

Symbolic Reachability Analysis Using Narrowing and its Application to Protocol Analysis

**Jose Meseguer and Prasanna Thati
University of Illinois at Urbana Champaign**

June 2004

Outline of the Talk

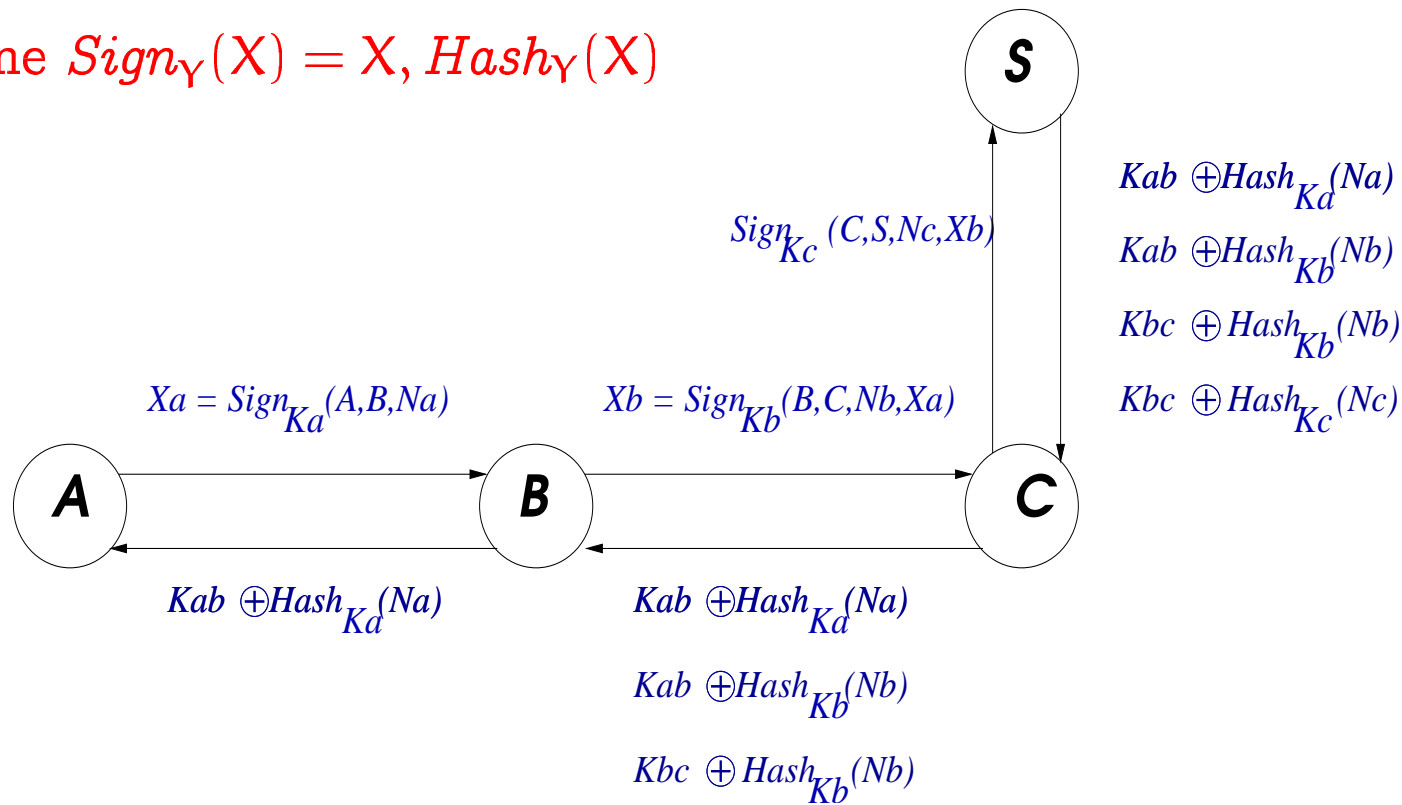
- **Perfect cryptography assumption** and its pitfalls.
 - Does not account for algebraic properties of cryptographic primitives.
- **Narrowing modulo equations**: A symbolic analysis technique for infinite state systems.
- Application of the technique to analysis of security protocols
 - **Relax the assumption** of perfect cryptography by accounting for algebraic properties of cryptographic primitives.

