
A \to B : &\quad \{N_b\}_{pk_B}
\end{aligned}
$$

Are there computational attacks?

Needham-Schroeder-Lowe's protocol

$$
\begin{aligned}
A \rightarrow B : &\quad \{N_a, A\}_{pk_B} \\
B \rightarrow A : &\quad \{N_a, N_b, B\}_{pk_A} \\
A \rightarrow B : &\quad \{N_b\}_{pk_B}
\end{aligned}
$$

Are there computational attacks?
yes, if encryption is malleable:

$$\mathcal{E}(N_a, N_b, B) \rightsquigarrow \mathcal{E}(N_a, N_b, C)$$

which is not excluded by the one-wayness condition:

$$Pr[x \stackrel{R}{\leftarrow} \{0,1\}^\eta; y \stackrel{R}{\leftarrow} \mathcal{A}(f(x)) : f(y) = x]$$

is negligible:

$f(\eta)$ is negligible, if $\forall c \in \mathbb{N} \exists \eta_0 \forall \eta \geq \eta_0 \cdot f(\eta) \leq \frac{1}{\eta^c}$

Needham-Schroeder-Lowe's protocol

$$
\begin{aligned}
A \to B : &\quad \{N_a, A\}_{pk_B} \\
B \to A : &\quad \{N_a, N_b, B\}_{pk_A} \\
A \to B : &\quad \{N_b\}_{pk_B}
\end{aligned}
$$

Are there computational attacks?

Conlcusion: There is a need for *formal verification of cryptographic protocols*.

Formal verification of cryptographic protocols

Rigorous (mathematical) definitions of:

1. The semantics of cryptographic protocols: their behaviors.
2. Their properties.

and methods to rigorously prove or disprove:

$$\Pi \models \varphi, \text{ i.e.,}$$

protocol Π satisfies property φ.
In this talk:

1. The complexity-theoretic approach
2. The symbolic approach
3. Their relation

The complexity-theoretic approach: semantics

- Data are bitstrings.
- Cryptographic primitives are randomized poly-time Turing machines.
- Protocols are communicating randomized poly-time Turing machines that may call cryptographic primitives.
- Adversaries are randomized poly-time Turing machines with rigorously defined capabilities and corruption model:
  - Static corruption vs. Dynamic corruption
  - Which information is leaked when a principle is corrupted

The complexity-theoretic approach: the properties

Two types of security definitions

- Game-based security definitions
- Simulation-based security definitions

The complexity-theoretic approach: the properties

- Game-based security definitions
  - The adversary makes some computation using the protocol as an oracle.
  - The behavior of the protocol (these oracles) depends on some randomly generated data (private keys, nonces, etc..)
  - The adversary has to answer a question concerning these data.

  Example: Real or random key.
  - Π a key-exchange protocol.
  - Π^* as Π except that a random value is exchanged instead of the session key.
  - The adversary should distinguish Π and Π^*.

The complexity-theoretic approach: the properties

- Simulation-based security definitions:
  - Define an ideal functionality Id that captures in an ideal way the *expected* properties of the protocol.
  - Define the real protocol: Rl.
  - Secrecy property: Simulation

    For any A there is an ideal adversary A^* such that $\mathcal{S}ys_{A,\mathsf{Rl}}$ and $\mathcal{S}ys_{A^*,\mathsf{Id}}$ are indistinsguishable.

Goldwasser-Micali'90, Micali-Rogaway'91, Beaver'92, Canetti'95, Shoup'99, Pflitzmann-Waidner'00, Canetti'01

The complexity-theoretic approach: the proofs

By reduction to the security of the cryptographic primitives or to a problem that is believed hard.

1. Fix a
   - A problem P that is believed to be hard: there is no rondomized Turing machine that solves P with non-negligible advantage.

The complexity-theoretic approach: the proofs

By reduction to the security of the cryptographic primitives or to a problem that is believed hard.

1. Fix a
   - A problem P that is believed to be hard: there is no rondomized Turing machine that solves P with non-negligible advantage.
2. Let A an adversary that breaks the protocol Π then A can be used to break P:
   - Given an adversary A against Π and φ, construct an adversary B against P and that uses A as a sub-routine such that

   (a) B is poly-time

   (b)

$$Adv_{A,\Pi,\varphi}(\eta) \leq Adv_{B,P}(\eta)$$

The complexity-theoretic approach: conclusions

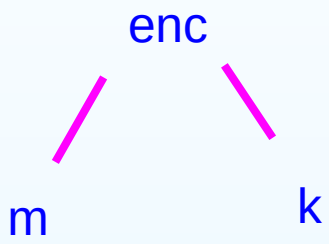
1. Detailed and realistic models.
2. Strong security garantees.

But

- Proof automation hard to achieve!
- Flawed proofs are not seldom!

