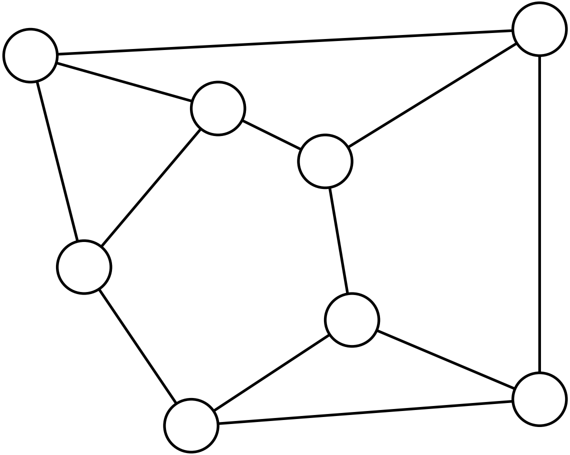
Local View Attack

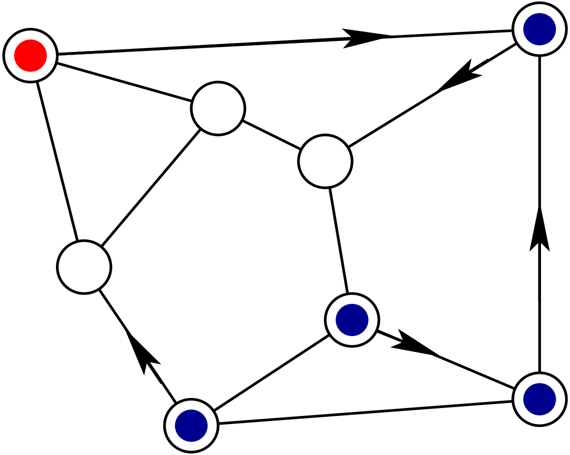
Marcin Gogolewski Marek Klonowski
Mirosław Kutyłowski

