# A process algebraic analysis of privacy-type properties in cryptographic protocols

Stéphanie Delaune

LSV, CNRS & ENS Cachan, France

Saturday, September 6th, 2014

# Cryptographic protocols everywhere !



### Cryptographic protocols

- small programs designed to secure communication (*e.g.* secrecy, authentication, anonymity, ...)
- use cryptographic primitives (e.g. encryption, signature, .....)

#### The network is unsecure!

Communications take place over a public network like the Internet.

# Cryptographic protocols everywhere !



### Cryptographic protocols

- small programs designed to secure communication (*e.g.* secrecy, authentication, anonymity, ...)
- use cryptographic primitives (e.g. encryption, signature, .....)

#### It becomes more and more important to protect our privacy.



 $\longrightarrow$  studied in [Arapinis *et al.*, 10]

An electronic passport is a passport with an RFID tag embedded in it.



The RFID tag stores:

- the information printed on your passport,
- a JPEG copy of your picture.

 $\longrightarrow$  studied in [Arapinis *et al.*, 10]

An electronic passport is a passport with an RFID tag embedded in it.



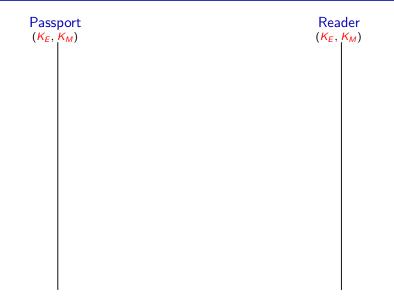
The RFID tag stores:

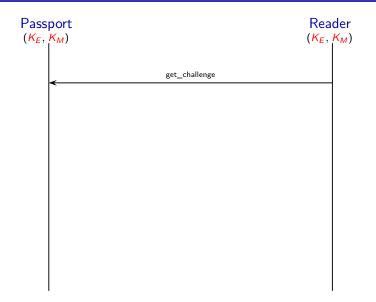
- the information printed on your passport,
- a JPEG copy of your picture.

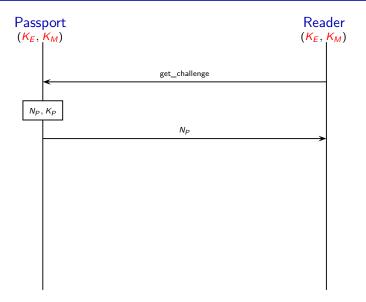
The Basic Access Control (BAC) protocol is a key establishment protocol that has been designed to also ensure unlinkability.

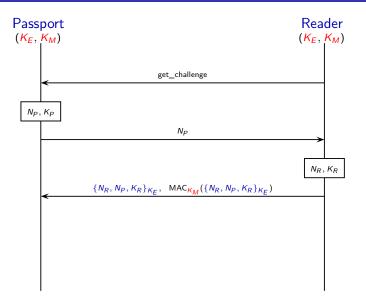
### ISO/IEC standard 15408

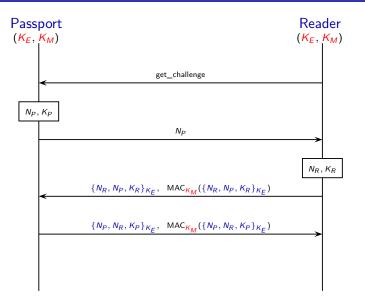
Unlinkability aims to ensure that a user may make multiple uses of a service or resource without others being able to link these uses together.

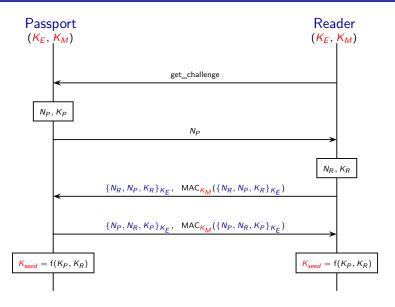












Informally, an observer/attacker can not observe the difference between the two following situations:

- a situation where the same passport may be used twice (or even more);
- a situation where each passport is used at most once.



