

# Trust in Crowds

## Probabilistic Behaviour in Anonymity Protocols

**Vladimiro Sassone**  
University of Southampton

**TGC 2010**  
**München 2010.2.24**

**(based on joint work with S. Hamadou & E. ElSalamouny)**

# Introduction

## Anonymity in Social Networks



Social Networks: very easy to collect private and sensitive information about individuals.

# Introduction

## Anonymity in Social Networks



**Social Networks:** very easy to collect private and sensitive information about individuals.



# Introduction

## Anonymity in Web Transactions



# Introduction

## Anonymity in Web Transactions





# Introduction

## Anonymity in Web Transactions



# Introduction

## Anonymity in Web Transactions



Google is watching you!



Google™



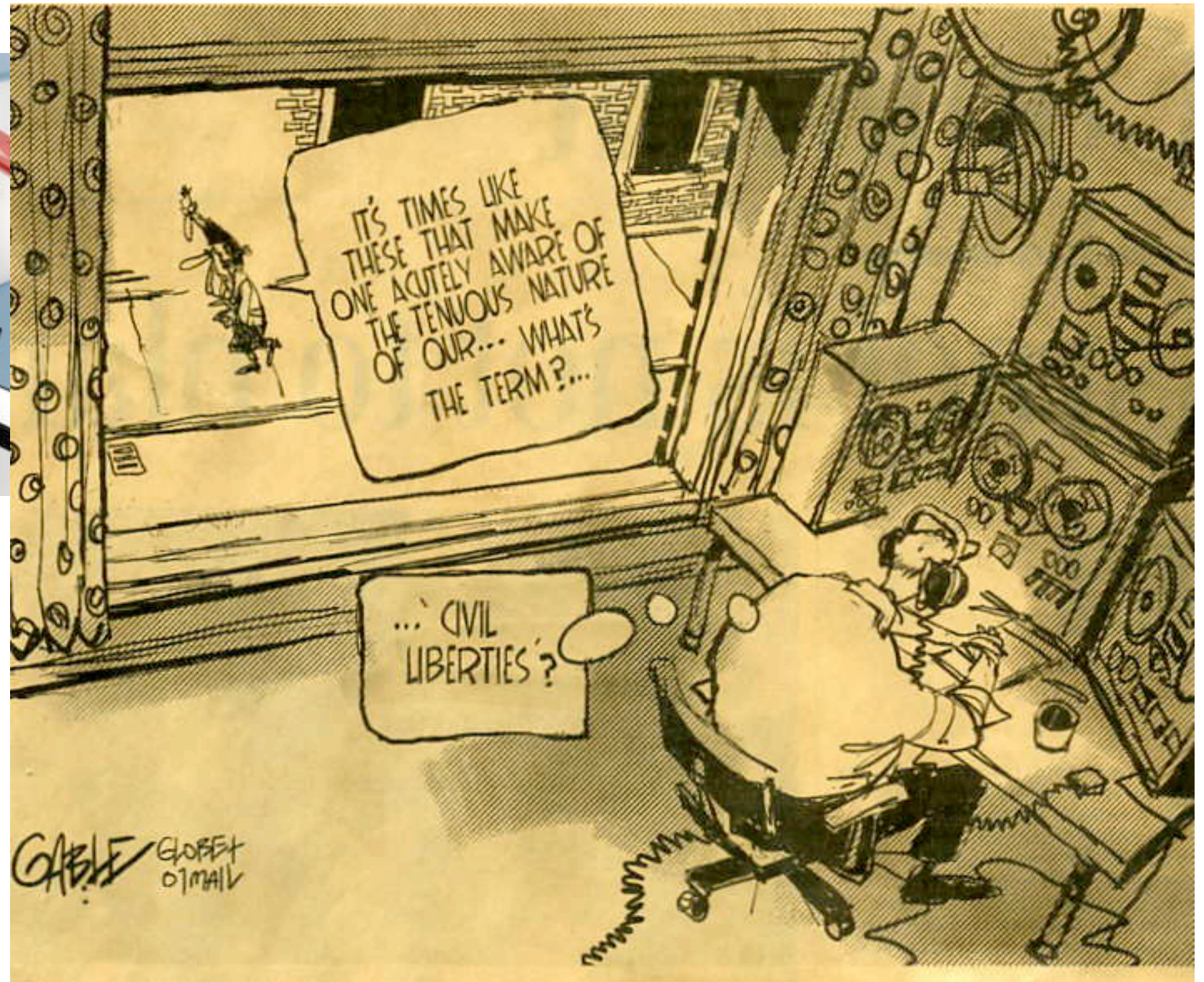


# Introduction

## Anonymity in Web Transactions



Google™





# Introduction

## Data Confidentiality



# Introduction

## Data Confidentiality



...of course, but also...





# Introduction

## Data Confidentiality



...of course, but also...



deduce high input from low output, in the fashion of information flow

# Introduction

## Anonymity Protocols (in general)

- ✦ Aims at obfuscating the link between private input (anonymous actions) and public (observable) output
- ✦ Attacker tries to infer the hidden info from his observation of the protocol



