Anonymity and Rapid Mixing in Cryptographic Protocols

Mirosław Kutyłowski

this talk concerns joint work with Artur Czumaj, Marcin Gomułkiewicz and Marek Klonowski

WARTACRYPT'04
The 4th Central European Conference on Cryptology

Significant achievements of cryptography

- data encryption
- digital signatures
- establishing keys between remote parties
- authentication protocols
- ...
- but problems with anonymity of communication

Communication systems

- messages can be kept secret
- authentication through MAC - nothing can be changed without being noticed
- how to hide that two parties are communicating??

Need of anonymity in communication

- business to business communication
- consumer protection
- privacy protection
- economic and political security of a country

Assumptions about an adversary

Many models possible, each of them might be relevant

- passive
  - adversary can eavesdrop the whole traffic
  - adversary can eavesdrop a constant fraction of traffic
- active – adversary can insert and delete messages
  - everywhere
  - at a constant fraction of nodes

Anonymity techniques – all-to-all

- everybody sends an encoded message to all possible recipients at every moment

Anonymity techniques – all-to-all

- everybody sends an encoded message to all possible recipients at every moment

- works only for a small number of participants
- can be implemented in a token ring

Mixes

- a number of messages enter a mix simultaneously
- they are recoded by a mix
- and permuted at random before outputting
- no connection between input and output can be derived appropriate encoding

Networks of Mixes

- cascades of mixes – mixes run by different parties
- parallel processing – using small mixes to permute large number of packets

Networks of Mixes

- cascades of mixes – mixes run by different parties
- parallel processing – using small mixes to permute large number of packets

Major problem:

- how many mixes are to be used

DC nets

- Gumiś or Mixer wish to send a bit to me without revealing who sends
- they toss a coin, the result is b
- if X does not send a bit, he sends b,
- the sender sends b for transmitting 0, and $1 - b$ for transmitting 1
- decoding: XOR of bits received

DC nets

- Gumiś or Mixer wish to send a bit to me without revealing who sends
- they toss a coin, the result is b
- if X does not send a bit, he sends b,
- the sender sends b for transmitting 0, and $1 - b$ for transmitting 1
- decoding: XOR of bits received

- perfect anonymity
- problems with scalability

Bulletin Board

- a shared broadcast channel
- encrypted messages
- everybody can receive, but who can decode??

Onions

- messages are sent along (random) paths chosen by the sender
- each server on the path knows only the predecessor and the successor on the path
- retreiving any other information (final destination, source,...) from the onion is infeasible

Anonymity

What does *anonymity* mean?

- one cannot deduce a destination of a message sent by a single user

Anonymity

What does *anonymity* mean?

- one cannot deduce a destination of a message sent by a single user
  OR

Anonymity

What does *anonymity* mean?

- one cannot deduce a destination of a message sent by a single user
  OR
- any significant data on the protocol participants cannot be deduced

Why anonymity definition is important

Important case - electronic election schemes

- Eve analyses the votes, and derives probabilities that Alice voted for X, for each single X
- if probability distribution is close to uniform, then the scheme is often told to preserve anonymity.

Why anonymity definition is important

Important case - electronic election schemes

- Eve analyses the votes, and derives probabilities that Alice voted for X, for each single X
- if probability distribution is close to uniform, then the scheme is often told to preserve anonymity.

FALSE!

Why anonymity definition is important

- Eve may be unable to derive preferences of Alice

Why anonymity definition is important

- Eve may be unable to derive preferences of Alice
- but can deduce that Eve and Jurek voted for the same party with probability 99%

Why anonymity definition is important

- Eve may be unable to derive preferences of Alice
- but can deduce that Eve and Jurek voted for the same party with probability 99%
- it remains to buy the information from Jurek

Prior work

Ron Berman, Amos Fiat, Amnon Ta-Shma say:

- *Literally dozens (hundreds?) of papers since, dedicated conferences, etc., etc.*
- *Many implementations*
- *Typical paper:*
  Attack on prior protocol(s)
  Suggest new protocol
  Repeat
- *Very few attempts to give rigorous definitions, let alone proofs*
- *Notable exception: Rackoff and Simon, 1993*

k-anonymity

- used in databases with sensitive information
- each user has to be undistinguishable from some k other users

k-anonymity

- used in databases with sensitive information
- each user has to be undistinguishable from some k other users

