Intruder deduction problem in presence of guessing attacks

Stéphanie Delaune
École Normale Supérieure de Cachan
Laboratoire Spécification et Vérification
61, avenue du Président Wilson
94 235 Cachan Cedex - France
delaune@lsv.ens-cachan.fr

May 28, 2003

Abstract

We present a decidability result in the context of the verification of cryptographic protocols in presence of data which take value in a finite known set. Since the perfect cryptography assumption is unrealistic for cryptographic protocols that employ weak data, we extend the conventional Dolev-Yao model to consider guessing attacks, where an intruder guesses the values of weak data and verify these guesses. We show that the intruder deduction problem, i.e. the existence of guessing attack, can be decided in polynomial time for the extended Dolev-Yao model.

1 Introduction

While the automatic verification of cryptographic protocols is undecidable, even with several restrictions, it has obtained a lot of attention these last years. In particular, the *intruder deduction problem*, which corresponds to the security decision problem in presence of a passive eavesdropper, is a significant question to the verification problem as well as to the search for attacks.

In most approaches, the underlying cryptographic primitives are based on the so called "Dolev-Yao" model [3]. This model is justified by the *perfect cryptography assumption*, that there is no way to obtain knowledge about the plaintext encrypted in a ciphertext without knowing the key.

This abstraction happened to be accurate in many works concerned with the search of attacks or the proof of cryptographic protocols, but it may be too strong in some particular situations. For instance, when we want to take into account attacks based on algebraic properties [2, 1], or, like in this paper, so called guessing attacks [4, 5].

Example 1 *Consider the following naive vote protocol:*

$$A \quad \rightarrow S: \quad \{m\}_{pub(S)}$$

A encrypts its vote m with the public key of the vote server S. The server decrypts the message with its private key. The requirement is that, only A and S know m.

m cannot be deduced using the standard Dolev-Yao model. However, if we assume that m belongs to a finite set $\mathcal{D}$ known to an intruder, then m can be computed: the attacker can encrypt all the possible values of m with $pub(S)$, he obtains this way a set of values $\{\{m'\}_{pub(S)} | m' \in \mathcal{D}\}$ that he can compare with $\{m\}_{pub(S)}$, which was already intercepted. When the eavesdropper finds the computed message which matches the intercepted message, he gets m (assuming injectivity of encryption).

Example 2 *Consider the two-messages handshake transaction (see [4]), which is often used in protocols:*

$$\begin{aligned} A \quad &\rightarrow B: \quad \{n\}_k \\ B \quad &\rightarrow A: \quad \{n+1\}_k \end{aligned}$$

A generates a random number n and encrypts it with a predetermined secret symmetric key that is shared between A and B. B decrypts the message,

computes $n+1$, and encrypts the result before returning it to A. The cryptosystem is symmetric. In the standard Dolev-Yao model, an intruder cannot get k nor n.

But, if we assume that k is weak, i.e. k is a value in a finite set known by the intruder, then the protocol is vulnerable: the intruder tries to decrypt both messages with a possible value for the key k, he obtains two values. If the second is the increment of the first, then the attacker has guessed the correct value of k (assuming standard properties of the cryptosystem).

The presence of weak data can allow the intruder to do guessing attacks. We shall use the definition of guessing attacks from [5] which generalizes the definition of [4]:

A guessing attack consists of the intruder guessing a value g, and then verifying it. The verification will be by the intruder using g to produce a value v, which we call the *verifier* and can take a number of different forms:

1. the intruder knew v initially. (cf. example 1)
2. the intruder produced v in two distinct ways from g. (cf. example 2)
3. v is an asymmetric key, and the intruder knows v's inverse from somewhere.

The main contribution of this paper is the formalization of such attacks (following the lines of [5]) and a proof that the intruder deduction problem is still in PTIME.

2 Intruder deduction problem

We assume that messages are terms built over a given alphabet $\mathcal{F}$ of function symbols containing constants, pairing $\langle \_, \_ \rangle$, encryption $\{\_\}\_$, and a unary symbol $\_^{-1}$.