Wrocław University of Technology

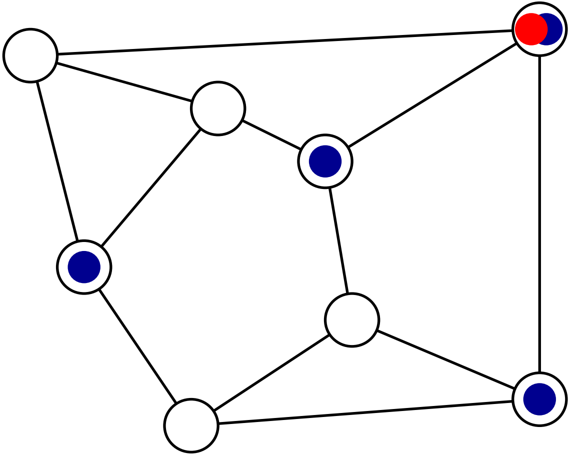
Idea of Mixing



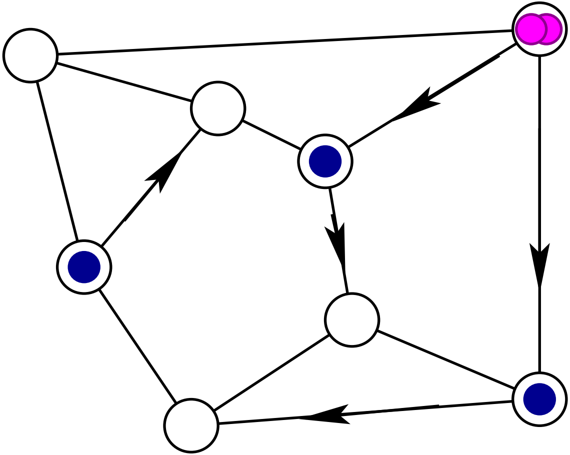
Idea of Mixing



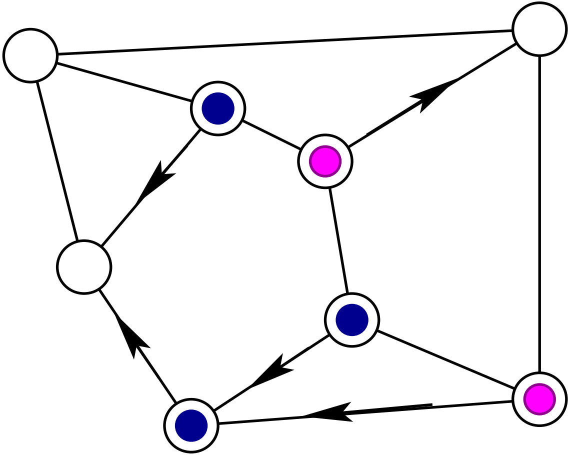
Idea of Mixing



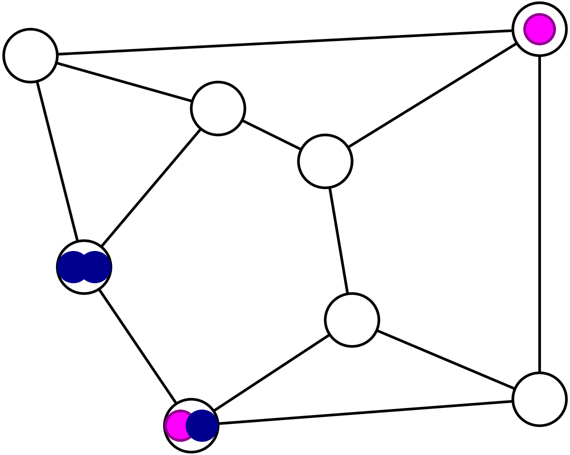
Idea of Mixing



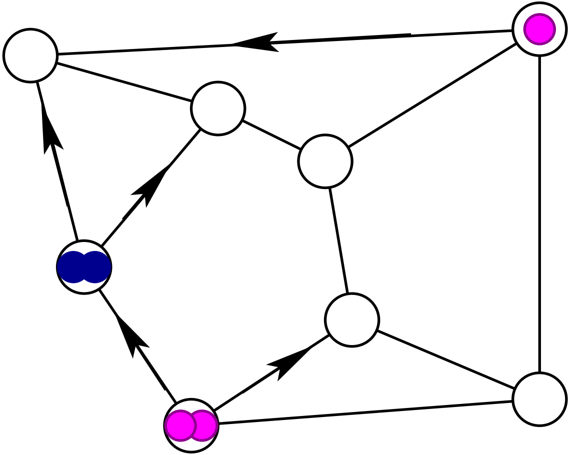
Idea of Mixing



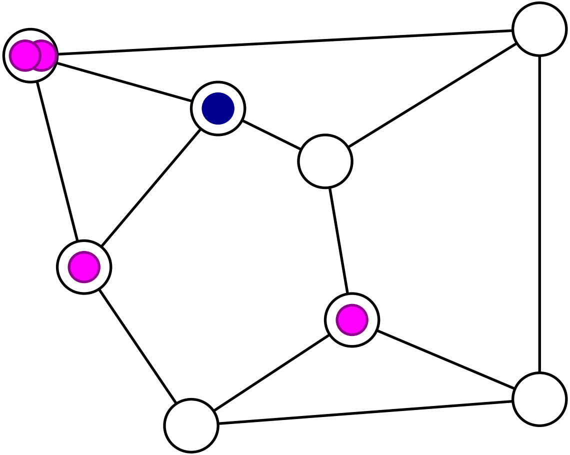
Idea of Mixing



Idea of Mixing



Idea of Mixing



Onions

1. a user A sending M to B determines a path C_1, C_2, ..., C_λ, B, where each C_i is chosen **independently at random with uniform probability distribution**
2. an onion O containing M is created as follows:

$$\begin{aligned} O_\lambda &= E_B(M, \mathrm{random}_{\lambda+1}) \\ O_i &= E_{Pub_i}(C_{i+1}, O_{i+1}, \mathrm{random}_i) \quad \text{for } i < \lambda \\ O &= O_1 \end{aligned}$$