# This presentation

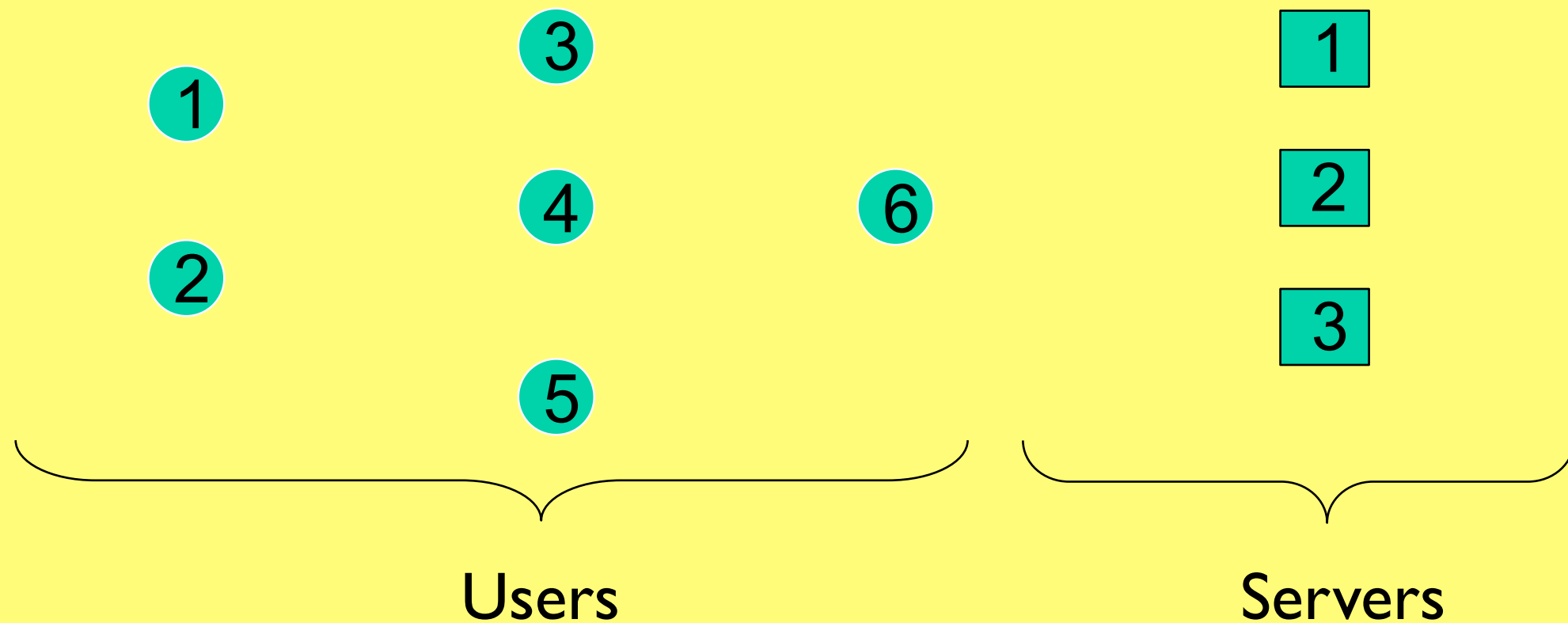
## Trust in the Crowds anonymity protocol

- ✦ Extend the Crowds protocol to a scenario where:
  - ✦ Each principal may suddenly become corrupt.
  - ✦ Principal behaviour is influenced by a trust relationship.
- ✦ Work:
  - ✦ Study the impact of these assumptions on the protocol.
  - ✦ Establish necessary and sufficient criteria for choosing a policy able to achieve a desired level of privacy.

# Crowds

## The protocol

✦ **Crowds** [Reiter and Rubin 1998]: allows internet users to perform anonymous web transactions.