Among these constants symbols, we distinguish constants keys. We consider symmetric key as well as asymmetric or public-key systems. A key k is symmetric if $k^{-1} = k$ and we assume that composed keys are symmetric.

We consider the congruence generated by the equation $x^{-1^{-1}} = x$. If we orient from left to right this equation, we get a convergent rewrite system. Hence every term t has a unique normal form.

To formalize the *intruder deduction problem*, we shall distinguish in the intruder's knowledge the "strongly known" messages (or "known" for short) and the "weakly known" messages.

Intuitively, the first ones are messages that the intruder knows exactly.

The second ones are messages which take their values in a finite set, known to the intruder, so the attacker can pick a value in the set and, if he has enough information, he can verify whether his guess is correct or not. If the guess is correct, we can assume that the message is "strongly known" by the intruder.

Definition 1 *(Atomic term)*
An atomic term is a constant, or the inverse k^{-1} of a constant key symbol k.

We formulate the *intruder deduction problem* in the following way:

Given a finite set of "strongly known" messages T, a finite set of atomic "weakly known" data T' and a (presumably) secret s, can the intruder deduce s from T and T'.

We introduce a new model to represent the intruder capabilities, and we prove a decidability theorem for the new set of deduction rules.

3 Extended Dolev-Yao model

In this section, we describe how guessing attacks can be modeled by adapting the standard intruder model.

The new model, called extended Dolev-Yao model, is presented in figure 1. We introduce two forms of sequents:

- $T/T' \vdash u$ means that if the intruder "strongly knows" messages in T and "weakly knows" atomic messages in T', he can (strongly) deduce the message u.
- $T/T' \vdash' u$ means that if the intruder "strongly knows" messages in T and "weakly knows" messages in T', he can weakly deduce the message u. In other words, he can deduce that u belongs to a finite set that he can compute.

So, the rules (A, P, UL, UR, E, D) represent the capacity of the intruder to do strong deduction from strong hypothesis, whereas the rules (A′, P′, UL′, UR′,

$$\text{Axiom (A)}\quad \frac{u \in T}{T/\emptyset \vdash u} \qquad \text{Pairing (P)}\quad \frac{T/T_1' \vdash u \;\; T/T_2' \vdash v}{T/T_1', T_2' \vdash \langle u, v\rangle}$$

$$\text{Unpairing (UL)}\quad \frac{T/T' \vdash \langle u, v\rangle}{T/T' \vdash u} \qquad \text{Unpairing (UR)}\quad \frac{T/T' \vdash \langle u, v\rangle}{T/T' \vdash v}$$

$$\text{Encryption (E)}\quad \frac{T/T_1' \vdash u \;\; T/T_2' \vdash v}{T/T_1', T_2' \vdash \{u\}_v} \qquad \text{Decryption (D)}\quad \frac{T/T_1' \vdash \{u\}_v \;\; T/T_2' \vdash v^{-1}}{T/T_1', T_2' \vdash u}$$

$$\text{Axiom (A}')\quad \frac{}{T/u \vdash' u}\; u \; atomic\; term \qquad \text{Pairing (P}')\quad \frac{T/T_1' \vdash' u \;\; T/T_2' \vdash' v}{T/T_1', T_2' \vdash' \langle u, v\rangle}$$

$$\text{Unpairing (UL}')\quad \frac{T/T' \vdash' \langle u, v\rangle}{T/T' \vdash' u} \qquad \text{Unpairing (UR}')\quad \frac{T/T' \vdash' \langle u, v\rangle}{T/T' \vdash' v}$$

$$\text{Encryption (E}')\quad \frac{T/T_1' \vdash' u \;\; T/T_2' \vdash' v}{T/T_1', T_2' \vdash' \{u\}_v} \qquad \text{Decryption (D}')\quad \frac{T/T_1' \vdash' \{u\}_v \;\; T/T_2' \vdash' v^{-1}}{T/T_1', T_2' \vdash' u}$$

$$\text{Weakening (W)}\quad \frac{u \in T}{T/\emptyset \vdash' u}$$

$$\text{Compare (C)}\quad \frac{P_1\left\{\begin{array}{c}\vdots\\ \hline T/T_1' \vdash' u\end{array}\right. \text{(R1)} \qquad P_2\left\{\begin{array}{c}\vdots\\ \hline T/T_2' \vdash' v\end{array}\right. \text{(R2)}}{T/T_1', T_2' \vdash w}$$

where:

(i). $w \in T_1' \cup T_2'$

(ii). P_1 and P_2 are normal proofs

(iii). R1 $\neq$ R2 or $\{u, x_1, .., x_n\} \neq \{v, y_1, ..., y_n\}$

(iv). $R(u, v)$ where $R = Id \cup \{(k, k^{-1}) |$ k is a key $\}$

Figure 1: The extended Dolev-Yao intruder capabilities.

E', D') represent the capacity of the intruder to do weak deduction from weak hypothesis.

The weakening rule (W) expressed that strongly known messages are a special case of weakly known messages.

Last but not least, the rule (C) which mix the two forms of sequents is used to formalize the verification of guessed (weak) data w. *(cf. definition of guessing attacks given in introduction)*

The second condition, (P_1 and P_2 are normal proofs, which is mentioned to apply the rule (C)) is necessary to prevent certain false attacks. This condition will prohibit deduction steps that simply undo previous steps.

Definition 2 *(Normal proof)*
A proof P of $T/T' \diamond u$ is normal if there is no subtree of P whose root is labeled with $T/T'_1 \diamond_1 v$ and which contains itself a strict subtree whose root is labeled with $T/T'_2 \diamond_1 v$. $\diamond, \diamond_1 \in \{\vdash, \vdash'\}$.

Example 3 *We continue example 1.*
Assume that $T = \{\{m\}_{pub(S)}, pub(S)\}$ and that m is weak, then we have the derivation drawn in figure 2.

The intruder guesses a value for m, produces the verifier $\{m\}_{pub(S)}$, (*left subtree*) and verifies his guess since he knows the verifier initially.

Remark 1 *In a proof, there is at most one instance of the rule Compare per branch.*

4 Results

Our goal in this section is to show that the intruder deduction problem, that we can reformulate in the following way:

Given two finite sets of messages T and T', and a secret s, can we derive a proof of $T/T'_1 \vdash s$ such that $T'_1 \subseteq T'$?

can be decided in polynomial time. To show this result, we prove a locality theorem [6] for the new set of deduction rules.

If T is a finite set of terms, $St(T)$ is the set of subterms of terms in T. The number of elements in $St(T)$ is linear in the size of T (the size of a set of terms is defined as usual, as the sum of the number of nodes in each member of T).

Theorem 1 *(locality theorem) If there is a proof of $T/T' \vdash u$, then there is a normal proof of $T/T' \vdash u$ in which only subterms of terms in $T \cup T' \cup \{u\}$ appear.*

Proof:
We prove the following results simultaneously by induction on the size of the proof of $T/T' \vdash u$:

1. a normal proof of $T/T' \vdash u$ contains only terms in $St(T \cup T' \cup \{u\})$.
2. if the last inference rule of a normal proof of $T/T' \vdash u$ is a decomposition rule, $(\mathsf{A}, \mathsf{UL}, \mathsf{UR}, \mathsf{D}, \mathsf{A}', \mathsf{UL}', \mathsf{UR}', \mathsf{D}', \mathsf{W}, \mathsf{C})$, then this proof contains only terms in $St(T \cup T')$.

Consider all possible cases for the last inference:

- Assume that the last rule is (C), *(see fig 1)*

  We distinguish two cases:

  - $u = v$
    The proofs P_1 and P_2 can not end with the same instance of the same rule (iii), so we can assume (w.l.o.g) that P_1 ends with a decomposition rule and, by induction hypothesis (2), involves only terms in $St(T \cup T'_1)$. By induction hypothesis (1), the proof P_2 involves only terms in $St(T \cup T'_2 \cup \{v\})$. Since $v = u$, $u \in St(T \cup T'_1)$ and $w \in T'_1 \cup T'_2$ (i), we deduce that P involves only terms in $St(T \cup T'_1 \cup T'_2)$.
  - u is an asymmetric key and v its inverse
    Assume (w.l.o.g) that $v = u^{-1}$ and u, v are in normal form. The last inference rule of P_2 is necessarily a decomposition rule, so it is similar to the first case.

