Narrowing Analysis of Crypto Protocols

Catherine Meadows, NRL, José Meseguer, UIUC
Prasanna Thati, CMU

HCSS'05, Baltimore, March 2005

Motivation

It is well-known that some protocols that had been proved secure under the usual Dolev-Yao perfect cryptography assumption have been broken using algebraic knowledge about the cryptographic functions.

For specific algebraic theories describing the properties of some given cryptographic functions, special-purpose methods have been devised to prove protocols secure even when an attacker uses such algebraic knowledge.

But can we have a general theoretical framework in which we can analyze different protocols assuming in each case different cryptographic functions and different algebraic theories for such functions?

Motivation (II)

In this talk we propose rewriting logic as such as general theoretical framework, so that crypto protocols are formally specified as rewrite theories.

We then define narrowing as a powerful symbolic reachability analysis method to reason about protocol security when an attacker knows the algebraic properties of the protocol's cryptographic functions.

We plan to combine the generality of the narrowing analysis with the powerful language-based state reduction techniques of the NRL Protocol Analyzer (NPA). In this way we will get: (i) more powerful analysis techniques; and (ii) a generalization of the NPA technology to handle different algebraic crypto properties beyond Dolev Yao.

Contents of the Talk

- Rewriting Logic
- Narrowing Reachability Analysis
- Applications to Security Protocol Analysis
- Relationships with the NRL Protocol Analyzer
- Related Work and Concluding Remarks

Rewriting Logic in a Nutshell

A rewrite theory $\mathcal{R}$ is a triple $\mathcal{R} = (\Sigma, E, R)$, with:

- (Σ, E) an equational theory with signature Σ and equations E

- R a set of labeled rewrite rules of the form $l : t \longrightarrow t' \ \ if \ \ cond$, with l a label, t and t' Σ-terms, and $cond$ a condition (i.e., a guard).

Intuitively, $\mathcal{R}$ specifies a concurrent system, whose states are elements of the initial algebra $T_{\Sigma/E}$ specified by (Σ, E), and whose concurrent transitions are specified by the rules R.

Crypto Protocols as Rewrite Theories

In particular, we can formally specify a cryptographic protocol $\mathcal{P}$ as a rewrite theory $\mathcal{R}_{\mathcal{P}} = (\Sigma_{\mathcal{P}}, E_{\mathcal{P}}, R_{\mathcal{P}})$ where:

- $\Sigma_{\mathcal{P}}$ is a signature specifying the different data types, functions, and state constructors used by the protocol;
- $E_{\mathcal{P}}$ is a set of equations specifying the algebraic properties of the functions used by $\mathcal{P}$, including those of its crypto functions;
- $R_{\mathcal{P}}$ specifies the transition rules of the protocol.

Narrowing Reachability Analysis

We propose narrowing as a general deductive procedure for solving reachability problems of the form

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

in a given rewrite theory.

- The terms t_i and t'_i denote sets of states.
- For what subset of states denoted by t_i are the states denoted by t'_i reachable?
- No finiteness assumptions about the state space.

Narrowing

- Generalize narrowing form a technique to solve equational goals to one for solving reachability goals.
- Provides a weakly complete semi-decision procedure under reasonable executability assumptions about the given rewrite theory.
- But strongly complete for several interesting classes of rewrite theories, including crypto protocol specifications.
- Complements other analysis techniques for infinite state systems, such as abstraction, theorem-proving, and tree-automata based reachability analysis.

Application: Bounded-Process Security Protocol Analysis

- Safety properties of security protocols, such as secrecy and authenticity can be characterized as reachability problems that are decidable when the number of protocol sessions is bounded.
- Algebraic properties of cryptographic primitives have to be taken into account during this analysis since they can be exploited to find attacks.
- So far an ad-hoc analysis procedures for each cryptographic primitive with different algebraic properties.
- Narrowing modulo equations provides a uniform procedure for a wide range of cryptograhic primitives with different algebraic properties.

More on Unconditional Rewrite Theories

Model a given system as a rewrite theory $\mathcal{R} = (\Sigma, E, R)$, where

Σ : is a signature of operators
E : is a set of equations of form $t = t'$
R : is a set of rewrite rules of form $l \to r$

- The relations $=_E$ and $\to_{R/E}$ over terms are defined as usual.
- The equivalence classes of $=_E$ over ground terms describe a system's state space.
- The relation $\to_{R/E}$ describes the system dynamics.

Reachability Goals

Reachability goal G

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

- A substitution σ is a solution of G if for all $1 \leq i \leq n$, we have

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

- A substitution σ is a trivial solution of G if for all $1 \leq i \leq n$, we have

$$\sigma(t_i) =_E \sigma(t'_i)$$

i.e., σ is an E-unifier for the system of equations

$$t_1 = t'_1\ \wedge \ldots \wedge\ t_n = t'_n$$

Solutions of a Reachability Goal

A set of substitutions S is said to be a complete set of solutions of G if

- Every $\sigma \in S$ is a solution of G, and
- For any solution ρ of G there is $\sigma \in S$ and some η such that $\rho|_{Var(G)} =_E (\eta \circ \sigma)|_{Var(G)}$, i.e there is a.