Informally, an observer/attacker can not observe the difference between the two following situations:

- a situation where the same passport may be used twice (or even more);
- a situation where each passport is used at most once.

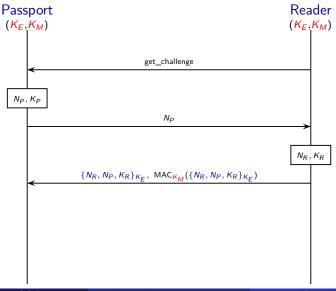


### More formally,

(we still have to formalize the processes and the notion of equivalence)

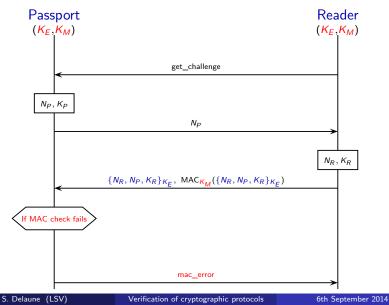
### French electronic passport

 $\rightarrow$  the passport must reply to all received messages.



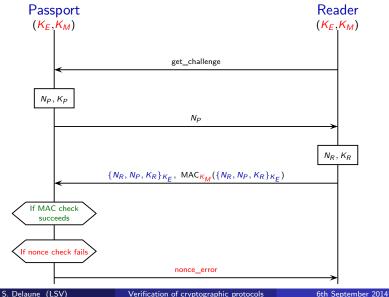
### French electronic passport

 $\rightarrow$  the passport must reply to all received messages.



### French electronic passport

 $\rightarrow$  the passport must reply to all received messages.



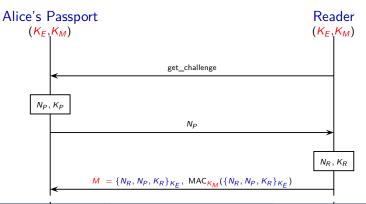
### Attack against unlinkability

An attacker can track a French passport, provided he has once witnessed a successful authentication.

### Attack against unlinkability

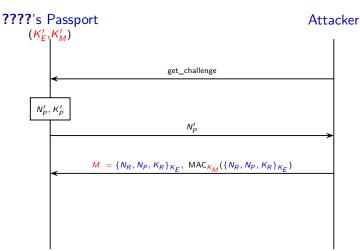
An attacker can track a French passport, provided he has once witnessed a successful authentication.

Part 1 of the attack. The attacker eavesdropes on Alice using her passport and records message M.



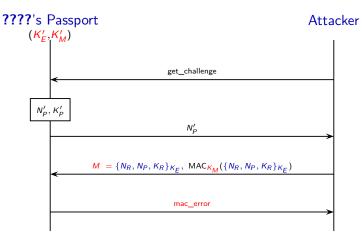
#### Part 2 of the attack.

The attacker replays the message M and checks the error code he receives.



#### Part 2 of the attack.

The attacker replays the message M and checks the error code he receives.

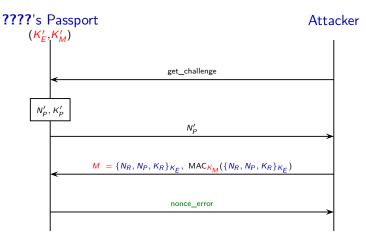


 $\implies$  MAC check failed  $\implies$   $K'_M \neq K_M \implies$  ???? is not Alice

S. Delaune (LSV)

#### Part 2 of the attack.

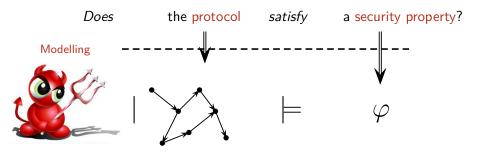
The attacker replays the message M and checks the error code he receives.