The Perfect Cryptography Assumption

- *There is no way to obtain knowledge about the plaintext pertaining to a ciphertext without knowing the key.*
- *There is enough redundancy in the cryptosystem that a cipher-text can only be generated using encryption with the appropriate key.*
- **Advantage:** Helps abstract away the cryptosystem and analyze the protocol itself.
- **Drawback:** An attack that takes advantage of a particular property of a specific cryptosystem cannot be found.

A Recursive Authentication (Bull's) Protocol

Define $Sign_Y(X) = X, Hash_Y(X)$



An Attack on Bull's Protocol

- The protocol's objective is to establish pairwise session keys K_{ab} between A and B, and K_{bc} between B and C.
- In particular, K_{ab} should be a **secret** from C.
- But C can **derive** K_{ab} from the messages it receives from the server S.

$$\begin{aligned}K_{ab} \oplus Hash_{K_b}(N_b) \oplus K_{bc} \oplus Hash_{K_b}(N_b) &= K_{ab} \oplus K_{bc} \\K_{bc} \oplus Hash_{K_c}(N_c) \oplus Hash_{K_c}(N_c) &= K_{bc} \\K_{ab} \oplus K_{bc} \oplus K_{bc} &= K_{ab}\end{aligned}$$

- If the algebraic properties of XOR encryption are ignored, then one can show that the secrecy is **not** violated.

Algebraic Properties of Some Cryptographic Primitives

- XOR Encryption:

$$(M_1 \oplus M_2) \oplus M_3 = M_1 \oplus (M_2 \oplus M_3)$$

$$M_1 \oplus M_2 = M_2 \oplus M_1$$

$$0 \oplus M = M$$

$$M \oplus M = 0$$

- Diffie-Hellman Encryption: Commutativity, associativity, and identity laws for multiplication operator, and

$$\text{Exp}(t, 1) = t$$

$$\text{Exp}(\text{Exp}(t, t'), t'') = \text{Exp}(t, t'.t'')$$

- Block Encryption:

$$\{(M_1, M_2)\}_k = (\{M_1\}_k, \{M_2\}_k)$$

Bounded-Process Security Protocol Analysis

- **Undecidability:** Security properties such as secrecy and authenticity are in general undecidable, but are **decidable** if the number of protocol sessions is **bounded**.
- **State of the art:** An **adhoc** procedure for each cryptographic primitive with a different set of algebraic properties.
- **Our proposal:** A symbolic analysis technique that can **uniformly handle** a broad class of algebraic properties.

Modeling Security Protocols as Rewrite Theories

- Model a protocol as a **rewrite theory** $\mathcal{R} = (\Sigma, E, R)$, where
 - Σ : is a signature of operators (message constructors)
 - E : is a set of equations of form $t = t'$ (algebraic properties)
 - R : is a set of rewrite rules of form $l \rightarrow r$ (execution rules)
- The equivalence classes of **ground terms** modulo E describes the system's **state space**.
- The rewrite relation $\rightarrow_{R/E}$ over terms that is defined modulo E describes the system **dynamics**.

