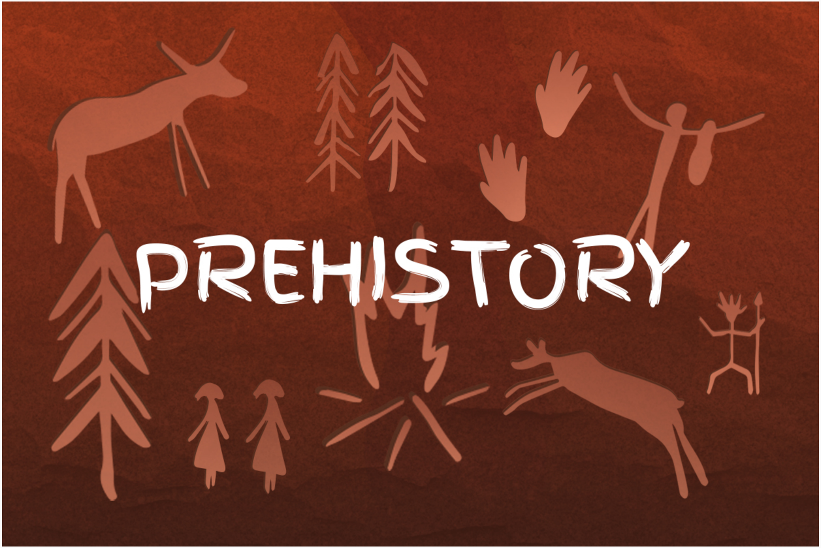
Formal verification of security protocols
The History

Véronique CORTIER & Stéphanie DELAUNE

June 16, 2022

Prehistory - past millennium



Prehistory - past millennium

I enjoyed working on the ~~the peasant knights of the year 1000 at lake paladru~~ simultaneous rigid reachability problem but ... you know, at dinner time...

Sure, let's go for security protocols!

Early Dolev-Yao model - FOCS 1981

On the Security of Public Key Protocols

DANNY DOLEV AND ANDREW C. YAO, MEMBER, IEEE

Abstract—Recently the use of public key encryption to provide secure network communication has received considerable attention. Such public key systems are usually effective against passive ... tap the lines and try to decipher th... however, that an improperly des... active saboteur, one who may ... message being transmitted. Several models ... security of protocols can be discussed precisely. Algorithms and characteri- zations that can be used to determine protocol security in these models are given.

I. Introduction

THE USE of public key encryption [1], [11] to provide secure network communication has received considerable attention [2], [7], [8], [10]. Such public key systems are usually very effective against a "passive" eavesdropper, namely, one who merely taps the communication line and

issues can be discussed precisely. The models we introduce ... to study the security problem for families of ... on the behavior of ...

$$E_x D_x = D_x E_x = 1$$

... key encryption ...). In a public key system, every user X has an *encryption function* E_x and a *decryption function* D_x, both are mappings from $\{0,1\}^*$ (the set of all finite binary sequences) into $\{0,1\}^*$. A secure public directory contains all the (X, E_x) pairs, while the decryption function D_x is known only to user X. The main requirements on E_x, D_x are:

1) $E_x D_x = D_x E_x = 1$, and
2) knowing $E_x(M)$ and the public directory does not reveal anything about the value M.

Early Dolev-Yao model - FOCS 1981

On the Security of Public Key Protocols

DANNY DOLEV AND ANDREW C. YAO, MEMBER, IEEE

Abstract—**Recently the use of public key encryption to provide secure network communication has received considerable attention. Such public key systems are usually effective against passive ... tap the lines and try to decipher th... however, that an improperly des... active saboteur, one who may ... message being transmitted. Several models ... security of protocols can be discussed precisely. Algorithms and characteri-zations that can be used to determine protocol security in these models are given.**