 $\implies$  MAC check succeeded  $\implies$   $K'_M = K_M \implies$  ???? is Alice

S. Delaune (LSV)

### Outline



Outline of the remaining of this talk

- Modelling cryptographic protocols and their security properties
- ② Designing verification algorithms
- $\longrightarrow$  we focus here on privacy-type security properties

# Modelling cryptographic protocols and their security properties

### Applied pi calculus

### [Abadi & Fournet, 01]

basic programming language with constructs for concurrency and communication

 $\longrightarrow$  based on the  $\pi$ -calculus [Milner *et al.*, 92] ...

$$P, Q := 0$$
null process  
in(c, x).P  
input  
out(c, u).P  
if  $u = v$  then P else Q conditional  
P | Q  
iP  
new n.P  
new n.P

### Applied pi calculus

### [Abadi & Fournet, 01]

basic programming language with constructs for concurrency and communication

 $\longrightarrow$  based on the  $\pi$ -calculus [Milner *et al.*, 92] ...

... but messages that are exchanged are not necessarily atomic !

### Messages are abstracted by (ground) terms

Ground terms are built over a set of names  $\mathcal{N}$ , and a signature  $\mathcal{F}$ .

$$egin{array}{cccc} {
m t} & ::= & n & {
m name} \; n \ & & & & & & \\ & & & & & & & f(t_1,\ldots,t_k) & {
m application} \; {
m of symbol} \; f \in \mathcal{F} \end{array}$$

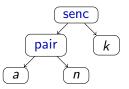
### Messages are abstracted by (ground) terms

Ground terms are built over a set of names  $\mathcal{N}$ , and a signature  $\mathcal{F}$ .

$$egin{array}{cccc} {\mathfrak t} & ::= & n & { ext{name }n} \ & & | & f(t_1,\ldots,t_k) & { ext{application of symbol }f\in \mathcal F} \end{array}$$

### Example: representation of $\{a, n\}_k$

- Names: n, k, a
- constructors: senc, pair,



### Messages are abstracted by (ground) terms

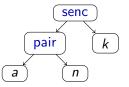
Ground terms are built over a set of names  $\mathcal N,$  and a signature  $\mathcal F.$ 

$$egin{array}{cccc} {\mathfrak t} & ::= & n & { ext{name }n} \ & & \mid & f(t_1,\ldots,t_k) & { ext{application of symbol }f\in \mathcal F} \end{array}$$

 $\longrightarrow$  The term algebra is equipped with an equational theory E.

### Example: representation of $\{a, n\}_k$

- Names: *n*, *k*, *a*
- constructors: senc, pair,
- destructors: sdec,  $proj_1$ ,  $proj_2$ .



 $\longrightarrow$  sdec(senc(x, y), y) = x, proj<sub>1</sub>(pair(x, y)) = x, proj<sub>2</sub>(pair(x, y)) = y.

Cryptographic primitives are modelled using function symbols

- encryption/decryption: senc/2, sdec/2
- concatenation/projections:  $\langle , \rangle/2$ , proj<sub>1</sub>/1, proj<sub>2</sub>/1
- mac construction: mac/2



 $\longrightarrow$  sdec(senc(x, y), y) = x, proj<sub>1</sub>( $\langle x, y \rangle$ ) = x, proj<sub>2</sub>( $\langle x, y \rangle$ ) = y. Nonces  $n_r$ ,  $n_p$ , and keys  $k_r$ ,  $k_p$ ,  $k_e$ ,  $k_m$  are modelled using names Cryptographic primitives are modelled using function symbols

- encryption/decryption: senc/2, sdec/2
- concatenation/projections:  $\langle , \rangle/2$ , proj<sub>1</sub>/1, proj<sub>2</sub>/1
- mac construction: mac/2



 $\longrightarrow$  sdec(senc(x, y), y) = x, proj<sub>1</sub>( $\langle x, y \rangle$ ) = x, proj<sub>2</sub>( $\langle x, y \rangle$ ) = y. Nonces  $n_r$ ,  $n_p$ , and keys  $k_r$ ,  $k_p$ ,  $k_e$ ,  $k_m$  are modelled using names