3. so it looks like an onion with many layers: $E_{Pub_2}(\ldots E_{Pub_{\lambda-1}}(C_\lambda, E_B(M, \ldots), \ldots), \ldots)$

Onions

1. a user A sending M to B determines a path C_1, C_2, ..., C_λ, B, where each C_i is chosen **independently at random with uniform probability distribution**
2. an onion O containing M is created as follows:

$$\begin{aligned} O_\lambda &= E_B(M, \mathrm{random}_{\lambda+1}) \\ O_i &= E_{Pub_i}(C_{i+1}, O_{i+1}, \mathrm{random}_i) \ \ \text{for } i < \lambda \\ O &= O_1 \end{aligned}$$

3. so it looks like an onion with many layers: $E_{Pub_2}(\ldots E_{Pub_{\lambda-1}}(C_\lambda, E_B(M, \ldots), \ldots), \ldots)$

Onions

1. a user A sending M to B determines a path C_1, C_2, ..., C_λ, B, where each C_i is chosen **independently at random with uniform probability distribution**
2. an onion O containing M is created as follows:

$$
\begin{aligned}
O_\lambda &= E_B(M, \mathrm{random}_{\lambda+1}) \\
O_i &= E_{Pub_i}(C_{i+1}, O_{i+1}, \mathrm{random}_i) \quad \text{for } i < \lambda \\
O &= O_1
\end{aligned}
$$

3. so it looks like an onion with many layers: $E_{Pub_2}(\ldots E_{Pub_{\lambda-1}}(C_\lambda, E_B(M, \ldots), \ldots), \ldots)$

Onions

In fact, some additional measures might be necessary:

- timestamps (for preventing repetition attack)
- uniform onion length
- uniform distribution
- ...

Processing of Onions

1. several messages enters a server
2. they are recoded cryptographically:
   - one layer is removed from each onion by decoding with the private key

$$\underbrace{E_{Pub_i}(C_{i+1}, \overbrace{E_{Pub_{i+1}}(\ldots)}^{O_{i+1}}, \ldots)}_{O_i}$$

   - the address for the next hop is retrieved,
   - and the onion O_{i+1} to be sent there
   - the new onions are sent to the next hop locations

Processing of Onions

1. several messages enters a server
2. they are recoded cryptographically:
   - one layer is removed from each onion by decoding with the private key

$$\underbrace{E_{Pub_i}(C_{i+1}, \overbrace{E_{Pub_{i+1}}(\ldots)}^{O_{i+1}}, \ldots)}_{O_i}$$

     - the address for the next hop is retrieved,
     - and the onion O_{i+1} to be sent there
   - the new onions are sent to the next hop locations

Idea of Mixing with Onions

if two onions enter the same server, then they “mix”:

- a proper encoding ensures that without the private key of the server one cannot link the incoming and the outgoing onions

Idea of Mixing with Onions

- if there are sufficiently many onions, then they meet quite often
- if two onions meet, then they “mix”
- if the paths are long enough, then there are enough “mixing” so that an adversary cannot find anymore who is communicating with whom

Idea of Mixing with Onions

- if there are sufficiently many onions, then they meet quite often
- if two onions meet, then they “mix”
- if the paths are long enough, then there are enough “mixing” so that an adversary cannot find anymore who is communicating with whom

Idea of Mixing with Onions

- if there are sufficiently many onions, then they meet quite often
- if two onions meet, then they “mix”
- if the paths are long enough, then there are enough “mixing” so that an adversary cannot find anymore who is communicating with whom

Central Question

- how big must be the path length λ so that anonymity goals are reached?

The intuition is that

- a small λ should be enough,
- anonymity level grows with λ so that for a big λ the adversary cannot get any information.

Central Question

- how big must be the path length λ so that anonymity goals are reached?

The intuition is that

- a small λ should be enough,
- anonymity level grows with λ so that for a big λ the adversary cannot get any information.