I. INTRODUCTION

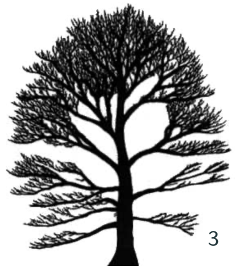
THE USE of public key encryption [1], [11] to provide secure network communication has received considerable attention [2], [7], [8], [10]. Such public key systems are usually very effective against a "passive" eavesdropper, namely, one who merely taps the communication line and

issues can be discussed precisely. The models we introduce ... the security problem for families of ... on the behavior of ... key encryption ... In a public key system, every user X has an *encryption function* E_x and a *decryption function* D_x, both are mappings from $\{0, 1\}^*$ (the set of all finite binary sequences) into $\{0, 1\}^*$. A secure public directory contains all the (X, E_x) pairs, while the decryption function D_x is known only to user X. The main requirements on E_x, D_x are:

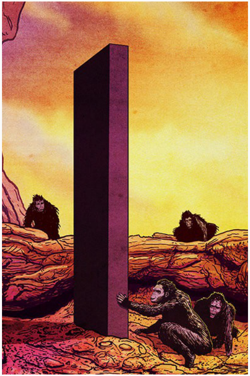
1) $E_x D_x = D_x E_x = 1$, and
2) knowing $E_x(M)$ and the pu... reveal anything about the val...

$$E_x D_x = D_x E_x = 1$$

Yeah... They are working with words.
Let's go for trees!



The Graal



The Needham-Schroeder protocol

$$\begin{aligned} A &\rightarrow B: \quad \{A, N_A\}_{\mathrm{pub}(B)} \\ B &\rightarrow A: \quad \{N_A, N_B\}_{\mathrm{pub}(A)} \\ A &\rightarrow B: \quad \{N_B\}_{\mathrm{pub}(B)} \end{aligned}$$

and its *man-in-the-middle* attack

Victory! 

Decidable class for security protocols [Icalp'01,RTA'03]

- one variable per rule
- no nonces
- at least doubly exponential...

Victory! 

Decidable class for security protocols [Icalp'01,RTA'03]

- one variable per rule
- no nonces
- at least doubly exponential...

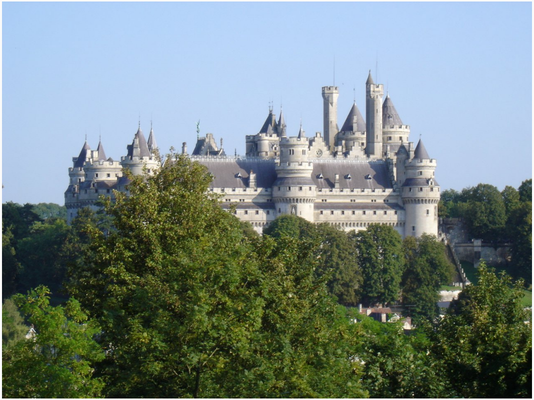
Decidability for bounded sessions Rusinowitch, Turuani [CSFW'01]

- CL-ATSE tool
- works for a small number of sessions (2-3)

Birth of ProVerif Blanchet [CSFW'01]

- Forget about decidability ;-)
- Needham-Schroeder with nonces

Middle Ages (2000-2010)



Pierrefonds castle (Barbizon Édition 2011)

Luc à Hubert, qui se bat avec les rideaux électriques :

Tu vas y arriver, c'est un automate fini. [Best-Of, 2011]

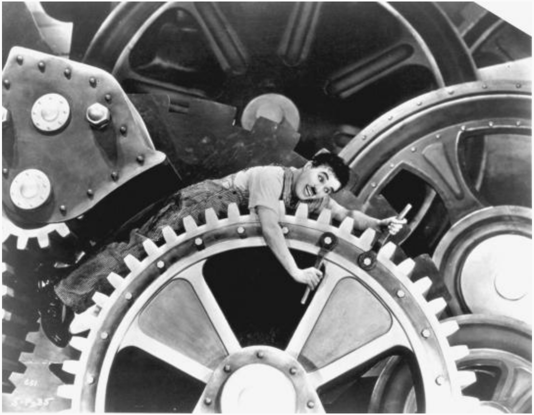
More primitives

⟶ The aim was to take into account the algebraic properties of cryptographic primitives to model them in a more faithful way.