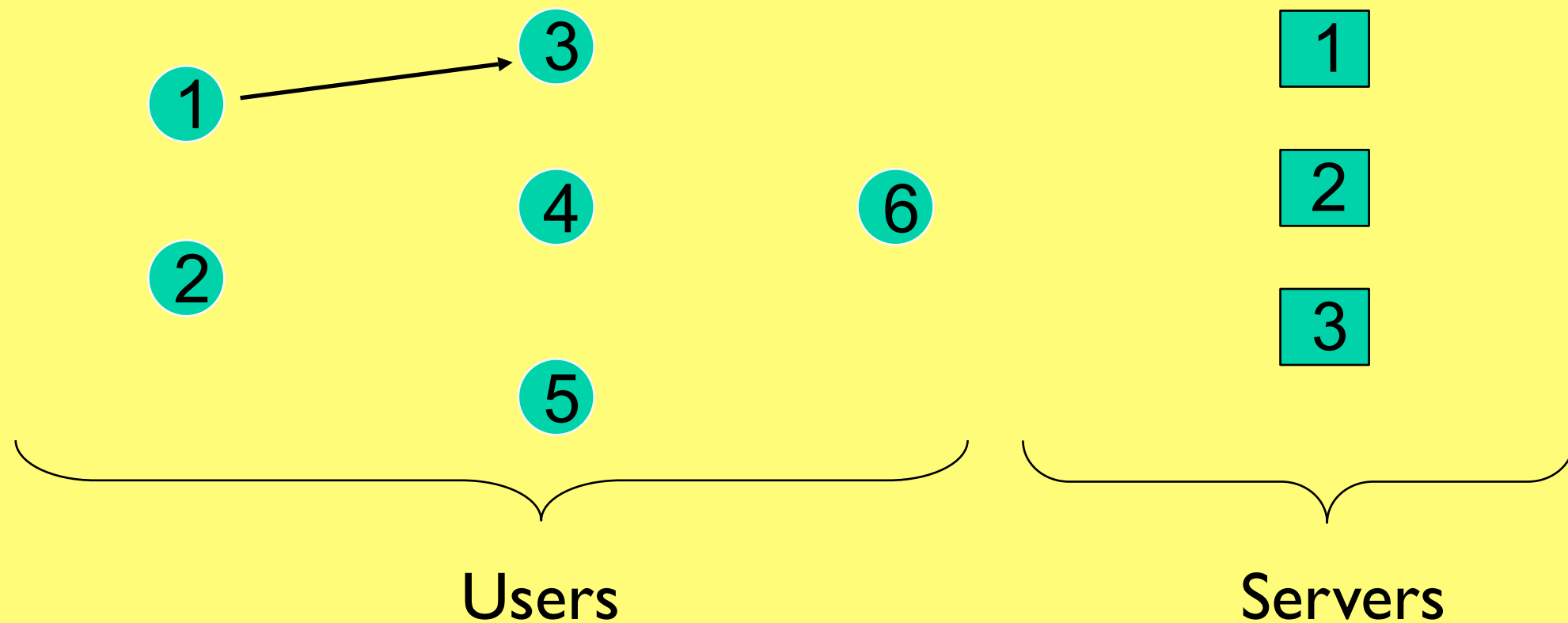




# Crowds

## The protocol

✦ **Crowds** [Reiter and Rubin 1998]: allows internet users to perform anonymous web transactions.

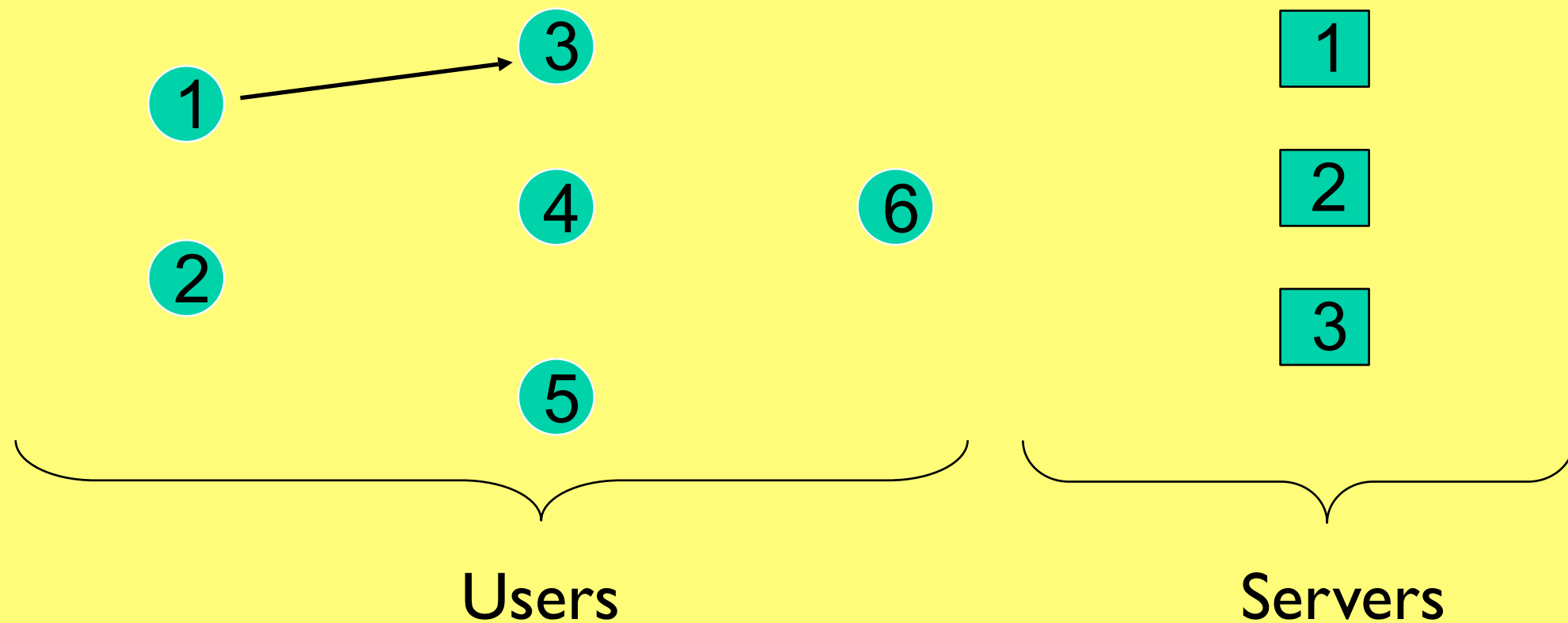


# Crowds

## The protocol

✦ **Crowds** [Reiter and Rubin 1998]: allows internet users to perform anonymous web transactions.