Network Assumptions

Different connectivities models:

- complete connection graph (every node can be the next hop)
- sparse connection graph

Global versus local view:

- everybody knows all servers (*global view*), or
- each node knows only a specific subset of nodes (*local view*)

Network Assumptions

Different connectivities models:

- complete connection graph (every node can be the next hop)
- sparse connection graph

Global versus local view:

- everybody knows all servers (*global view*), or
- each node knows only a specific subset of nodes (*local view*)

Network Assumptions

Different connectivities models:

- complete connection graph (every node can be the next hop)
- sparse connection graph

Global versus local view:

- everybody knows all servers (*global view*), or
- each node knows only a specific subset of nodes (*local view*)

Network Assumptions

Different connectivities models:

- complete connection graph (every node can be the next hop)
- sparse connection graph

Global versus local view:

- everybody knows all servers (*global view*), or
- each node knows only a specific subset of nodes (*local view*)

Traffic Analysis

An adversary tries to break anonymity features

1. he collects traffic information (in a passive or an active way)
2. he makes computations resulting with some substantial information on probability distribution of possible destinations of a message (or a group of messages), **which is not known before**

- a protocol is good if this additional knowledge through traffic analysis is marginal

Traffic Analysis

An adversary tries to break anonymity features

1. he collects traffic information (in a passive or an active way)
2. he makes computations resulting with some substantial information on probability distribution of possible destinations of a message (or a group of messages), **which is not known before**

- a protocol is good if this additional knowledge through traffic analysis is marginal

Traffic Analysis

An adversary tries to break anonymity features

1. he collects traffic information (in a passive or an active way)
2. he makes computations resulting with some substantial information on probability distribution of possible destinations of a message (or a group of messages), **which is not known before**

- a protocol is good if this additional knowledge through traffic analysis is marginal

Adversary Models

- global passive - the adversary can see the whole traffic (Rackoff, Simon)
- limited passive - the adversary can monitor only a constant fraction of connections established in advance (Berman, Fiat, Ta-Shma)
- global active - the adversary can insert, delete and modify messages

Adversary Models

- global passive - the adversary can see the whole traffic (Rackoff, Simon)
- limited passive - the adversary can monitor only a constant fraction of connections established in advance (Berman, Fiat, Ta-Shma)
- global active - the adversary can insert, delete and modify messages

Adversary Models

- global passive - the adversary can see the whole traffic (Rackoff, Simon)
- limited passive - the adversary can monitor only a constant fraction of connections established in advance (Berman, Fiat, Ta-Shma)
- global active - the adversary can insert, delete and modify messages

Provable Anonymity

Estimations on the parameter λ sufficient to that traffic analysis does not reveal almost any information

global passive adversary

- Rackoff, Simon: λ polylogarithmic in the number of nodes, heavy traffic, an extra assumption about paths, STOC'93
- Czumaj, Kutyłowski: $\lambda = O(\log^2 n)$ is enough, SODA'98 (no full version published)

limited passive adversary

- Berman, Fiat, Ta-Shma: $\lambda = O(\log^4 n)$ is enough, FC'2004
- Gomułkiewicz, Klonowski, Kutyłowski: $\lambda = \Theta(\log n)$, ISC'2004

Provable Anonymity

Estimations on the parameter λ sufficient to that traffic analysis does not reveal almost any information

global passive adversary

- Rackoff, Simon: λ polylogarithmic in the number of nodes, heavy traffic, an extra assumption about paths, STOC'93
- Czumaj, Kutyłowski: $\lambda = O(\log^2 n)$ is enough, SODA'98 (no full version published)

limited passive adversary

- Berman, Fiat, Ta-Shma: $\lambda = O(\log^4 n)$ is enough, FC'2004
- Gomułkiewicz, Klonowski, Kutyłowski: $\lambda = \Theta(\log n)$, ISC'2004