- [V. Bernat, 2006]: Théories de l'intrus pour la vérification des protocoles cryptographiques;
- [S. Delaune, 2006]: Vérification des protocoles cryptographiques et propriétés algébriques;
- [Bursuc, 2009]: Contraintes de déductibilité dans une algèbre quotient : réduction de modèles et applications à la sécurité.

ETAPS 2004, Barcelona

Modern Times (2010-2020)

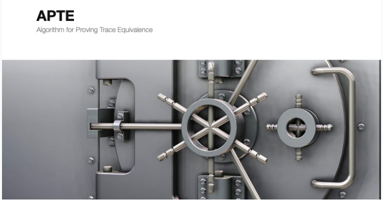


Google me dit que la solution à ce problème est dans un papier que j'ai écrit... [Hubert, 2018]

More properties and more automation

$\longrightarrow$ the importance of equivalence properties to model e.g. anonymity, unlinkability

[V. Cheval, 2012]: Automatic verification of cryptographic protocols : privacy-type properties



More properties and more automation

⟶ the importance of equivalence properties to model e.g. anonymity, unlinkability

[V. Cheval, 2012]: Automatic verification of cryptographic protocols : privacy-type properties

DEEPSEC

DEciding Equivalence Properties in SECurity protocols

https://deepsec-prover.github.io

Success story ! A formal analysis of unlinkability of the BAC protocol

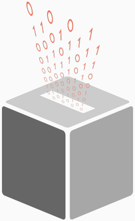
A novel approach - CCSA approach

⟶ to obtain security guarantees in the computational setting.

- [G. Scerri, 2015]: Proof of security protocols revisited;
- [A. Koutsos, 2019]: Preuves symboliques de propriétés d'indistinguabilité calculatoire;
- [C. Jacomme, 2020]: Preuves de protocoles cryptographiques : méthodes symboliques et attaquants puissants.

The framework is now implemented in the Squirrel prover – `https://squirrel-prover.github.io`

Nowdays ...




Google apps
talk


https://


Crédit Mutuel
Crédit Mutuel
MasterCard


#Tous
AntiCovid


Union européenne
République française
PASSEPORT

Contactless systems ...

... so near and yet so far !

Contactless payment

A few figures regarding 2020 (France):

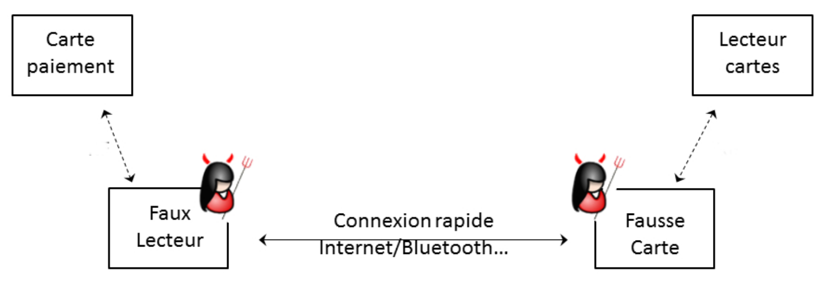
- 4.6 billion of transactions were paid contactless (40%);
- 6 out of 10 transactions of less than 50 euros.

Contactless payment

A few figures regarding 2020 (France):

- 4.6 billion of transactions were paid contactless (40%);
- 6 out of 10 transactions of less than 50 euros.

Contactless payment is vulnerable to relay attack:



⟶ **How to prevent such an attack?**

Distance bounding protocols

They aim to ensure authentication and physical proximity.

- more than 40 protocols have been designed since 1993;
- included in the EMV specification (payment) since 2016.

[Avoine *et al.*, ACM Computing Surveys, 2019]

Distance bounding protocols

They aim to ensure authentication and physical proximity.

- more than 40 protocols have been designed since 1993;
- included in the EMV specification (payment) since 2016.

[Avoine *et al.*, ACM Computing Surveys, 2019]

How it works (or not) !

$$P \to V : \; commit(m, k)$$

$$\boxed{\begin{array}{l} V \to P : \; chall \\ P \to V : \; chall \oplus m \end{array}}$$

$$P \to V : \; k, \; Sign_P(\langle m, chall \oplus m \rangle)$$

[Brands and Chaum, 93]

Distance bounding protocols

They aim to ensure authentication and physical proximity.