Flips a biased coin  $p_f$



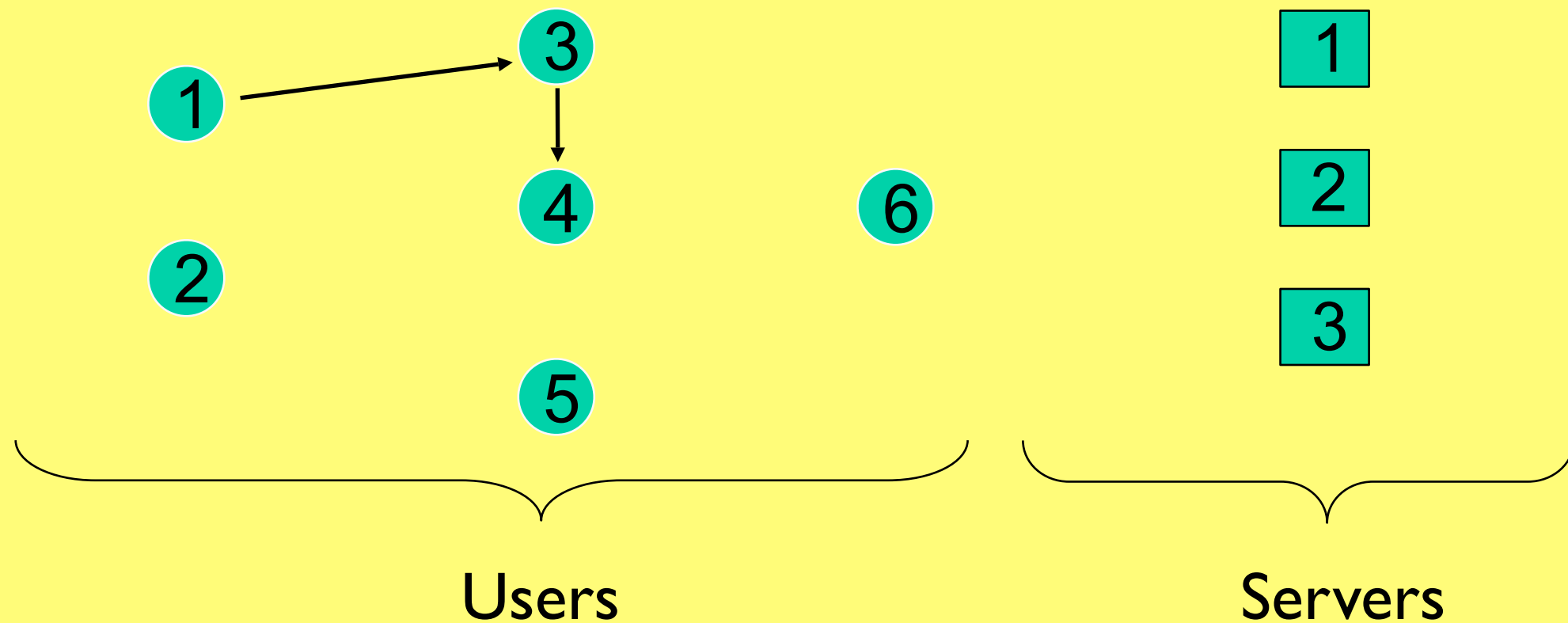


# Crowds

## The protocol

✦ **Crowds** [Reiter and Rubin 1998]: allows internet users to perform anonymous web transactions.