- The others cases are very similar.

Theorem 2 *The intruder deduction problem $T/T' \vdash s$, can be decided in polynomial time in the extended Dolev-Yao intruder model.*

Proof: *(sketch)*
In this inference system, the proofs have a very particular form, only the rules $(\mathsf{A}', \mathsf{UL}', \mathsf{UR}', \mathsf{D}', \mathsf{E}', \mathsf{W})$ are used until an instance of the rule Compare. Here, we deduce that a weak data is finally strongly known and after, we do strongly deduction as in a Dolev-Yao model. So, to solve the intruder deduction problem in presence of weak data, it is sufficient to:

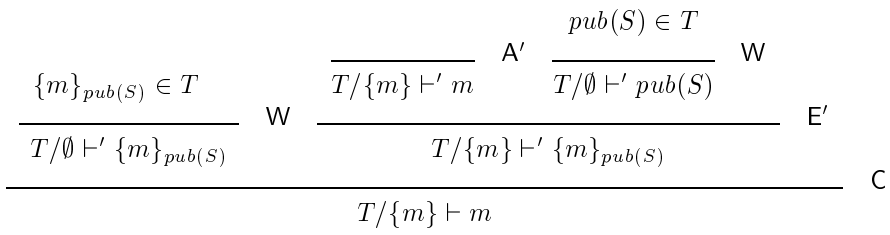

Figure 2: Example 3

- find among the weak data, those that the intruder can strongly deduce
- add these to the intruder knowledge, and then solve the intruder deduction problem in the standard Dolev-Yao model

The second point can be decided in polynomial time, this result can be easily derived from a theorem by McAllester [6]. For the first point:

- we code inference system of figure 1 as a set S of Horn clauses
- thanks to theorem 1, determine if there exists a proof of w is reducible to HORN-SAT for the (finite) set of instances of clause of S by terms of $St(T \cup T')$. The size of this set is polynomial.

5 Conclusion

We have extended the Dolev-Yao model to take into account guessing attacks and we have shown that the intruder deduction problem is still PTIME in presence of weak data.

This work can be extended to solve the reachability problem (*in presence of weak data*) with a bounded number of sessions. Verifying whether a protocol is secure is equivalent to deciding whether a particular sequence of protocol messages representing the attack is reachable, that is, the intruder can use the protocol to construct this sequence. Therefore, the reachability problem is simply an ordered set of intruder deduction problems, and is in co-NP [7].

References

[1] Y. Chevalier, R. Kuester, M. Rusinowitch, and M. Turuani. An NP decision procedure for protocol insecurity with xor. 2003. To appear.

[2] H. Comon-Lundh and V. Shmatikov. Constraint solving, exclusive or and the decision of confidentiality for security protocols assuming a bounded number of sessions. June 2003. To appear.

[3] D. Dolev and A. Yao. On the security of public key protocols. In *IEEE Transactions on Information Theory, 29DBLP(2):198-208*, 1983.

[4] L. Gong, T. M. A. Lomas, R. M. Needham, and J. H. Saltzer. Protecting poorly chosen secrets from guessing attacks. In *IEEE Journal on Selected Areas in Communications, Vol.11, No.5, June, 1993, pp.648-656*, 1993.

[5] G. Lowe. Analysing protocols subject to guessing attacks. In *Proceedings of the Workshop on Issues in the Theory of Security (WITS '02)*, 2002.

[6] D. McAllester. Automatic recognition of tractability in inference relations. In *Journal of the ACM, 40(2):284-303*, 1993.

[7] M. Rusinowitch and M. Turani. Protocol insecurity with finite number of sessions is NP-complete. *Proc. 14th IEEE Computer Security Foundations Workshop, pages174-190*, 2001.