- more than 40 protocols have been designed since 1993;
- included in the EMV specification (payment) since 2016.

[Avoine *et al.*, ACM Computing Surveys, 2019]

How it works (or not) !

$$P \rightarrow V : \quad commit(m, k)$$

$$V \rightarrow P : \quad chall$$

$$P \rightarrow V : \quad chall \oplus m$$

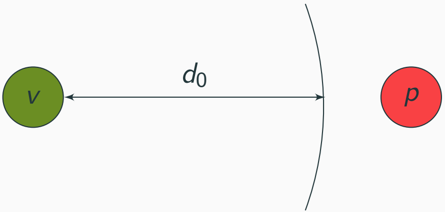
$$2 \times dist(V, P) \leq \Delta t \times c$$

$$P \rightarrow V : \; k, \; Sign_P(\langle m, chall \oplus m \rangle)$$

[Brands and Chaum, 93]

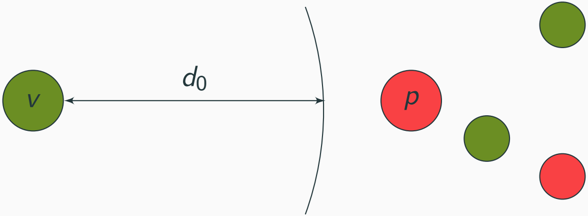
Distance fraud

A malicious prover should not be able to successfully complete a session with an honest verifier who is far away



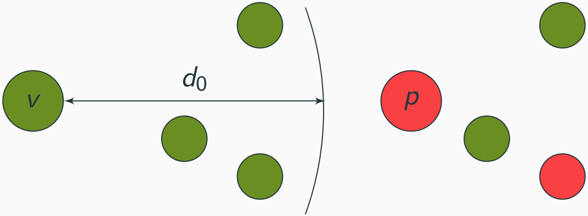
Distance fraud

A malicious prover should not be able to successfully complete a session with an honest verifier who is far away



Distance fraud (including distance hijacking)

A malicious prover should not be able to successfully complete a session with an honest verifier who is far away (even with the help of some honest agents in the neighbourhood)

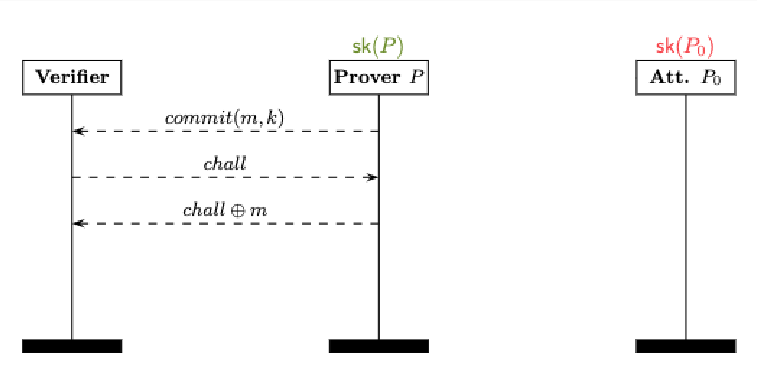


$\longrightarrow$ Distance hijacking attack has been overlooked until 2012.

[Cremers *et al.*, S&P'12]

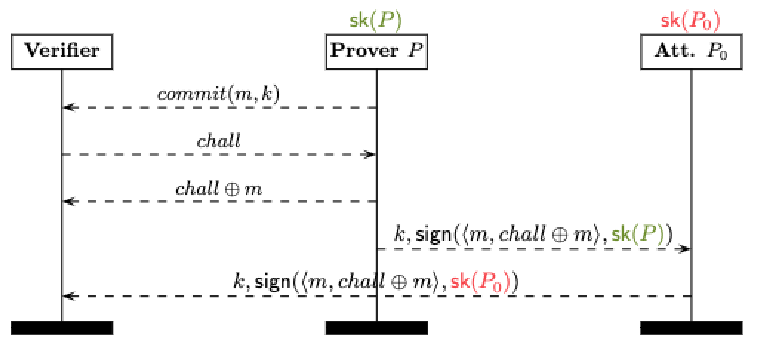
Distance Hijacking attack

P is in the neighboorhood of V whereas P_0 (dishonest) is far away.