Problem:

- low level of anonymity
- suitable if one can control knowledge of an adversary and block further querries

Anonymity set

- let A be the set of all user that are the recipients of a message with a positive probability
- A called the anonymity set of this message
- anonymity measure: the size of A

Anonymity set

- let A be the set of all user that are the recipients of a message with a positive probability
- A called the anonymity set of this message
- anonymity measure: the size of A

Problem:

- if this size is low, then anonymity is poor
- if this size is high, it does not necessarily mean that anonymity is high, probabilities can differ substantially

Highest probabilities

- anonymity measure: **the highest probability in the anonymity set**
- motivation: high probability means there is a quite probable location, even if many other locations are possible

Entropy and anonymity set

- consider probabilities of locations in the anonymity set
- anonymity measure: **entropy of this probability distribution**
- motivation: entropy says how many bits in average are required to specify the location in the anonymity set

Problems with these definitions

- only one user considered
- dependencies among users may be crucial

Problems with these definitions

- only one user considered
- dependencies among users may be crucial

An obvious example: “pseudo-mix”

- input: n messages on positions 0 through $n-1$

Problems with these definitions

- only one user considered
- dependencies among users may be crucial

An obvious example: “pseudo-mix”

- input: n messages on positions 0 through $n-1$
- cryptographic recoding of messages – as usual

Problems with these definitions

- only one user considered
- dependencies among users may be crucial

An obvious example: “pseudo-mix”

- input: n messages on positions 0 through $n-1$
- cryptographic recoding of messages – as usual
- “permuting”:
  - $r < n$ chosen uniformly at random

Problems with these definitions

- only one user considered
- dependencies among users may be crucial

An obvious example: “pseudo-mix”

- input: n messages on positions 0 through $n-1$
- cryptographic recoding of messages – as usual
- “permuting”:
  - $r < n$ chosen uniformly at random
  - a decoded message from position i moved to position $i + r \mod n$.

Problems with these definitions

- only one user considered
- dependencies among users may be crucial

An obvious example: “pseudo-mix”

- input: n messages on positions 0 through $n-1$
- cryptographic recoding of messages – as usual
- “permuting”:
  - $r < n$ chosen uniformly at random
  - a decoded message from position i moved to position $i + r \mod n$.
- a message of adversary reveals r and thereby all anonymity is gone
- well, entropy for a single message is maximal

Traffic analysis

consider a communication network, with (unbreakable) cryptographic recoding of the messages at the network nodes

Traffic analysis

consider a communication network, with (unbreakable) cryptographic recoding of the messages at the network nodes

- how much gains an adversary by observing the traffic?

Traffic analysis

consider a communication network, with (unbreakable) cryptographic recoding of the messages at the network nodes

- how much gains an adversary by observing the traffic?
- sometimes an adversary knows everything (the routes of messages do not cross, while the adversary see all links)

Traffic analysis

consider a communication network, with (unbreakable) cryptographic recoding of the messages at the network nodes

- how much gains an adversary by observing the traffic?
- sometimes an adversary knows everything (the routes of messages do not cross, while the adversary see all links)
- destinations and sources often cannot be hidden, only linking them might be difficult

Viewpoint without traffic information

- for each node the adversary knows:
  - how many messages are initially sent,
  - how many messages are finally delivered

Viewpoint without traffic information

- for each node the adversary knows:
  - how many messages are initially sent,
  - how many messages are finally delivered
- random variable π:
  $\pi(j) = i$ iff the ith message goes to the jth destination place
- probability distribution of π summarizes all information which an adversary can use

View with traffic information

- the same as before, but additionally adversary knows which links have been used for communication

View with traffic information

- the same as before, but additionally adversary knows which links have been used for communication
- sometimes it is evident that a certain message could not be delivered somewhere - no path exists

Probability distribution

- now conditional probabilities:

$$\Pr[\pi|c]$$

where c is traffic information

- different c influence conditional probability in a different way,

Probability distribution

- now conditional probabilities:

$$\Pr[\pi|c]$$

  where c is traffic information
- different c influence conditional probability in a different way,
- goal of anonymity system: conditional probability distribution should be almost the same as the original one, – not always possible