We are interested in finding a complete set of solutions for a given goal G in an unconditional rewrite theory $\mathcal{R}$.

"Implementing" the Relation $\rightarrow_{R/E}$ as $\rightarrow_{R \cup \Delta, B}$

- $\rightarrow_{R/E}$-reducibility is undecidable in general. But if the rules R are coherent with the equations E $\rightarrow_{R/E}$-rewriting becomes decidable.
- We assume $E = \Delta \cup B$ such that Δ is convergent modulo B, and B has a finite and complete unification algorithm.
- Define $t \rightarrow_{R \cup \Delta, B} t'$ if
  - a position ω in t
  - an equation $l = r$ in Δ, or a rule $l \rightarrow r$ in R, and
  - a substitution σ such that for $t|_\omega =_B \sigma(l)$

  such that $t' = t[\omega \leftarrow \sigma(r)]$.
- $\rightarrow_{R \cup \Delta, B}$-reducibility is decidable.

$R \cup \Delta, B$-Narrowing

Define $t \overset{\sigma}{\mapsto}_{R \cup \Delta, B} t'$ if there is

- a position ω in t
- an equation $l = r$ in Δ or a rule $l \to r$ in R, and
- a B-unifier σ for $t|_\omega = l$

such that $t' = \sigma(t[\omega \leftarrow r])$.

LEMMA $t \overset{\sigma}{\mapsto}_{R \cup \Delta, B} t'$ implies $\sigma(t) \to_{R \cup \Delta, B} t'$.

Narrowing the Goals

- For goals

$$G \;:\; t_1 \to t_2 \;\wedge \ldots \wedge\; t_{2n-1} \to t_{2n}$$
$$G' \;:\; t'_1 \to t'_2 \;\wedge \ldots \wedge\; t'_{2n-1} \to t'_{2n}$$

  we define $G \overset{\sigma}{\mapsto}_{R\cup\Delta,B} G'$ if there is an odd i such that
  - $t_i \overset{\sigma}{\mapsto}_{R\cup\Delta,B} t'_i$, and
  - $j \neq i$ implies $t'_j = \sigma(t_j)$.
- Note that only the left side of each conjunct is narrowed.
- The idea is that is $G \overset{\sigma}{\mapsto}_{R\cup\Delta,B} G'$ and ρ is a solution of G' then $\rho \circ \sigma$ is a solution of G.
- For $\sigma = \sigma_n \circ \ldots \circ \sigma_1$ define $G \overset{\sigma}{\mapsto}^*_{R\cup\Delta,B} G'$ if $G \overset{\sigma_1}{\mapsto}_{R\cup\Delta,B} \ldots \overset{\sigma_n}{\mapsto}_{R\cup\Delta,B} G'$.

Solving Reachability Goals by Narrowing

THEOREM [Soundness] If $G \overset{\sigma}{\mapsto}{}^{*}_{R \cup \Delta, B} G'$ and ρ is a trivial solution of G', then $\rho \circ \sigma$ is a solution of G.

This gives us an algorithm for solving reachability goals, that explores the narrowing tree starting from G.

THEOREM [Weak Completeness] The narrowing algorithm finds a set of solutions that is complete with respect to R/E-normalized solutions σ, i.e. σ such that $\sigma(x)$ cannot be rewritten for any x.

But the algorithm may not find substitutions that are not R/E-normalized.

Some Strong Completeness Results

- Narrowing algorithm is complete for the case of topmost rewrite theories, i.e. where terms can be rewritten only at the topmost position (root).
- Narrowing is also complete for theories that are topmost modulo an associative, commutative operator.
- This covers a number of practical cases, including
  - Concurrent object-based systems, including crypto protocols
  - Petri nets
  - Context free grammars
- Strong completeness results for linear rewrite theories and goals, and for back-and-forth narrowing.

Application: Bounded Process Security Protocol Analysis

- 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"*

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

- An attacker can attempt to break the protocol by replacing messages sent by honest principals with fake messages that it can construct using the messages it has observed so far.

The Dolev Yao Inference System with XOR (Sample)

$$\textit{(Ax)} \quad K, M \vdash M$$

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

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

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

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`

$$
\begin{array}{rcl@{\qquad\qquad}rcl}
(M_1 \oplus M_2) \oplus M_3 & = & M_1 \oplus (M_2 \oplus M_3) & 0 \oplus M & = & M \\
M_1 \oplus M_2 & = & M_2 \oplus M_1 & M \oplus M & = & 0
\end{array}
$$

This set of equations is known to have a unification algorithm.