Provable Anonymity

Estimations on the parameter λ sufficient to that traffic analysis does not reveal almost any information

global passive adversary

- Rackoff, Simon: λ polylogarithmic in the number of nodes, heavy traffic, an extra assumption about paths, STOC'93
- Czumaj, Kutyłowski: $\lambda = O(\log^2 n)$ is enough, SODA'98 (no full version published)

limited passive adversary

- Berman, Fiat, Ta-Shma: $\lambda = O(\log^4 n)$ is enough, FC'2004
- Gomułkiewicz, Klonowski, Kutyłowski: $\lambda = \Theta(\log n)$, ISC'2004

Provable Anonymity

Estimations on the parameter λ sufficient to that traffic analysis does not reveal almost any information

global passive adversary

- Rackoff, Simon: λ polylogarithmic in the number of nodes, heavy traffic, an extra assumption about paths, STOC'93
- Czumaj, Kutyłowski: $\lambda = O(\log^2 n)$ is enough, SODA'98 (no full version published)

limited passive adversary

- Berman, Fiat, Ta-Shma: $\lambda = O(\log^4 n)$ is enough, FC'2004
- Gomułkiewicz, Klonowski, Kutyłowski: $\lambda = \Theta(\log n)$, ISC'2004

Anonymity Set

For a message M its *anonymity set* is
the set of possible locations of an onion containing M at a given moment. (D.Kesdogan)

It is necessary that at least anonymity set of each message is big.

Anonymity Set

For a message M its *anonymity set* is
the set of possible locations of an onion containing M at a given moment. (D.Kesdogan)

It is necessary that at least anonymity set of each message is big.

Problem

The results above concerning global passive adversary use the assumption that nodes on the paths are chosen independently at random **from the same set of nodes by each user**.

1. anonymous referee of some other paper says: *it does not matter, as long as the sets used are large*
2. other people expect that it matters

Problem

The results above concerning global passive adversary use the assumption that nodes on the paths are chosen independently at random **from the same set of nodes by each user**.

1. anonymous referee of some other paper says: *it does not matter, as long as the sets used are large*
2. other people expect that it matters

Main Results – Overview

1. Attacks for the case when the sets of servers known by the users differ.
2. Phase transition phenomenon: if a user knows less than $\approx 52\%$ of servers, and knowledge of others is independent of other users, then anonymity breaks down.
3. Above the phase transition point $\approx 52\%$ the anonymity set starts to grow almost linearly.
4. Except for very small values the size of anonymity set of a message does not grow with λ!

Main Results – Overview

1. Attacks for the case when the sets of servers known by the users differ.
2. Phase transition phenomenon: if a user knows less than $\approx 52\%$ of servers, and knowledge of others is independent of other users, then anonymity breaks down.
3. Above the phase transition point $\approx 52\%$ the anonymity set starts to grow almost linearly.
4. Except for very small values the size of anonymity set of a message does not grow with λ!

Main Results – Overview

1. Attacks for the case when the sets of servers known by the users differ.
2. Phase transition phenomenon: if a user knows less than $\approx 52\%$ of servers, and knowledge of others is independent of other users, then anonymity breaks down.
3. Above the phase transition point $\approx 52\%$ the anonymity set starts to grow almost linearly.
4. Except for very small values the size of anonymity set of a message does not grow with λ!

Main Results – Overview

1. Attacks for the case when the sets of servers known by the users differ.
2. Phase transition phenomenon: if a user knows less than $\approx 52\%$ of servers, and knowledge of others is independent of other users, then anonymity breaks down.
3. Above the phase transition point $\approx 52\%$ the anonymity set starts to grow almost linearly.
4. Except for very small values the size of anonymity set of a message does not grow with λ!