Probability distribution

- now conditional probabilities:

$$\Pr[\pi|c]$$

  where c is traffic information
- different c influence conditional probability in a different way,
- goal of anonymity system: conditional probability distribution should be almost the same as the original one, – not always possible
- modified goal: get this property for almost all c, i.e. whp

Distance of probability distributions

Variation distance

- two probability distributions μ_1 and μ_2 over a finite space Ω
- definition of variation distance:

$$\|\mu_1 - \mu_2\| = \frac{1}{2} \sum_{\omega \in \Omega} |\mu_1(\omega) - \mu_2(\omega)| \ .$$

Anonymity definition based on variation distance

$$\|\Pi - \Pi|C\| \leq \ldots$$

where Π is probability distribution of π,
Π is probability distribution of π conditioned upon traffic information

Anonymity definition based on mutual information

- information theoretic approach
- roughly speaking: how many information bits on Π is given by C

Anonymity definition based on mutual information

- information theoretic approach
- roughly speaking: how many information bits on Π is given by C
- roughly equivalent to the previous definition – conversions possible (Berman, Fiat, Ta-Shma)

Results on the onion protocol

Adversary with full knowledge on the traffic

Results on the onion protocol

Adversary with full knowledge on the traffic

- Rackoff, Simon (ACM STOC'93):
  polylogarithmic time (degree 11),
  special assumption: at stage i the messages stay inside groups of cardinality 2^i

Results on the onion protocol

Adversary with full knowledge on the traffic

- Rackoff, Simon (ACM STOC'93): polylogarithmic time (degree 11), special assumption: at stage i the messages stay inside groups of cardinality 2^i
- Czumaj, Kanarek, Kutyłowski, Loryś (ACM SODA'99): under the same assumptions - time $O(\log^2 n)$

Results on the onion protocol

Adversary with partial knowledge on the traffic

Results on the onion protocol

Adversary with partial knowledge on the traffic

- Berman, Fiat, Ta-Shma (FC'2004) – adversary model, $O(\log^4 n)$ steps for n messages and variation distance $1/n$

Results on the onion protocol

Adversary with partial knowledge on the traffic

- Berman, Fiat, Ta-Shma (FC'2004) – adversary model, $O(\log^4 n)$ steps for n messages and variation distance $1/n$
- Gomułkiewicz, Klonowski, Kutyłowski (ISC'2004) – $O(\log n)$ steps,
  optimal result

Rapid mixing and anonymity

consider a stochastic process of transmitting messages at random

- at every step the messages are recoded at the nodes and
- sent further to a random destination (chosen independently)

Rapid mixing and anonymity

consider a stochastic process of transmitting messages at random

- at every step the messages are recoded at the nodes and
- sent further to a random destination (chosen independently)
- the adversary can see where the messages are sent (conditional probabilities are considered)

Rapid mixing and anonymity

consider a stochastic process of transmitting messages at random

- at every step the messages are recoded at the nodes and
- sent further to a random destination (chosen independently)
- the adversary can see where the messages are sent (conditional probabilities are considered)

How many steps are needed until probability distribution becames close to the uniform distribution?

Stationary distribution

- a probability distribution over the set of states is **stationary** if applying a single step of the process does not change the probability distribution,

Stationary distribution

- a probability distribution over the set of states is **stationary** if applying a single step of the process does not change the probability distribution,
- example: initially: a uniform distribution over permutations of k elements,
  apply a permutation chosen according to some distribution S
  result: again a uniform distribution over the set of permutations of k elements.

Rapid mixing techniques

- given a stochastic process $\mathcal{P}$ with a uniform distribution u
- show that after t steps the probability distribution of the process started in an arbitrary state is close to u

Rapid mixing techniques

- given a stochastic process $\mathcal{P}$ with a uniform distribution u
- show that after t steps the probability distribution of the process started in an arbitrary state is close to u

How to construct such a proof?