Flips a biased coin  $p_f$

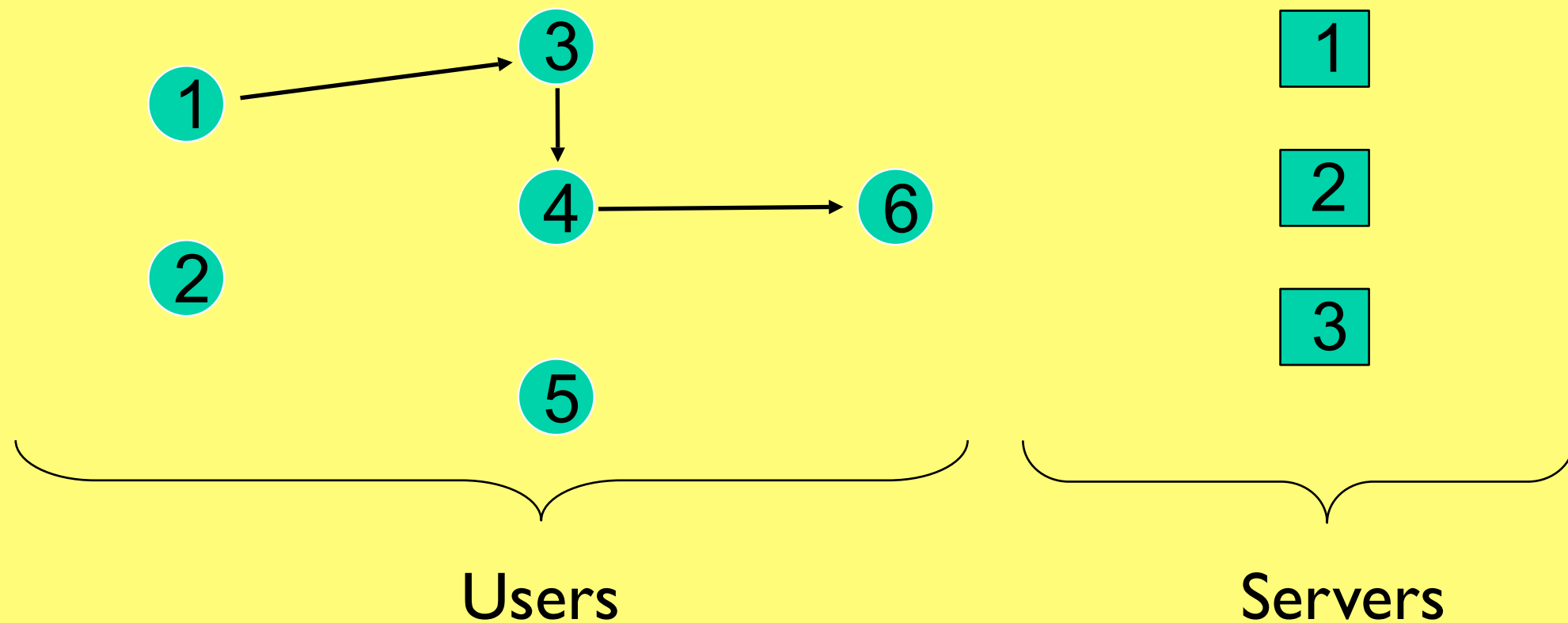


# Crowds

## The protocol

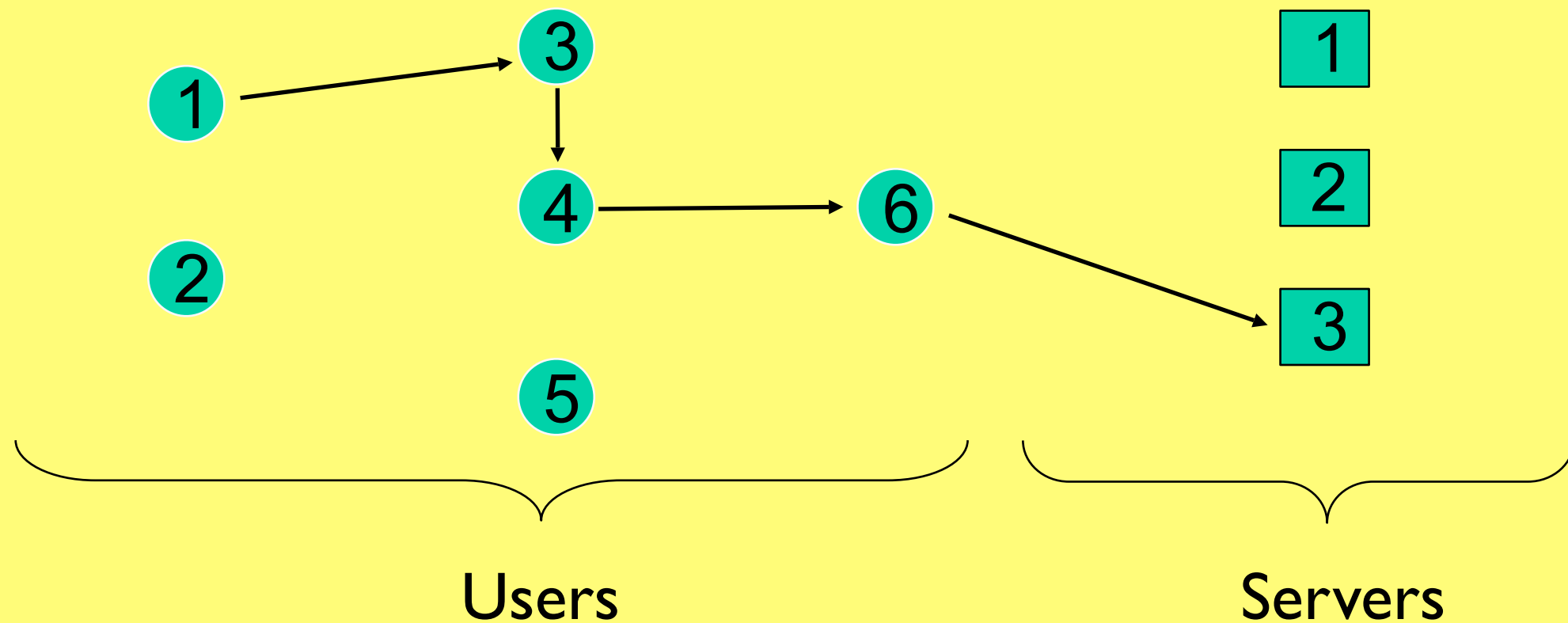
✦ **Crowds** [Reiter and Rubin 1998]: allows internet users to perform anonymous web transactions.

Flips a biased coin  $p_f$



✦ **Crowds** [Reiter and Rubin 1998]: allows internet users to perform anonymous web transactions.