#### Modelling Passport's role

$$\begin{split} P_{\mathsf{BAC}}(k_E, k_M) &= \mathsf{new} \ n_P.\mathsf{new} \ k_P.\mathsf{in}(\langle z_E, z_M \rangle). \\ & \text{if } z_M = \mathsf{mac}(z_E, k_M) \ \texttt{then if } n_P = \mathsf{proj}_1(\mathsf{proj}_2(\mathsf{sdec}(z_E, k_E))) \\ & \quad \mathsf{then out}(\langle m, \mathsf{mac}(m, k_M) \rangle) \\ & \quad \mathsf{else out}(\mathsf{nonce\_error}) \\ & \quad \mathsf{else out}(\mathsf{mac\_error}) \\ \end{split}$$

#### $\mathsf{Semantics} \to:$

Сомм	$out(c,u).P \mid in(c,x).Q  ightarrow P \mid Q\{u/x\}$
THEN	if $u = v$ then $P$ else $Q  o P$ when $u =_{E} v$
Else	if $u = v$ then P else $Q \rightarrow Q$ when $u \neq_{E} v$

#### Semantics $\rightarrow$ :

Сомм	$out(c,u).P \mid in(c,x).Q  ightarrow P \mid Q\{u/x\}$
THEN	if $u = v$ then $P$ else $Q  o P$ when $u =_{E} v$
Else	if $u = v$ then P else $Q \rightarrow Q$ when $u \neq_{E} v$

#### closed by

• structural equivalence ( $\equiv$ ):

$$P \mid Q \equiv Q \mid P, \quad P \mid 0 \equiv P, \quad \dots$$

• application of evaluation contexts:

$$\frac{P \to P'}{\text{new}\,n.\,P \to \text{new}\,n.\,P'} \qquad \frac{P \to P'}{P \mid Q \to P' \mid Q}$$

Privacy-type properties are modelled as equivalence-based properties

testing equivalence between P and Q,  $P \approx_t Q$ 

for all processes A, we have that:

 $(A \mid P) \Downarrow_c$  if, and only if,  $(A \mid Q) \Downarrow_c$ 

where  $P \Downarrow_c$  means that P can evolve and emits on public channel c.

Privacy-type properties are modelled as equivalence-based properties

testing equivalence between P and Q,  $P \approx_t Q$ 

for all processes A, we have that:

 $(A \mid P) \Downarrow_c$  if, and only if,  $(A \mid Q) \Downarrow_c$ 

where  $P \Downarrow_c$  means that P can evolve and emits on public channel c.

Example 1:  $\operatorname{out}(a, s) \approx^?_t \operatorname{out}(a, s')$ 

Privacy-type properties are modelled as equivalence-based properties

testing equivalence between P and Q,  $P \approx_t Q$ 

for all processes A, we have that:

 $(A \mid P) \Downarrow_c$  if, and only if,  $(A \mid Q) \Downarrow_c$ 

where  $P \Downarrow_c$  means that P can evolve and emits on public channel c.

Example 1:

$$\operatorname{out}(a, \mathbf{s}) \not\approx_t \operatorname{out}(a, \mathbf{s}')$$

 $\longrightarrow$  A = in(a, x).if x = s then out(c, ok)

Privacy-type properties are modelled as equivalence-based properties

testing equivalence between P and Q,  $P \approx_t Q$ 

for all processes A, we have that:

$$(A \mid P) \Downarrow_c$$
 if, and only if,  $(A \mid Q) \Downarrow_c$ 

where  $P \Downarrow_c$  means that P can evolve and emits on public channel c.

Example 2:

new s.out(a, senc(s, k)).out(a, senc(s, k'))  
$$\stackrel{?}{\approx_t}$$
  
new s, s'.out(a, senc(s, k)).out(a, senc(s', k'))

## Security properties - privacy

Privacy-type properties are modelled as equivalence-based properties

testing equivalence between P and Q,  $P \approx_t Q$ 

for all processes A, we have that:

$$(A \mid P) \Downarrow_c$$
 if, and only if,  $(A \mid Q) \Downarrow_c$ 

where  $P \Downarrow_c$  means that P can evolve and emits on public channel c.