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

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

  in the Dolev-Yao inference system.

- A substitution σ such that

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

  represents an attack that violates the property.

Modeling the Dolev-Yao System as a Rewrite Theory

- Represent the Dolev-Yao inference system as a rewrite theory that is topmost modulo associativity, commutativity and identity axioms of $\wedge$

- Represent the given set of constraints as a reachability goal.

- Exploit the topmost completeness results for narrowing to find all solutions of the reachability goal.

Modeling the Dolev-Yao System as a Rewrite Theory

- The equations E include the algebraic laws for cryptographic primitives (for example, `xor`)

- A Dolev-Yao inference rule

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

  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 `Eq` 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

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

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

$$K_1 \vdash M_1 \rightarrow \mathit{true} \ \wedge \ \ldots \wedge K_n \vdash M_n \rightarrow \mathit{true}$$

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

A Variant of Lowe's Fix for NSPK Protocol

Consider the following variant of Lowe's fix ($\oplus$ vs. pairing)

1. $A \to B \;:\; \{(N_A, A)\}_{pb(B)}$
2. $B \to A \;:\; \{(N_A \oplus B, N_B)\}_{pb(A)}$
3. $A \to B \;:\; \{N_B\}_{pb(B)}$

This is represented in the protocol model as

FixedInitiator(A, B, N_A):

(I_1) $\Rightarrow \{(N_A, A)\}_{pb(B)}$

(I_2) $\{(N_A \oplus B, X_2)\}_{pb(A)} \Rightarrow \{X_2\}_{pb(B)}$

FixedResponder(A, B, N_B):

(R_1) $\{(X_1, A)\}_{pb(B)} \Rightarrow \{(X_1 \oplus B, N_B)\}_{pb(A)}$

(R_2) $\{N_B\}_{pb(B)} \Rightarrow$

Analysing the Protocol Using Narrowing

- Consider a protocol instance with three principals

$$\begin{array}{lcl} a & : & \mathit{Initiator}(a, c, n_a) \\ b & : & \mathit{Responder}(a, b, n_b) \\ c & : & \text{dishonest principal} \end{array}$$

- The nonce n_b is to be kept secret from the intruder c.
- The ordering of actions I_1, R_1, I_2, R_2 generates the following constraints

$$K_0, \{(n_a, a)\}_{\mathit{pb}(c)} \vdash \{(X_1, a)\}_{\mathit{pb}(b)}$$
$$K_0, \{(n_a, a)\}_{\mathit{pb}(c)}, \{(X_1 \oplus b, n_b)\}_{\mathit{pb}(a)} \vdash \{(n_a \oplus c, X_2)\}_{\mathit{pb}(a)}$$
$$K_0, \{(n_a, a)\}_{\mathit{pb}(c)}, \{(X_1 \oplus b, n_b)\}_{\mathit{pb}(a)}, \{X_2\}_{\mathit{pb}(c)} \vdash \{n_b\}_{\mathit{pb}(b)}$$
$$K_0, \{(n_a, a)\}_{\mathit{pb}(c)}, \{(X_1 \oplus b, n_b)\}_{\mathit{pb}(a)}, \{X_2\}_{\mathit{pb}(c)} \vdash n_b$$

Analysing the Protocol Using Narrowing

Narrowing finds the solution

$$\sigma = \{n_a \oplus b \oplus c/X_1, n_b/X_2\}$$

which corresponds to the following attack

$$\begin{array}{lcl}
1.\ a \to c & : & \{(n_a, a)\}_{pb(c)} \\
2.\ c(a) \to b & : & \{(n_a \oplus b \oplus c, a)\}_{pb(b)} \\
3.\ b \to c(a) & : & \{(n_a \oplus b \oplus c \oplus b, n_b)\}_{pb(a)} \\
4.\ c \to a & : & \{(n_a \oplus b \oplus c \oplus b, n_b)\}_{pb(a)} \\
5.\ a \to c & : & \{n_b\}_{pb(c)}
\end{array}$$

that critically makes use of the equality $n_a \oplus b \oplus c \oplus b = n_a \oplus c$

An Alternative Approach

Narrowing with rules modulo equational axioms can be applied to many different rewrite theories. The example just given illustrates its application to "Dolev-Yao" theories of the form $DY(B)$, which are parameterized by the algebraic properties B of the cryptographic properties of interest (in our example the `xor` equations).

Alternatively, we can formalize the actions of agents in a given protocol P (that may use crytographic functions satisfying some axioms B) as well as the capabilities of an intruder that can make use of the B properties, as the rules of a rewrite theory $\mathcal{R}_P$ which will also be topmost. Therefore, narrowing with $\mathcal{R}_P$ will be strongly complete and will provide a semi-decision procedure for finding security attacks in P.