Distance Hijacking attack

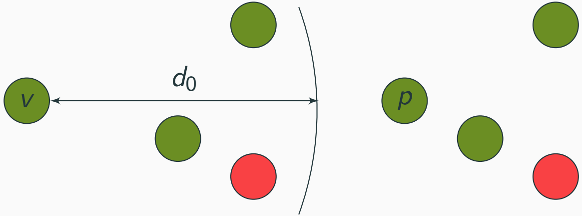
P is in the neighboorhood of V whereas P_0 (dishonest) is far away.



$\longrightarrow$ At the end, V ends the protocol successfully with P_0 whereas P_0 is far away.

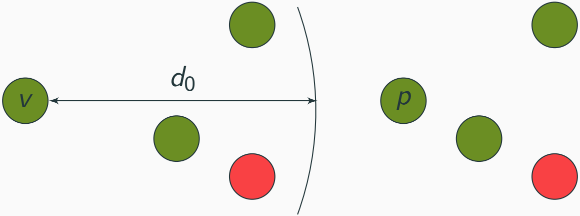
Mafia fraud

An attacker should not be able to abuse a far away honest prover to pass the protocol.



Mafia fraud

An attacker should not be able to abuse a far away honest prover to pass the protocol.



$\longrightarrow$ A payment protocol should resist to mafia fraud

Symbolic security analysis

We need a framework that takes into account:

- transmission delay, location of participants, mobility issues, ...
- low-level operators and their algebraic properties, such as exclusive-or.

Some existing works:

- Formalisation in Isabelle/HOL [Basin *et al.*, CSF'09]
- Distance-hijacking attack [Cremers *et al.*, S&P'12]

$\longrightarrow$ lack of automation to support the security analysis.

Some contributions

A lot of progress has been done!

1. A framework to model distance, location, transmission delay, mobility, . . . ;
2. Formal symbolic definitions of the different types of fraud;
3. Reduction results to allow the use of the ProVerif tool.

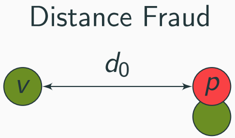
[PhD thesis of A. Debant, 2020]

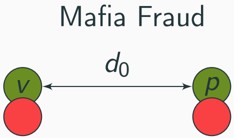
A Tamarin-based framework has been developed concurrently by S. Mauw *et al.*

[PhD thesis of J. Toro-Pozo, 2019]

Reduction results

When analysing distance fraud and mafia fraud, we can restrict ourselves to the anlaysis of the following configurations:



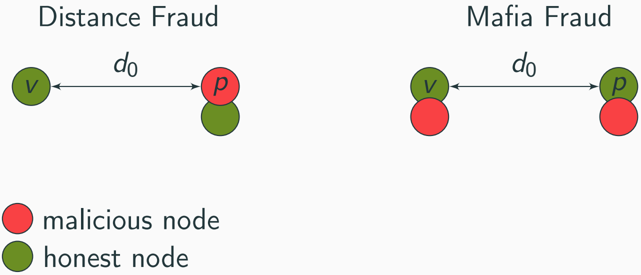


malicious node

honest node

Reduction results

When analysing distance fraud and mafia fraud, we can restrict ourselves to the anlaysis of the following configurations:



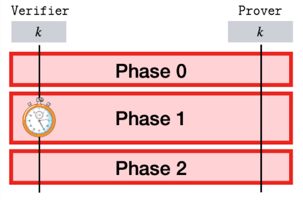
$\longrightarrow$ These topologies are enough even when considering mobility – as soon as agents do not move too fast.

[Boureanu, Chothia, Debant & Delaune, CCS'20]

Analysis using ProVerif

We can encode this fixed topology relying on the phase mechanism of ProVerif.

- phase 0: slow initialization phase
- phase 1: rapid phase
- phase 2: slow verification phase



⟶ Remote agents do *not* act in phase 1 !

Efficient analysis (few minutes or even less) for most of the protocols using the latest version of ProVerif.