Case: Alice does not Know all Servers

Assumptions:

- W – the set of servers known by Alice
- N – the set of all servers
- $|W|/|N| < c$, c is a constant
- the other users know N
  (or know a smaller random subset that has been chosen independently from W)
- each server generates exactly one onion

Idea

Consider message M sent by Alice

1. Let position A belong to the anonymity set $\mathcal{A}$ of M at step t. Consider onions sent out of $\mathcal{A}$ at step $t+1$
   1. if an onion goes to a position from $N \setminus W$, then it does not contain M,
      its destination is not included in the $\mathcal{A}$ after step $t+1$,
   2. if a onion goes into some server B in W, then we have to include B in the anonymity set $\mathcal{A}$ after step $t+1$
2. the anonymity set can both grow and shrink at step $t+1$

Idea

Consider message M sent by Alice

1. Let position A belong to the anonymity set $\mathcal{A}$ of M at step t. Consider onions sent out of $\mathcal{A}$ at step $t+1$
   1. if an onion goes to a position from $N \setminus W$, then it does not contain M,
      its destination is not included in the $\mathcal{A}$ after step $t+1$,
   2. if a onion goes into some server B in W, then we have to include B in the anonymity set $\mathcal{A}$ after step $t+1$
2. the anonymity set can both grow and shrink at step $t+1$

Idea

Consider message M sent by Alice

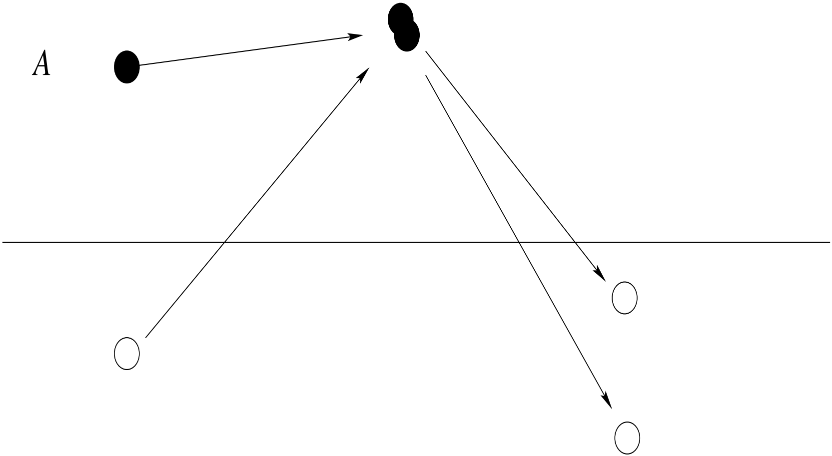
1. Let position A belong to the anonymity set $\mathcal{A}$ of M at step t. Consider onions sent out of $\mathcal{A}$ at step $t+1$
    1. if an onion goes to a position from $N \setminus W$, then it does not contain M,
       its destination is not included in the $\mathcal{A}$ after step $t+1$,
    2. if a onion goes into some server B in W, then we have to include B in the anonymity set $\mathcal{A}$ after step $t+1$
2. the anonymity set can both grow and shrink at step $t+1$

Idea

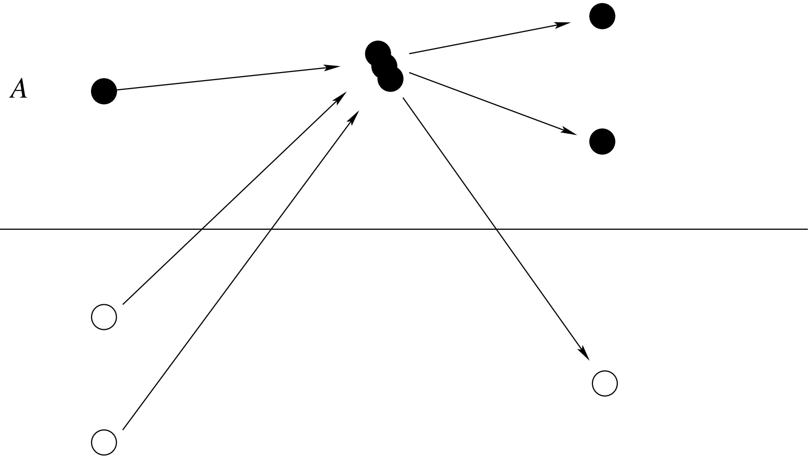
Consider message M sent by Alice

1. Let position A belong to the anonymity set $\mathcal{A}$ of M at step t. Consider onions sent out of $\mathcal{A}$ at step $t+1$
    1. if an onion goes to a position from $N \setminus W$, then it does not contain M,
       its destination is not included in the $\mathcal{A}$ after step $t+1$,
    2. if a onion goes into some server B in W, then we have to include B in the anonymity set $\mathcal{A}$ after step $t+1$
2. the anonymity set can both grow and shrink at step $t+1$