Encoding Safety Properties as Reachability Goals

Safety properties such as secrecy and authenticity can be characterized as a **reachability goal** of the form

$$(\exists \vec{x}') t_1(\vec{x}') \rightarrow t'_1(\vec{x}') \wedge \dots \wedge t_n(\vec{x}') \rightarrow t'_n(\vec{x}')$$

- The terms t_i and t'_i denote **sets of states**.
- t_i : **initial** states, t'_i : **undesirable** states (leaked secret etc).
- Violation of the safety property is the same as **solvability** of the goal, i.e. existence of a substitution σ such that

$$\sigma(t_i) \rightarrow_{R/E}^* \sigma(t'_i)$$

Solving the Goals by Narrowing Modulo Equations

- We generalize **narrowing modulo equations** from a technique for solving unification problems in equational theories to one for solving reachability problems in rewrite theories.
- A **complete semi-decision** procedure under reasonable assumptions about the rewrite theory. Covers a broad class of infinite state systems including bounded process security protocols.
- Can be readily **automated**, and complements other approaches such as abstraction, theorem-proving, and tree-automata based approximations to reachability problems.

The Protocol Model

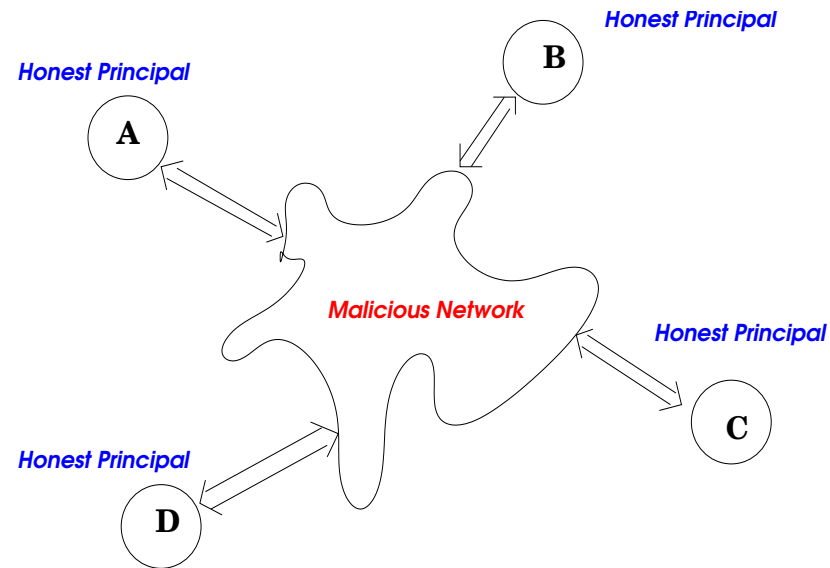
- A protocol specifies a **finite** list of actions for each of the finite number of participating principals.
- An action $M_1 \Rightarrow M_2$ is interpreted as: “on receiving a message matching the pattern M_1 , send the corresponding message M_2 ”

$$M ::= Var \mid Atoms \mid (M_1, M_2) \mid \{M\}_k \mid M_1 \oplus M_2$$

- Example action: Principal A challenging the identity of another principal who claims to be x

$$\{x\}_{pb(A)} \Rightarrow \{N_A\}_{pb(x)}$$

The Attacker Model



The attacker has complete control over the network. It can

- **Observe** every message that is exchanged.
- **Replace** a message sent by an honest principal with a fake message that it can **build**.

The Dolev Yao Inference System

$$(Ax) \quad K, M \vdash M$$

$$(Xor) \quad \frac{K \vdash M_1 \quad K \vdash M_2}{K \vdash M_1 \oplus M_2}$$

$$(Decr) \quad \frac{K \vdash \{M\}_k \quad K \vdash k^{-1}}{K \vdash M}$$

$$(Eq) \quad \frac{K \vdash M_1 \quad M_1 =_E M_2}{K \vdash M_2}$$

The judgment $K \vdash M$ is read as: “an attacker that knows all the messages in the set K can construct the message M ”.

E contains the following equations for xor

$$\begin{array}{ll} (M_1 \oplus M_2) \oplus M_3 & = \quad M_1 \oplus (M_2 \oplus M_3) & 0 \oplus M & = \quad M \\ M_1 \oplus M_2 & = \quad M_2 \oplus M_1 & M \oplus M & = \quad 0 \end{array}$$

Security Properties as Solvability of Constraints

- The violation of safety properties such as secrecy and authenticity can be reduced to solving a constraint set of the form

$$K_1 \vdash M_1 \wedge \dots \wedge K_n \vdash M_n$$

in the Dolev-Yao inference system.