Flips a biased coin  $p_f$



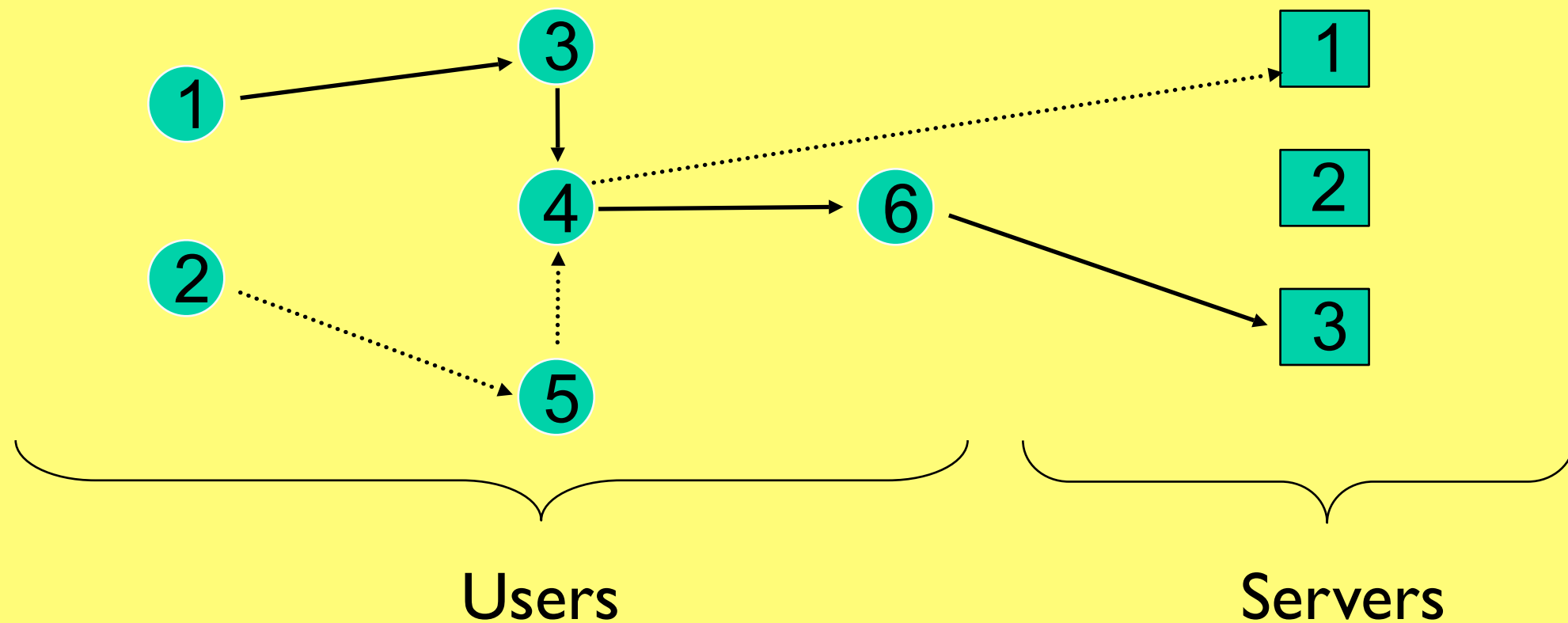


# Crowds

## The protocol

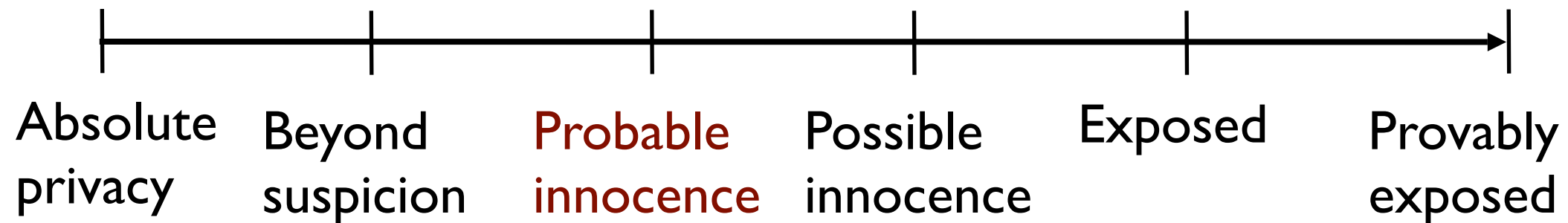
✦ **Crowds** [Reiter and Rubin 1998]: allows internet users to perform anonymous web transactions.

Flips a biased coin  $p_f$



# Probable Innocence

## Informal definition



*“A sender is probably innocent if, from the attacker's point of view, the sender appears no more likely to be the originator than to not be the originator”*

## Formal definition

- ✦ Members:  $m$  members participating in the protocol
  - ✦  $n$  honest members
  - ✦  $c=(m-n)$  corrupt members or collaborating attackers
- ✦ Anonymous events: a random variable  $A$  distributed over  $\{a_1, a_2, \dots, a_n\}$ , where  $a_i$  indicates that the honest user  $i$  is the initiator of the message.
- ✦ Observable events: a random variable  $O$  distributed over  $\{o_1, o_2, \dots, o_n\}$ , where  $o_i$  indicates that user  $i$  is honest and forwards the message to a corrupted user. In this case we say that user  $i$  is detected.



# Probable Innocence

## Formal definition

**Definition** [Reiter and Ruben, 98]:  
a protocol satisfies probable innocence if

$$\forall i \ p(o_i \mid a_i) \leq 1/2$$

# Probable Innocence

## Formal definition

**Definition** [Reiter and Ruben, 98]:  
a protocol satisfies probable innocence if

$$\forall i \ p(o_i \mid a_i) \leq 1/2$$

**Definition** [Halpern and O'Neill, 05]:

$$\forall i \ p(a_i \mid o_i) \leq 1/2$$

# Probable Innocence

## Formal definition

**Definition** [Reiter and Ruben, 98]:  
a protocol satisfies probable innocence if

$$\forall i \ p(o_i \mid a_i) \leq 1/2$$

Wrong



**Definition** [Halpern and O'Neill, 05]:

$$\forall i \ p(a_i \mid o_i) \leq 1/2$$

Right





# Probable Innocence

## Formal definition

**Proposition:** if the a priori distribution is uniform then

$$\forall i \ p(o_i | a_i) = p(a_i | o_i)$$

**Proof:** by Bayes theorem we have

$$p(o_j | a_i)p(a_i) = p(a_i | o_j)p(o_j)$$

If **A** is uniformly distributed then (in Crowds) **O** is uniformly distributed too. Hence  $p(a_i) = p(o_j) = 1/n$

# Probable Innocence

## Extended

### Definition:

a protocol satisfies  $\alpha$ -probable innocence ( $0 \leq \alpha \leq 1$ ) if

$$\forall i \ p(a_i \mid o_i) \leq \alpha$$

### Proposition:

a protocol satisfies  $\alpha$ -probable innocence if and only if

$$1 + n(1-\alpha)/p_f \leq m$$

# Overview

## Trust in Crowds

- ✦ Extend the Crowds protocol to a more realistic scenario:
  - ✦ Associate to each principal  $i$  a probability  $1 - t_i \in [0, 1]$  to become corrupt.
  - ✦ The forwarding process is governed by a policy  $q_i \in [0, 1]$  which together with the forwarding factor  $p_f$  determines the probability that each member  $i$  is chosen as a forwarder.
  
- ✦ Results:
  - ✦ Analyse the impact of such probabilistic behaviour of principals.
  - ✦ Establish necessary and sufficient criteria for choosing an appropriate forwarding policy to achieve required privacy level.

# Overview

## Trust in Crowds

- ✦ Extend the Crowds protocol to a more realistic scenario:
  - ✦ Associate to each principal  $i$  a probability  $1 - t_i \in [0, 1]$  to become corrupt.
  - ✦ The forwarding process is governed by a policy  $q_i \in [0, 1]$  which together with the forwarding factor  $p_f$  determines the probability that each member  $i$  is chosen as a forwarder.

observe this is at meta-level, a parameter of the analysis

- ✦ Results:
  - ✦ Analyse the impact of such probabilistic behaviour of principals.
  - ✦ Establish necessary and sufficient criteria for choosing an appropriate forwarding policy to achieve required privacy level.