Some case studies

Protocols	MF	DH	TF
Basin's toy example [7]	✓	✓	✓
Brands and Chaum [15]			
• Signature	✓	✗	*o.o.s.*
• Fiat-Shamir	✓	✗	✗
CRCS No-revealing sign [52]			
• No-revealing sign	✓	✓	✗
• Revealing sign	✓	✗	✗
Eff-PKDB [40]			
• No protection *(new)*	✓	✓	✓
• Protected *(new)*	✓	✓	✓
Hancke and Kuhn	✓	✓	✗
MAD (One-Way) [17]	✓	✗	*o.o.s.*

...

Going back to contactless payment

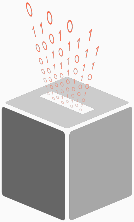
Protocols	MF	DH	TF
MasterCard RRP [33]	✓	✗	✗
NXP [39]	✓	✗	✗
PaySafe [20]	✓	✗	✗

Going back to contactless payment

Protocols	MF	DH	TF
MasterCard RRP [33]	✓	✗	✗
NXP [39]	✓	✗	✗
PaySafe [20]	✓	✗	✗

None of these protocols are implemented in our credit cards !

Electronic voting



- Possibly more convenient
  ⟶ for voters: vote from home, or abroad
  ⟶ for authorities: easier to record and tally votes
- May allow for more "democracy"
  ⟶ complex tally process (Condorcet, STV, IRV)
  ⟶ can be used more often
  ⟶ complex legal rules (a voter may vote from any place in her state)

Confidentiality of the votes

Vote privacy

"No one should know how I voted"

Confidentiality of the votes

Vote privacy

"No one should know how I voted"

Better: Receipt-free / Coercion-resistant

"No one should know how I voted,
even if I am willing to tell my vote! "

Confidentiality of the votes

Vote privacy

"No one should know how I voted"

Better: Receipt-free / Coercion-resistant

"No one should know how I voted,
even if I am willing to tell my vote! "

- vote buying
- coercion

Confidentiality of the votes

Vote privacy

"No one should know how I voted"

Better: Receipt-free / Coercion-resistant

"No one should know how I voted,

even if I am willing to tell my vote! "

- vote buying
- coercion

Everlasting privacy: no one should know my vote, even when the cryptographic keys will be eventually broken.

Verifiability

Individual Verifiability: a voter can check that

- *cast as intended:* their ballot contains their intended vote
- *recorded as cast*: their ballot is in the ballot box.

Universal Verifiability: everyone can check that

- *tallied as recorded*: the result corresponds to the ballot box.
- *eligibility*: ballots have been casted by legitimate voters.

You should verify the election,
not the system.

Verifiability

Individual Verifiability: a voter can check that

- *cast as intended:* their ballot contains their intended vote
- *recorded as cast*: their ballot is in the ballot box.

Universal Verifiability: everyone can check that

- *tallied as recorded*: the result corresponds to the ballot box.
- *eligibility*: ballots have been casted by legitimate voters.

You should verify the election,
not the system.

Even better: accountability

- the system tells whom to blame
- eases dispute resolution

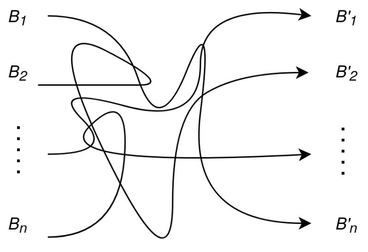
How to tally an election?

Homomorphic tally

$$\mathrm{dec}(B_1 B_2) = \mathrm{dec}(B_1) + \mathrm{dec}(B_2)$$

Limited form of voting

Mixing



Leaks too much information

Italian attacks

Some voting systems (Condorcet, STV) let you chose any permutation of the candidates.

1. Alice
2. Bob
3. Charlie
4. Diane
5. Eve

⋮

1. Alice
2. Eve
3. Diane
4. Greg
5. Bob

⋮

1. Alice
2. Kolmogorov
3. Zarathustra
4. Leto
5. Puccini

⋮

1. Eve
2. Bob
3. Freya
4. Diane
5. Alice

⋮

1. Eve
2. Charlie
3. Alice
4. Bob
5. Diane

⋮

Some voting systems (Condorcet, STV) let you chose any permutation of the candidates.