- A substitution σ such that

$$\sigma(K_1) \vdash \sigma(M_1) \wedge \dots \wedge \sigma(K_n) \vdash \sigma(M_n)$$

represents an attack that violates the property.

Modeling the Constraint System as a Rewrite Theory

- Represent this constraint system as a rewrite theory.
- Represent the given set of constraints as a reachability goal.
- Use narrowing to find solutions (attacks) of the reachability goal.

Modeling the Constraint System as a Rewrite Theory

- The signature Σ includes constructors for messages and judgements.
- The equations E include the algebraic laws for cryptographic primitives.
- A Dolev-Yao inference rule

$$\frac{K \vdash M_1 \quad K \vdash M_2}{K \vdash M_3}$$

is modeled as the rewrite rule

$$K \vdash M_3 \rightarrow K \vdash M_1 \wedge K \vdash M_2$$

that rewrites multisets of judgments.

- The E_q rule is **implicit** since rewriting is defined modulo equations.

Modeling the Constraints as Reachability Goals

- Then a substitution σ is a solution of the constraints

$$K_1 \vdash M_1 \wedge \dots \wedge K_n \vdash M_n$$

if and only if it is a solution of the reachability goal

$$K_1 \vdash M_1 \rightarrow true \wedge \dots \wedge K_n \vdash M_n \rightarrow true$$

- Narrowing modulo equations will enumerate a **complete** set of solutions, i.e. it will find all the attacks.

Narrowing the Judgements

- Given
 - A judgement J .
 - A rule $J_1 \rightarrow J_2 \wedge J_3$
 - A **E-unifier** σ for J and J_1 , i.e. $\sigma(J) =_E \sigma(J_1)$
- Then define $J \overset{\sigma}{\rightsquigarrow} \sigma(J_2) \wedge \sigma(J_3)$

LEMMA $J \overset{\sigma}{\rightsquigarrow} J' \wedge J''$ implies $\sigma(J) \rightarrow J' \wedge J''$.

Narrowing the Reachability Goals

- If $J_i \xrightarrow{\sigma} J' \wedge J''$ then for

$$G : J_1 \rightarrow true \wedge \dots \wedge J_i \rightarrow true \wedge \dots \wedge J_n \rightarrow true$$

define $G \xrightarrow{\sigma} G'$, where

$$G' : \sigma(J_1) \rightarrow true \wedge \dots \wedge J' \rightarrow true \wedge J'' \rightarrow true \wedge \dots \wedge \sigma(J_n) \rightarrow true$$

- The idea is that if ρ is a solution of G' then $\rho \circ \sigma$ is a solution of G .
- For $\sigma = \sigma_n \circ \dots \circ \sigma_1$ define $G \xrightarrow{\sigma^*} G'$ if $G \xrightarrow{\sigma_1} \dots \xrightarrow{\sigma_n} G'$.

A Complete Procedure for Solving Reachability Goals

THEOREM [Soundness] If $G \xrightarrow{\sigma}^* G'$ to a trivial goal

$$G' : true \rightarrow true \wedge \dots \wedge true \rightarrow true$$

then σ is a solution of G .

- This suggests an algorithm for solving reachability goals, that explores the **narrowing tree** starting from G .
- Since narrowing derivations may not terminate, the tree has to be expanded in a **fair** manner.

THEOREM [Completeness] The narrowing algorithm finds a complete set of solutions of any given goal G .

A Few Remarks

- Narrowing modulo equations provides a **uniform** procedure for verifying security protocols by also taking into account the algebraic properties of cryptographic primitives.
- It can be readily **automated**. Complements other techniques for verification of infinite state systems such as abstraction and theorem proving.
- Currently investigating narrowing **strategies** that will give that cut down the search space while retaining completeness, thus improving the efficiency of the narrowing procedure.