# Overview

## Trust in Crowds

- ✦ Extend the Crowds protocol to a more realistic scenario:
  - ✦ Associate to each principal  $i$  a probability  $1 - t_i \in [0, 1]$  to become corrupt.
  - ✦ The forwarding process is governed by a policy  $q_i \in [0, 1]$  which together with the forwarding factor  $p_f$  determines the probability that each member  $i$  is chosen as a forwarder.

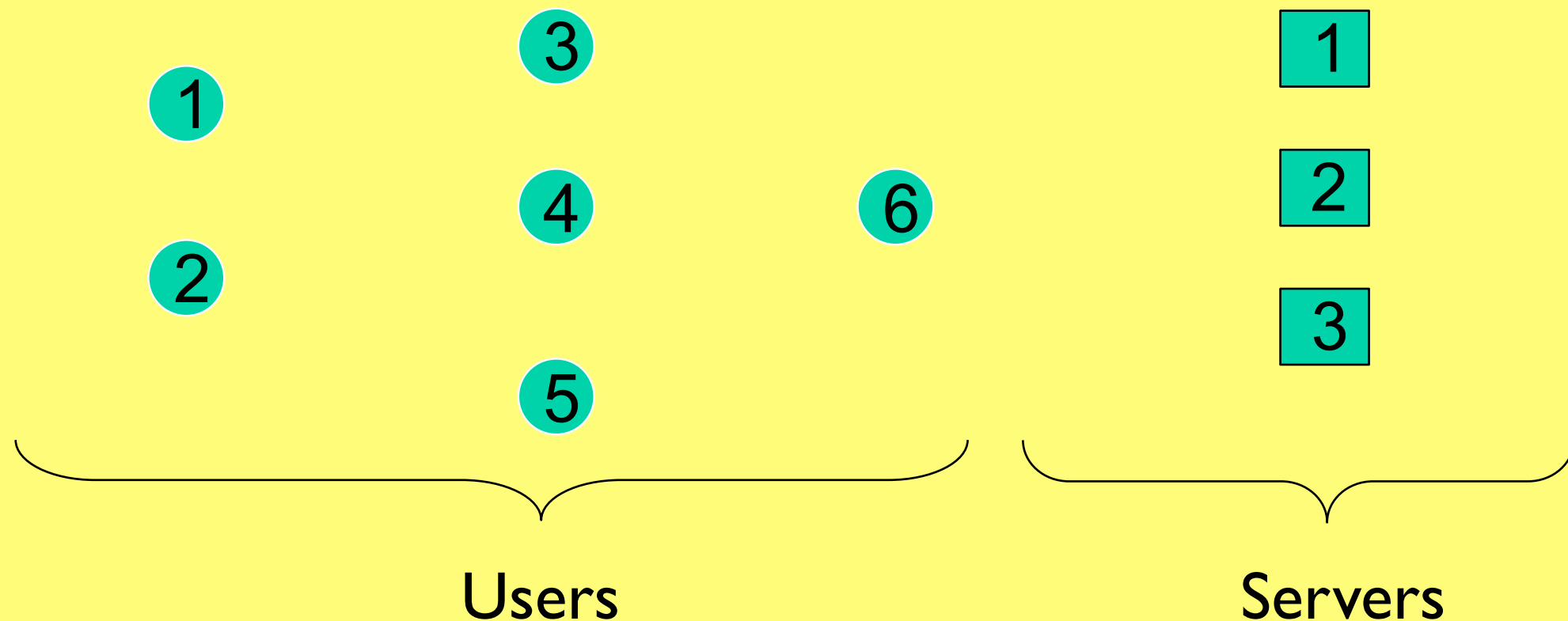
observe this is at meta-level, a parameter of the analysis

Can be established experimentally, eg by the “*blender*” using Bayesian method, eg the Beta trust model

- ✦ Results:
  - ✦ Analyse the impact of such probabilistic behaviour of principals.
  - ✦ Establish necessary and sufficient criteria for choosing an appropriate forwarding policy to achieve required privacy level.