Example 2:

new s.out(a, senc(s, k)).out(a, senc(s, k'))  

$$\approx_t$$
  
new s, s'.out(a, senc(s, k)).out(a, senc(s', k'))

 $\longrightarrow A = in(a, x).in(a, y).if (sdec(x, k) = sdec(y, k')) then out(c, ok)$ 

### Security properties - privacy

Privacy-type properties are modelled as equivalence-based properties

testing equivalence between P and Q,  $P \approx_t Q$ 

for all processes A, we have that:

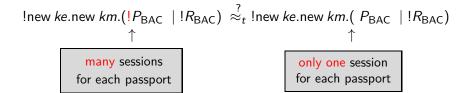
 $(A \mid P) \Downarrow_c$  if, and only if,  $(A \mid Q) \Downarrow_c$ 

where  $P \Downarrow_c$  means that P can evolve and emits on public channel c.

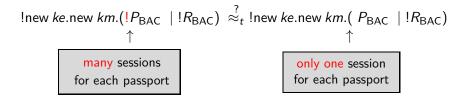
Question: Are the two following processes in testing equivalence?

$$\mathsf{new}\,s.\mathsf{out}(a,s) \stackrel{?}{pprox}{}_t \;\; \mathsf{new}\,s.\mathsf{new}\,k.\mathsf{out}(a,\mathsf{senc}(s,k))$$

#### [Arapinis et al, 2010]



#### [Arapinis et al, 2010]

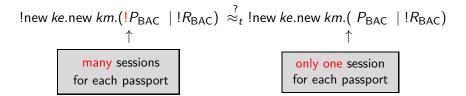


Vote privacy

[Kremer and Ryan, 2005]

 $V_A(yes) \approx_t V_A(no)$ 

[Arapinis et al, 2010]



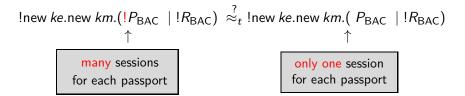
Vote privacy

[Kremer and Ryan, 2005]

$$V_A(yes) \mid V_B(no) \approx_t V_A(no) \mid V_B(yes)$$

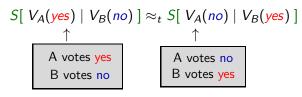
$$\uparrow \qquad \uparrow$$
A votes yes
B votes no
B votes yes

[Arapinis et al, 2010]



Vote privacy

[Kremer and Ryan, 2005]



ightarrow often requires some assumptions  $S[\_]$ 

# Designing verification algorithms for privacy-type properties

# Difficulties when checking testing equivalence $P \approx_t Q$

testing equivalence is undecidable in general

# Difficulties when checking testing equivalence $P \approx_t Q$

#### testing equivalence is undecidable in general

#### Processes without replication

• We still have to consider any possible behavior for the attacker (for all quantification over processes).

 $\longrightarrow$  no hope to test each possible behavior of the attacker in turn

• Once the behavior of the attacker is fixed, we still have to decide whether the two sequences of messages that are outputted are indistinguishable or not.

 $\longrightarrow$  the so-called static equivalence problem.

### Static equivalence $\sigma \sim \sigma'$ (modulo E)

Two sequences of messages  $\sigma = \{w_1 \rightarrow u_1, \dots, w_n \rightarrow u_n\}$  and  $\sigma' = \{w_1 \rightarrow u'_1, \dots, w_n \rightarrow u'_n\}$  are in static equivalence when:

 $R_1\sigma =_E R_2\sigma \iff R_1\sigma' =_E R_2\sigma'$  for any recipes  $R_1, R_2$ 

### Static equivalence $\sigma \sim \sigma'$ (modulo E)

Two sequences of messages  $\sigma = \{w_1 \rightarrow u_1, \dots, w_n \rightarrow u_n\}$  and  $\sigma' = \{w_1 \rightarrow u'_1, \dots, w_n \rightarrow u'_n\}$  are in static equivalence when:

 $R_1\sigma =_E R_2\sigma \iff R_1\sigma' =_E R_2\sigma'$  for any recipes  $R_1, R_2$ 

Example: a simple voting protocol with 3 voters *a*, *b*, and *c*.