Coupling techniques

- define two processes $\mathcal{P}_A, \mathcal{P}_B$
- both are the copies of $\mathcal{P}$,

Coupling techniques

- define two processes $\mathcal{P}_A, \mathcal{P}_B$
- both are the copies of $\mathcal{P}$,
- but the choices of the first process may influence the second process

Coupling goal

- define dependencies so that the processes “converge”
  – (with probabilities growing with the number of steps) they reach the same state

Coupling goal

- define dependencies so that the processes “converge” – (with probabilities growing with the number of steps) they reach the same state
- key property – coupling lemma:

$$\text{variation distance after } t \text{ steps} \leq \mathrm{Pr}[\mathcal{P}_A \text{ and } \mathcal{P}_B \text{ differ after } t \text{ steps}].$$

Why coupling lemma holds?

- let $\mathcal{P}_B$ be started according to stationary distribution

Why coupling lemma holds?

- let $\mathcal{P}_B$ be started according to stationary distribution
- by definition of stationary distribution $\mathcal{P}_B$ will stay in this distribution after each step

Why coupling lemma holds?

- let $\mathcal{P}_B$ be started according to stationary distribution
- by definition of stationary distribution $\mathcal{P}_B$ will stay in this distribution after each step

what about $\mathcal{P}_A$?

- start $\mathcal{P}_A$ in an arbitrary state
- .. and use dependencies defined by coupling

Why coupling lemma holds?

- key point: if probability that two processes differ is at most p then probability distributions cannot differ by more than p.

Let's use coupling

- a universal tool for showing convergence
- no expertise in stochastic processes necessary - only combinatorial skills

Path coupling

- it suffices to consider processes that are almost in the same state
  - distances between process states should be defined
  - it suffices to consider pair of processes at distance 1

Example - anonymity for Chaum's scheme of electronic elections

- **proving security** of voters (well, with high probability)
- Gomułkiewicz, Klonowski, and myself, ESORICS'2003

Chaums's scheme

- visual cryptography for convincing voters
- essential point: decoding of votes
- several decoding authorities:
  - Authority 1 decodes all votes, permutes at random, the results given to Authority 2
  - Authority 2 decodes all votes, permutes at random, the results given to Authority 3
  - ...

Checking Authorities

- Authorities have to prove honesty of decoding and permuting
- selective proof (Randomized Partial Checking): for 50% randomly chosen positions permutation values must be revealed
- privacy concerns: may be it guarantees honesty of Authorities but at the price of voter's anonymity?

Checking Authorities

Modelling Randomized Partial Checking

after simple reformulation we get a process in which during a step

- elements on positions 1 through $n/2$ are permuted at random,
- elements on positions $n/2+1$ through n are permuted at random,
- a single permutation is applied to all elements
  even if this permutation is random, it is fixed when process is defined

Coupling proof

- after the first step we have “white” and “black” elements,
- path coupling: consider the states which differ by just one transposition τ

Definition of coupling

- if differences inside the same half, then define dependence:
  - if the first process chooses permutation ρ in this half,
  - then the second process chooses $\rho \circ \tau$
- with such a dependence the difference dissappears

Definition of coupling - difficult case

- the first process has an extra black element in the first half, the second process has an extra black element in the second half
- it does not work as before

How to couple? In two steps!

Constructing a coupling

- from the first half of white elements will go to the second half, while the rest will remain,
- similarly for the second half

Constructing a coupling

- from the first half of white elements will go to the second half, while the rest will remain,
- similarly for the second half
- solution idea: exchange the location of the extra black item of the second process with the places of white elements

Constructing a coupling

- constructing dependencies:
  - if the extra black element of the first process will go to **another** half
    then the extra black element of the second process takes a place to **remain** in its half

Constructing a coupling

- constructing dependencies:
  - if the extra black element of the first process will go to **another** half
    then the extra black element of the second process takes a place to **remain** in its half
  - if the extra black element of the first process will **remain** in its half
    then the extra black element of the second process takes a place to go to **another** half

Constructing a coupling

- constructing dependencies:
  - if the extra black element of the first process will go to **another** half
    then the extra black element of the second process takes a place to **remain** in its half
  - if the extra black element of the first process will **remain** in its half
    then the extra black element of the second process takes a place to go to **another** half
- minor technical difficulties: white elements do not split evenly between those that stay in the same half and those that go

Thanks for your attention!