The NRL Protocol Analyzer (NPA)

- A formal methods tool for the analysis of cryptographic protocols using standard Dolev-Yao model of intruder
- Written in Prolog, uses many techniques from logic programming
- User specifies insecure state using a combination of constants and existentially quantified variables
  - NPA works backwards from state to find path to it
- Starts with infinite search space
  - User proves a set of lemmas to cut down size
  - When a state is generated, lemmas used to determine whether it should be kept or discarded

Why We Are Interested in NPA

- Basic structure, using logic programming and narrowing, similar to rewriting framework
- Sophisticated automated techniques for generating lemmas give it great power
- Techniques have been validated by its application to a wide variety of substantial-sized protocols
  - e.g., IKE, IKEv2, Group Domain of Interpretation

What NPA Specifications Look Like

- Intruder and honest protocol rules represented same way
- Each protocol transition represented by Prolog-style clause
  - Main difference: conclusions may be the conjunction of a number of literals, instead of a single literal
- Terms in clauses made up variables and constants
- Terms in clauses obey rewrite rules describing behavior of encryption functions
- Uses narrowing in different way from rewriting framework, but basic structure of transition rules very similar

An Example NPA Lemma Generation Technique: Languages

- Start with a single seed term
- NPA automatically generates infinite set of terms from seed term and simultaneously proves that intruder can never learn any term in that set
- Generation of seed terms not automated, but have a set of standard heuristics for producing basic ones
- Languages have been shown to be related to other inductively defined structures used in crypto protocol analysis
  - e.g. strand space ideals, Schneider's rank functions

Strengths of NPA Languages

- Gives infinite set of terms intruder can't learn
  - Allows us to rule out infinite set of states that precede any given state
- In right circumstances, this puts a bound on number of states that need to be checked
  - Thus allows us to reduce unbounded session system to bounded system
- Have applied this technique effectively to a large number of real-life protocols

Limitations of NPA Languages

- Have not developed way of telling that we have succeeded in reducing to finite search space
    - Heuristics help, but ultimately we rely on trial and error to tell us when we have defined enough languages

Applying Language Techniques in Rewriting Framework

- Develop techniques similar to languages to prove that we can rule out infinite sets of constraints in a constraint-based protocol analysis on an unbounded system
- Prove theorems within rewrite rule framework that predict ability to reduce infinite set of constraints to finite set
- Generalize results to different inductively defined systems for different types of protocols for different types of equational theories
  - Rewriting framework, which parameterizes equational theories, has potential to be of great help here

Benefits NPA Techniques Could Bring to Rewriting Framework

- Benefit of experience of using NPA on real-world examples
  - IKE Version 1 and 2
  - Group Domain of Interpretation
- NPA techniques for reducing search space
  - Inductive techniques for reducing infinite to finite state space
  - Subsumption-based partial order reduction for cutting down size of finite space
  - A host of others

Benefits Rewriting Framework Could Bring to NPA Techniques

- Placing NPA state reduction techniques within narrowing framework has potential to provide help in developing a more rigorous understanding of the state reduction techniques

- Rewriting framework allows parametric treatment of different algebraic properties of cryptographic systems
  - Will allow us to extend NPA state reduction techniques in a natural way to other equational systems

Related Work and Conclusions

This is work in progress. We are planning to add narrowing capabilities to the Maude rewriting logic language and to experiment with a range of applications and case studies. We are also investigating how to best combine narrowing with the NPA state reduction techniques.

This work is related to:

1. work by Comon and Shmatikov, Millen and Shmatikov, and by Chevalier et al., extending security decision procedures to handle given algebraic properties of the cryptographic operations

2. work by Kapur, Meadows, Narendran, and Wang on unification algorithms for equational theories

axiomatizing modular exponentiation to analyze algorithms such as group Diffie-Hellman

3. The NPA as already discussed; formalisms such as CAPSL/CIL, MSR, and HLPSL/IF; and tools such as the CAPSL tools, OFMC, and the Maude/MSR tool.

Regarding (1), narrowing can provide a general framework for such, up to now ad-hoc extensions.

The work in (2) complements ours, in that with such methods we can build-in more of the equational constraint solving part and be more efficient.

The NPA and OFMC tools perform special-purpose kinds of narrowing that could be extended to narrowing modulo equational axioms for cryptographic functions using the methods presented here.

Regarding (1) and (3) a very interesting topic is investigating termination conditions for the narrowing procedure (with strategies) over appropriate rewrite theories. Approaches may include detecting loops, bounding the size of solutions, etc. Generalization of the NPA techniques for this purpose is at the top of our research agenda.

We expect that combining the genericity of narrowing to handle different crypto algebraic theories with the power of the NPA state reduction techniques will yield a very versatile protocol analysis methodology on which to base next-generation formal analysis tools.