1. Alice
2. Bob
3. Charlie
4. Diane
5. Eve

⋮

1. Alice
2. Eve
3. Diane
4. Greg
5. Bob

⋮

1. Alice
2. Kolmogorov
3. Zarathustra
4. Leto
5. Puccini

⋮

Italian Attack!
(vote buying)

1. Eve
2. Bob
3. Freya
4. Diane
5. Alice

⋮

1. Eve
2. Charlie
3. Alice
4. Bob
5. Diane

⋮

ElGamal vs Paillier

Joint work with Quentin Yang and Pierrick Gaudry

Our finding: ElGamal suitable even for complex tally procedures

ElGamal offers

- a standard security assumption
- much better tool support
- simple efficient threshold decryption

How to use ElGamal for MPC

- addition is easy: $\text{enc}(x) * \text{enc}(y) = \text{enc}(x + y)$
- ... but not multiplication (unlike Paillier)
- let's encrypt bit-wise
- one key primitive:

$$\text{CGATE}(\text{enc}(x), \text{enc}(b)) = \begin{cases} \text{enc}(x) & \text{if } b = 1 \\ \text{enc}(0) & \text{if } b = 0 \end{cases}$$

- all the rest is built upon CGATE, with some optimizations

How to use ElGamal for MPC

- addition is easy: $\mathrm{enc}(x) * \mathrm{enc}(y) = \mathrm{enc}(x+y)$
- ... but not multiplication (unlike Paillier)
- let's encrypt bit-wise
- one key primitive:

$$\mathrm{CGATE}(\mathrm{enc}(x), \mathrm{enc}(b)) = \begin{cases} \mathrm{enc}(x) & \text{if } b = 1 \\ \mathrm{enc}(0) & \text{if } b = 0 \end{cases}$$

- all the rest is built upon CGATE, with some optimizations

Theorem

CGATE *UC-implements the corresponding ideal trusted party.*

An MPC toolbox for tally hiding

Logical operations:

- $CG(\text{enc}(b_1), \text{enc}(b_2)) = \text{enc}(b_1 \wedge b_2)$.

Arithmetic operations:

- $Add(\text{enc}(x), \text{enc}(y)) = \text{enc}(x + y)$,
- $Sub(\text{enc}(x), \text{enc}(y)) = \text{enc}(x - y)$,
- $Gt(\text{enc}(x), \text{enc}(y)) = \text{enc}(x < y)$,
- But also multiplication and division.

More complex algorithms:

- Sort
- Find the s largest values

Another counting function? We can do it!

Tally hiding in the litterature

Ordinos: Voters select one candidate, the candidate(s) with the most votes are elected. Based on Paillier.

[Kuesters, Liedtke, Mueller, Rausch, Vogt (2020)]

(Very specific counting function.)

Tally hiding in the litterature

Ordinos: Voters select one candidate, the candidate(s) with the most votes are elected. Based on Paillier.

[Kuesters, Liedtke, Mueller, Rausch, Vogt (2020)]

(Very specific counting function.)

Condorcet based on homomorphic ElGamal

[Haines, Pattinson, Tiwari (2019)]

→ Privacy breach when two candidates are ranked equal.

Tally hiding in the litterature

Ordinos: Voters select one candidate, the candidate(s) with the most votes are elected. Based on Paillier.

[Kuesters, Liedtke, Mueller, Rausch, Vogt (2020)]

(Very specific counting function.)

Condorcet based on homomorphic ElGamal

[Haines, Pattinson, Tiwari (2019)]

→ Privacy breach when two candidates are ranked equal.

Majority Judgment: Voters grade each candidate, the one with the best median sequence is elected.

based on Paillier [Canard, Pointcheval, Santos, Traoré (2018)]

→ fails in not so rare cases (22% fail rate for 100 voters)

The case of STV

STV Algorithm

- emulates a multi-round election
- a fraction of votes is tranfered to remaining candidates

The case of STV

STV Algorithm

- emulates a multi-round election
- a fraction of votes is tranfered to remaining candidates

The ideal algorithm is not practical!