A Reason for Shrinking of Anonymity Set



A Reason for Expansion of Anonymity Set



Size of Anonymity Set

Fluctuations of the size of the anonymity set

- like branching process
- the process cannot die – one onion actually holds the message from Alice!
- if the anonymity set has cardinality m, then the expected change of the size of anonymity set:

$$\approx (n-m) \cdot \frac{|W|}{n} \cdot (1 - e^{-\frac{m-1}{n} - \frac{1}{|W|}}) - m \cdot (1 - \frac{|W|}{n}) + (1 - \frac{|W|}{n})$$

  - the first term make the size increase for small m
  - the second term make the size decrease for large m
- where is the equilibrium where the expected change is 0?

Size of Anonymity Set

Fluctuations of the size of the anonymity set

- like branching process
- the process cannot die – one onion actually holds the message from Alice!
- if the anonymity set has cardinality m, then the expected change of the size of anonymity set:

$$\approx (n-m)\cdot\frac{|W|}{n}\cdot(1-e^{-\frac{m-1}{n}-\frac{1}{|W|}})-m\cdot(1-\frac{|W|}{n})+(1-\frac{|W|}{n})$$

  - the first term make the size increase for small m
  - the second term make the size decrease for large m
- where is the equilibrium where the expected change is 0?

Size of Anonymity Set

Fluctuations of the size of the anonymity set

- like branching process
- the process cannot die – one onion actually holds the message from Alice!
- if the anonymity set has cardinality m, then the expected change of the size of anonymity set:

$$\approx (n-m) \cdot \frac{|W|}{n} \cdot (1 - e^{-\frac{m-1}{n} - \frac{1}{|W|}}) - m \cdot (1 - \frac{|W|}{n}) + (1 - \frac{|W|}{n})$$

  - the first term make the size increase for small m
  - the second term make the size decrease for large m
- where is the equilibrium where the expected change is 0?

Size of Anonymity Set

Fluctuations of the size of the anonymity set

- like branching process
- the process cannot die – one onion actually holds the message from Alice!
- if the anonymity set has cardinality m, then the expected change of the size of anonymity set:

$$\approx (n-m)\cdot\frac{|W|}{n}\cdot(1-e^{-\frac{m-1}{n}-\frac{1}{|W|}})-m\cdot(1-\frac{|W|}{n})+(1-\frac{|W|}{n})$$

  - the first term make the size increase for small m
  - the second term make the size decrease for large m
- where is the equilibrium where the expected change is 0?

Size of Anonymity Set

Fluctuations of the size of the anonymity set

- like branching process
- the process cannot die – one onion actually holds the message from Alice!
- if the anonymity set has cardinality m, then the expected change of the size of anonymity set:

$$\approx (n-m)\cdot\frac{|W|}{n}\cdot(1-e^{-\frac{m-1}{n}-\frac{1}{|W|}})-m\cdot(1-\frac{|W|}{n})+(1-\frac{|W|}{n})$$

  - the first term make the size increase for small m
  - the second term make the size decrease for large m
- where is the equilibrium where the expected change is 0?