The symbolic approach

- Replace bitstrings by terms (in a free algebra):
  - Encryption of m by k is the term $\{m\}_k$:



- No collision:
  - $n \neq n'$
  - $\{m\}_k = \{m'\}_{k'} \Rightarrow m = m' \wedge k = k'$
  - $\{m\}_k \neq \langle m_1, m_2 \rangle$
  - ...
- No randomization.
- Adversary: a single fixed process which is even less powerful than Turing machines.

The adversary: Dolev-Yao's model

- The adversary's knowledge: a set E of messages.
- The adversary's computation power:

$$\frac{m \in E}{E \vdash m} \qquad \frac{E \vdash m_1,\ E \vdash m_2}{E \vdash (m_1, m_2)}$$

$$\frac{E \vdash m,\ E \vdash k \in \mathcal{K}}{E \vdash \{m\}_k} \qquad \frac{E \vdash (m_1, m_2)}{E \vdash m_1}$$

$$\frac{E \vdash (m_1, m_2)}{E \vdash m_2} \qquad \frac{E \vdash \{m\}_k,\ E \vdash k^{-1}}{E \vdash m}$$

Expressiveness of the model

Secrecy often defined: there is no reachable configuration such that

$$E \vdash S$$

- The question $E \vdash m$ is P-complete.
- Turing-complete model.
  - Size of messages is unbounded.
  - Creation of fresh names (nonces) unbounded.
  - Number of transitions unbounded.

Decidability results for secrecy

nb. of sessions	nonce	size of mess.	secrecy
bounded	bounded	bounded	decidable
bounded	bounded	unbounded	NP-complete
unbounded	bounded	bounded	DEXPTIME
unbounded	bounded	unbounded	undecidable
unbounded	unbounded	bounded	undecidable

- [A.W. Roscoe'95, S. Schneider'96, J.C. Mitchell, M. Mitchell and U. Stern'97,Clarke, jha & Morrero'98]
- [Rusinowich&Turuani'01,Boreale'01, Amadio, Lugiez and Vanackère'01, Huima'99]
- [Durgin, Lincoln, Mitchell and Scederov'99]
- [Amdio, Lugiez and Vanackère'01, Comon, Cortier and Mitchell'01]

Partial Decision Methods

All these methods are based on fixpoint computation:

D. Monniaux'99, J. Goubault-Larecq'00 Tree Automate for representing the intruder knowledge, forward analysis.

B. Blanchet'01 Without nonces. Backward analysis, abstraction of Prolog programs.

Cortier, Mitchell, Ruess'01 Backward, termination not garanteed.

- Bozga, Lakhnech, Perin'03 symbolic fixpoint computation, termination guaranteed (widening and approximation).

Our Contributions - 1

Bounded number of sessions: NP-complete decision procedure

- secrecy, authentication (aliveness, weak agreement, agreement) and other prop.
- time sensitive cryptographic protocols
- unbounded initial intruder knowledge
- unbounded size of messages, but atomic keys

Bozga, Ene and Lakhnech in:

- FOSSACS'04
- CONCUR'04
- Journal of Logic and Algebraic Programming'05

Our contributions - 2

Opacity properties

Laurent Mazaré in

- Workshop on Issues in the Theory of Security (WITS'04)
- Workshop on Automated Reasoning for Security Protocol Analysis (ARSPA'04)
- The second international Workshop on Formal Aspects in Security and Trust (FAST'04)
- The third international Workshop on Formal Aspects in Security and Trust (FAST'05)

Our Contributions - 3

Unbounded number of sessions: Partial decision method

- combining the approach for bounded with abstract interpretation techniques
- unbounded initial intruder knowledge
- secrecy properties
- unbounded size of messages, but atomic keys

Bozga, Lakhnech and Perin in:

- TACAS'03
- CAV'03
- Journal STTT

Some practical results

Yahalom	OK	13.81
Needham-Schroeder Public Key	Attack	0.04
Needham-Schroeder Public Key (with key server)	Attack	0.90
Needham-Schroeder-Lowe	OK	0.03
Otway-Rees	OK	0.01
Kao-Chow_1	OK	0.79
Kao-Chow_2	OK	16.89
Neumann-Stubblebine	OK	0.04
Needham-Schroeder Symmetric Key	Don’t know	0.08
TMN	Attack	2.78
TMN-Lowe	Attack	3.04
Denning-Sacco	Don’t know	0.04
ISO Symmetric Key One-Pass Unilateral Authentication	OK	0.01
ISO Symmetric Key Two-Pass Unilateral Authentication	OK	0.01
ISO Symmetric Key Two-Pass Mutual Authentication	OK	0.01
Andrew Secure RPC	OK	0.02

Not discussed

Spi-Calculus and behavioral equivalence based methods: C. Fournet, Abadi, Rogaway, ...

Typing-based methods: Gordon, ...

Deduction based methods: L. Paulson. C. Meadows, J. Millen, ...