New South Wales (Australia): 21 seats, 346 candidates, 3,5 million votes

→ Requires 30GB of central memory to store the fractions, whose size roughly doubles at each selection of candidate

→ Let's go for rounding, in MPC

The case of STV

STV Algorithm

- emulates a multi-round election
- a fraction of votes is tranfered to remaining candidates

The ideal algorithm is not practical!

New South Wales (Australia): 21 seats, 346 candidates, 3,5 million votes

→ Requires 30GB of central memory to store the fractions, whose size roughly doubles at each selection of candidate

→ Let's go for rounding, in MPC

Literature

- [Wen, Buckland (2008)] no rounding
- [Benaloh, Moran, Naish, Ramchen, Teague (2010)]

→ both leak some information (way less than the ballots)

Precise complexity analysis

Example: Different trade-offs for the Condorcet counting function

Version	Leakage	Voters # exp.	Authorities # exp.	Authorities # comm.	Size of the transcript
[19]	adj. matrix **privacy breach** [i]	$10k^2$	$18ank^2$	2	$10ank^2$
ballots as list of integers (partial MPC)	adj. matrix	$8k \log k$	$30nak^2 \log k$	$2 \log k$	$27nak^2 \log k$
ballots as list of integers (full MPC)	$\varnothing$	$8k \log k$	$10nak^2(3 \log k + 5m)$ $+120mak^3$	$m(m+4k)$	$9nak^2(3 \log k + 5m)$ $+108mak^3$
ballots as matrices	adj. matrix	$\frac{51}{2}k^2$	$\frac{51}{2}nk^2$	0	$\frac{29}{2}nk^2$
ballots as matrices (naive, for comparison)	adj. matrix	$20k^3$	$20nk^3$	0	$20nk^3$

[i] [19] leaks, for each ballot, the number of candidates ranked at equality. In particular, who voted blank is known to everyone.

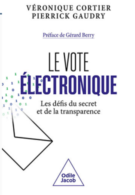
Figure 5. Leading terms of the cost of MPC implementations of Condorcet winners. n: number of voters, $m = \lceil \log(n+1) \rceil$, k: number of candidates, a: number of authorities.

With Quentin Yang and Pierrick Gaudry, eprint 2021/491

Bonus information

A popularization book

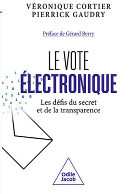
- Odile Jacob
- with Pierrick Gaudry



Bonus information

A popularization book

- Odile Jacob
- with Pierrick Gaudry



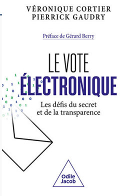
Belenios

- about 1400 elections / years, 100 000+ voters
- ongoing certification CSPN by ANSSI
- ongoing CNIL expertise

Bonus information

A popularization book

- Odile Jacob
- with Pierrick Gaudry



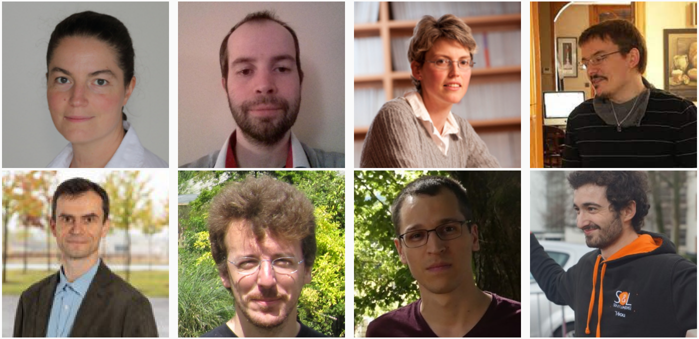
Belenios

- about 1400 elections / years, 100 000+ voters
- ongoing certification CSPN by ANSSI
- ongoing CNIL expertise

Participation to the e-voting elections - législatives 2022

- Approached by MEAE and ANSSI
- Request to be proxy-verifier for individual and universal verifiability
- They need compliance with level 3 of CNIL recommendations

Thank you Hubert!



Barbétretat Juin 2013

Hubert à l’apéro : « On va quand même pas glander !?! »