Size of Anonymity Set

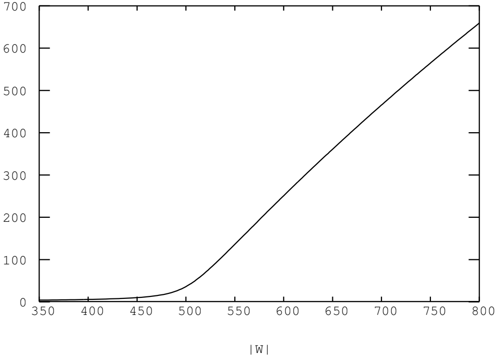
Fluctuations of the size of the anonymity set

- like branching process
- the process cannot die – one onion actually holds the message from Alice!
- if the anonymity set has cardinality m, then the expected change of the size of anonymity set:

$$\approx (n-m)\cdot\frac{|W|}{n}\cdot(1-e^{-\frac{m-1}{n}-\frac{1}{|W|}})-m\cdot(1-\frac{|W|}{n})+(1-\frac{|W|}{n})$$

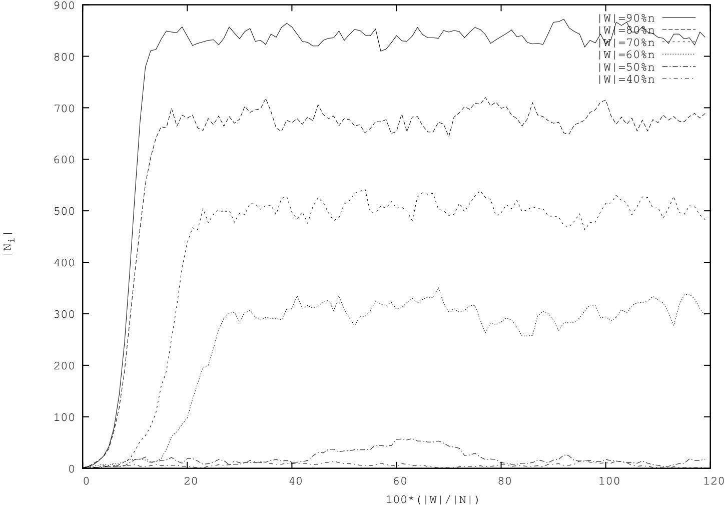
  - the first term make the size increase for small m
  - the second term make the size decrease for large m
- where is the equilibrium where the expected change is 0?

Plot of the Equilibrium Values



A network with 1000 nodes, x-axis: $|W|$, y-axis: equilibrium point

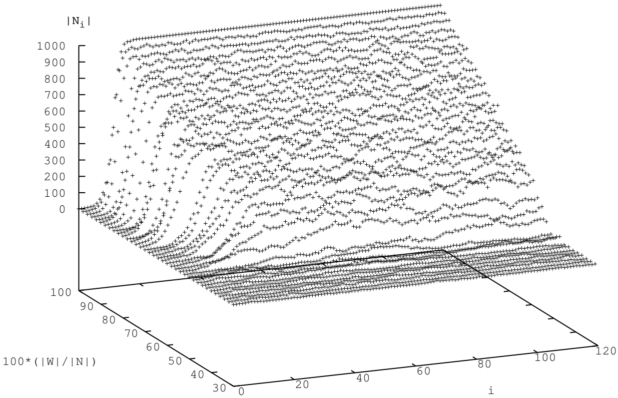
Size of Anonymity Set – Simulations



Different curves for different ratios $|W|/|N|$

Visualization

let us inspect 3D depiction of the experimental data



Skewed Probabilities

- probabilities of holding M are highly nonuniform in the anonymity set,
- for $|N| = 1000$, $|W| = 700$ we have still a fair chance to point to the position of M, if we take, say, the best 30 positions from the anonymity set.

Conclusions

- it is hard to achieve the same view of the network (it may evolve! immediate informing of the changes is problematic)
- if the network load is not heavy, be very careful with the global passive adversary,
- in the case of partial passive adversary everything is much safer

Conclusions

- it is hard to achieve the same view of the network (it may evolve! immediate informing of the changes is problematic)
- if the network load is not heavy, be very careful with the global passive adversary,
- in the case of partial passive adversary everything is much safer

Conclusions

- it is hard to achieve the same view of the network (it may evolve! immediate informing of the changes is problematic)
- if the network load is not heavy, be very careful with the global passive adversary,
- in the case of partial passive adversary everything is much safer