 $V \rightarrow S : \langle V, \{ vote \}_k \rangle$ 

### Static equivalence $\sigma \sim \sigma'$ (modulo E)

Two sequences of messages  $\sigma = \{w_1 \rightarrow u_1, \dots, w_n \rightarrow u_n\}$  and  $\sigma' = \{w_1 \rightarrow u'_1, \dots, w_n \rightarrow u'_n\}$  are in static equivalence when:

 $R_1\sigma =_E R_2\sigma \iff R_1\sigma' =_E R_2\sigma'$  for any recipes  $R_1, R_2$ 

Example: a simple voting protocol with 3 voters *a*, *b*, and *c*.

 $V \rightarrow S : \langle V, \{ vote \}_k \rangle$ 

•  $\sigma = \{w_1 \rightarrow \langle a, \{yes\}_k \rangle; w_2 \rightarrow \langle b, \{no\}_k \rangle; w_3 \rightarrow \langle c, \{yes\}_k \rangle\}, \text{ and}$ •  $\sigma' = \{w_1 \rightarrow \langle a, \{no\}_k \rangle; w_2 \rightarrow \langle b, \{yes\}_k \rangle; w_3 \rightarrow \langle c, \{yes\}_k \rangle\}.$ 

where a, b, c, yes, and no are public constants.

### Static equivalence $\sigma \sim \sigma'$ (modulo E)

Two sequences of messages  $\sigma = \{w_1 \rightarrow u_1, \dots, w_n \rightarrow u_n\}$  and  $\sigma' = \{w_1 \rightarrow u'_1, \dots, w_n \rightarrow u'_n\}$  are in static equivalence when:

$$R_1\sigma =_E R_2\sigma \iff R_1\sigma' =_E R_2\sigma'$$
 for any recipes  $R_1, R_2$ 

Example: a simple voting protocol with 3 voters a, b, and c.

 $V \rightarrow S : \langle V, \{ vote \}_k \rangle$ 

•  $\sigma = \{w_1 \rightarrow \langle a, \{yes\}_k \rangle; w_2 \rightarrow \langle b, \{no\}_k \rangle; w_3 \rightarrow \langle c, \{yes\}_k \rangle\}, \text{ and}$ •  $\sigma' = \{w_1 \rightarrow \langle a, \{no\}_k \rangle; w_2 \rightarrow \langle b, \{yes\}_k \rangle; w_3 \rightarrow \langle c, \{yes\}_k \rangle\}.$ 

where a, b, c, yes, and no are public constants.

 $\longrightarrow \sigma$  and  $\sigma'$  are not in static equivalence  $R_1 = \operatorname{proj}_2(w_1)$  and  $R_2 = \operatorname{proj}_2(w_3)$ 

### Static equivalence $\sigma \sim \sigma' \pmod{\mathsf{E}}$

Two sequences of messages  $\sigma = \{w_1 \rightarrow u_1, \dots, w_n \rightarrow u_n\}$  and  $\sigma' = \{w_1 \rightarrow u'_1, \dots, w_n \rightarrow u'_n\}$  are in static equivalence when:

 $R_1\sigma =_E R_2\sigma \iff R_1\sigma' =_E R_2\sigma'$  for any recipes  $R_1, R_2$ 

The static equvialence problem is decidable (even in PTIME) for many interesting equational theories useful to model cryptography primitives. [Abadi & Cortier, TCS 2006], [Cortier & Delaune, JAR 2012]