Symbolic and complexity-theoretic approaches

computational	Symbolic
Messages	
bitstrings	terms
Cyphering	
$\hat{m}' \stackrel{R}{\leftarrow} \mathcal{E}(\hat{m}, k)$	$\{m\}_k$ a term
Nonce	
randomly generated value	fresh name
Intrus	
TM prob. poly.	Système d'inférence
Attaque	
Pro. negligible	non-existence
more realistic semantics	automatic verification

Relation between the symbolic and complexity-theoretic models

Suppose that we proved a property of our protocol Π using a symbolic verification method.
Can we conclude anything on the computational behavior?

- If the cryptographic primitives are provably secure then the symbolic model is a safe abstraction of the computational model.
- Safe abstraction: Any computational attack has a symbolic abstraction except for negligible probability.
- A seminal paper: M. Abadi and P. Rogaway, Reconciling two views of cryptography. IFIP TCS'00.
  Deals with passive Intruder, Symmetric Keys, Encryption of Keys, Strong Secrecy

General proof Idea

Proof by reduction!

- Given an adversary $\mathcal{A}$ that breaks the protocol
- Build an adversary $\mathcal{B}$ that breaks the cryptographic primitives; $\mathcal{B}$ uses $\mathcal{A}$ as a sub-routine.

Provable Cryptography

Originated in the seminal paper Goldwasser-Micali'84: Probabilistic encryption.
Main idea:

1. Formal definitions of what a cryptographic primitive should do!
2. Proofs relative to problems assumed hard.
   Examples:
   - Discrete logarithm: given g^x compute x
   - Computational Diffie-Hellman: given g^x, g^y compute g^{xy}
   - Decisional Diffie-Hellman: distinguish (g^x, g^y, g^{xy}) and (g^x, g^y, g^r), where r is random.
   - RSA: given $n = pq$, p and q prime numbers, given $e \in \mathbb{Z}^*_{\varphi(n)}$ and $y = x^e mod\, n$ with $x \in \mathbb{Z}^*_n$, compute $y^d mod\, n$, where $ed \equiv_{\varphi(n)} 1$.

Indistinguishability

Game Adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$:

1. A pair od keys is computed: $(pk, sk) \stackrel{R}{\leftarrow} \mathcal{K}(\eta)$.
2. The adversary is given the public key, she computes two messages: $(m_0, m_1) \stackrel{R}{\leftarrow} \mathcal{A}_1(pk)$.
3. A bit b is chosen randomly: $b \stackrel{R}{\leftarrow} \{0,1\}$ and $c = \mathcal{E}(m_b, pk)$ is given to the adversary.
4. The adversary guesses b: $b' \stackrel{R}{\leftarrow} \mathcal{A}_2(c)$.

Advantage of $\mathcal{A}$:

$$Pr[b = b'] - \frac{1}{2}$$

Attacker's model

A strict hierarchy of security definitions is obtained according to whether $\mathcal{A}$ has access to $\mathcal{D}(\cdot, sk)$:

- IND-CPA, Chosen-Plaintext Attack: no access to $\mathcal{D}(\cdot, sk)$.
- IND-CCA1, Non-Adaptive Chosen-Cyphertext Attack: access to $\mathcal{D}(\cdot, sk)$ *only* before the challenge
- IND-CCA2, Adaptive Chosen-Cyphertext Attack: access to $\mathcal{D}(\cdot, sk)$ any time

Unforgeability

Similar definitions for signature and symmetric encryption.
Signature:

- Adversary tries to forge a valid message-signature pair without the private key.
- adversary is successful, if the following probability is not negligible:

$$\Pr[\mathcal{V}(m,\sigma) = 1 \mid (m,\sigma) = \mathcal{A}(K_v)]$$

Symmetric encryption: Indistinguishability + Unforgeability

State of the art

- D. Micciancio, B. Warinschi'04: Active Intruder, Asymmetric Keys, Authentication
- P. Laud'04: Active Intruder, Symmetric Keys, Strong Secrecy
- M. Backes et All'04: Active Intruder but no automation
- L. Mazaré, Y. Lakhnech and R. Janvier'05: Active Intruder, Asymmetric Keys, Symmetric Keys, Hashing and Signature all combined + some equational theories
- L. Mazaré and Y. Lakhnech'05: + Diffie-Hellman key exchange (modular exponentiation)
- L. Mazaré and Y. Lakhnech'05: opacity and e-voting/passive adversaries.

Concluding remarks

- An autopsy of the symbolic approach.
- A need for a rigorous approach for cryptographic primitives and protocols:
  - A *spealized* programming language/calculus for describing the algorithms.
  - A proof theory for this language: invariants, polynomial-time termination.
  - A proof theory to prove bounds on probabilities.

Other work on security in our group

- Hermes: a tool for the automatic verification of cryptographic protocols (L. Bozga)
- Verification of cryptographic API's (J.C. Courant, J.F. Monin, Y.L.).
- Formal model-based testing of security policies (J.C. Fernandez, Laurent Mounier; collaboration with LSR).
- Certification of SmartCard applications (M. Périn).