- $\longrightarrow$  Some automatic tools are available, e.g.
  - YAPA: http://www.lsv.ens-cachan.fr/~baudet/yapa/
  - KISS: http://www.lsv.ens-cachan.fr/~ciobaca/kiss/
  - FAST: http://www.infsec.ethz.ch/people/brunoco

# Testing equivalence (for processes with replication)

Some decidability results [Chrétien, Cortier & D., ICALP'13 & CONCUR'14]

- - restricted set of cryptographic primitives
- - some syntaxic restrictions on the shape of the processes

A more pragmatic approach	[Blanchet <i>et al.</i> , 2005]
ProVerif	http://www.proverif.ens.fr
• + various cryptographic primitives	
<ul> <li>termination is not guaranteed; diff-equivalence (too strong)</li> </ul>	

 $\longrightarrow$  These results are not suitable to analyse vote-privacy, or unlinkability of the BAC protocol.

### Cheval, Comon-Lundh & D.

CCS 2011

A procedure for deciding testing equivalence for a large class of processes implemented in a tool called APTE

#### Our class of processes:

- + non-trivial else branches, private channels, and non-deterministic choice;
- but no replication, and a fixed set of cryptographic primitives (signature, symmetric and asymmetric encryptions, hash function, mac, pairs).

### Cheval, Comon-Lundh & D.

A procedure for deciding testing equivalence for a large class of processes implemented in a tool called APTE

#### Our class of processes:

- + non-trivial else branches, private channels, and non-deterministic choice;
- but no replication, and a fixed set of cryptographic primitives (signature, symmetric and asymmetric encryptions, hash function, mac, pairs).

Similar results for restricted class of processes have been obtained in [Baudet, 05], [Dawson & Tiu, 10], [Chevalier & Rusinowitch, 10], [Chadha *et al.*, 12], ...

CCS 2011

#### Two main steps:

A symbolic exploration of all the possible traces
 The infinite number of possible traces (*i.e.* experiment) are represented by a finite set of symbolic traces.

 $\rightarrow$  this set is still huge (exponential) !

A decision procedure for deciding (symbolic) equivalence between sets of symbolic traces.

 $\longrightarrow$  this algorithm works quite well

#### Two main steps:

A symbolic exploration of all the possible traces
 The infinite number of possible traces (*i.e.* experiment) are represented by a finite set of symbolic traces.

 $\rightarrow$  this set is still huge (exponential) !

A decision procedure for deciding (symbolic) equivalence between sets of symbolic traces.

 $\longrightarrow$  this algorithm works quite well

#### Some applications

- unlinkability in RFID protocols (*e.g.* e-passport protocol)
- anonymity (e.g. private authentication protocol)

#### Two main steps:

A symbolic exploration of all the possible traces
 The infinite number of possible traces (*i.e.* experiment) are represented by a finite set of symbolic traces.

 $\rightarrow$  this set is still huge (exponential) !

A decision procedure for deciding (symbolic) equivalence between sets of symbolic traces.

 $\longrightarrow$  this algorithm works quite well

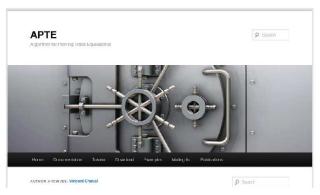
#### Main limitations

- e-voting protocols are still out of reach
- we can only handle very few sessions (state space explosion problem)

### APTE- Algorithm for Proving Trace Equivalence

#### http://projects.lsv.ens-cachan.fr/APTE

 $\longrightarrow$  developed by Vincent CHEVAL



#### $\longrightarrow$ written in Ocaml, around 12 KLocs

### It remains a lot to do for analysing privacy-type properties

- formal definitions of some sublte security properties (receipt-freeness, coercion-resistance, ...)
- algorithms (and tools!) for checking (automatically or not) testing equivalence for various cryptographic primitives;
- more composition results.



```
Main topics of the ANR JCJC - VIP project
(Jan. 2012 - Dec 2015)
http://www.lsv.ens-cachan.fr/Projects/anr-vip/
```

 $\longrightarrow$  a postdoc position is available on this project.