## The extended protocol

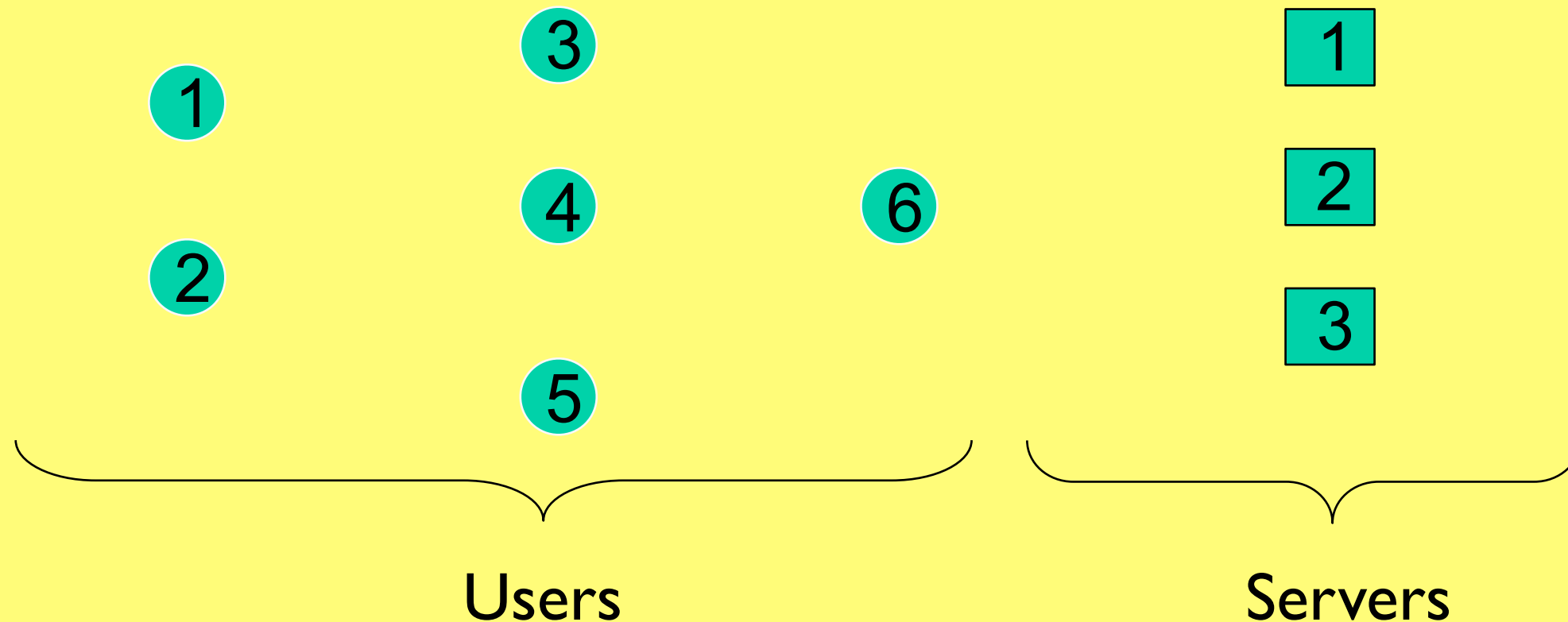
✦ **tCrowds** [here and now]: allows users anonymous web transactions in the presence of probabilistic principals' behaviours.



## The extended protocol

✦ **tCrowds** [here and now]: allows users anonymous web transactions in the presence of probabilistic principals' behaviours.

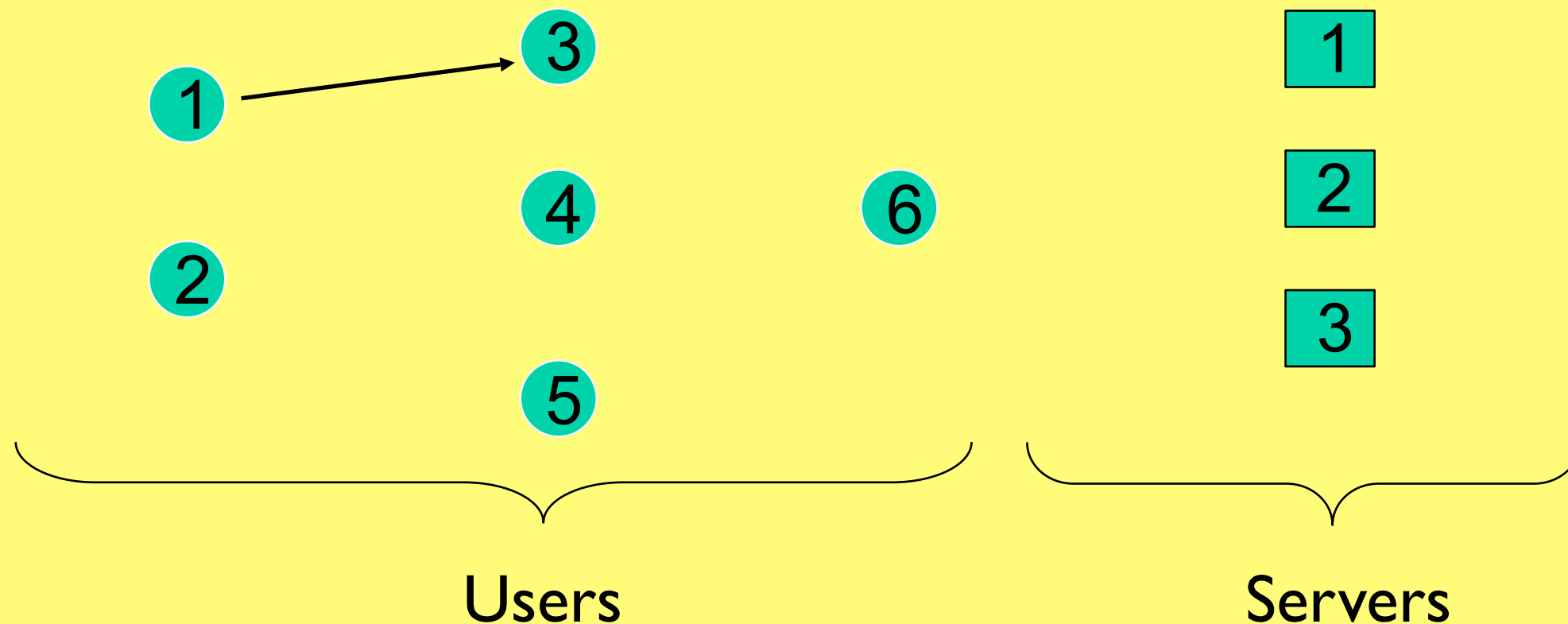
Initiator selects  
 $j$  with prob  $q_j$



## The extended protocol

✦ **tCrowds** [here and now]: allows users anonymous web transactions in the presence of probabilistic principals' behaviours.

Initiator selects  
 $j$  with prob  $q_j$

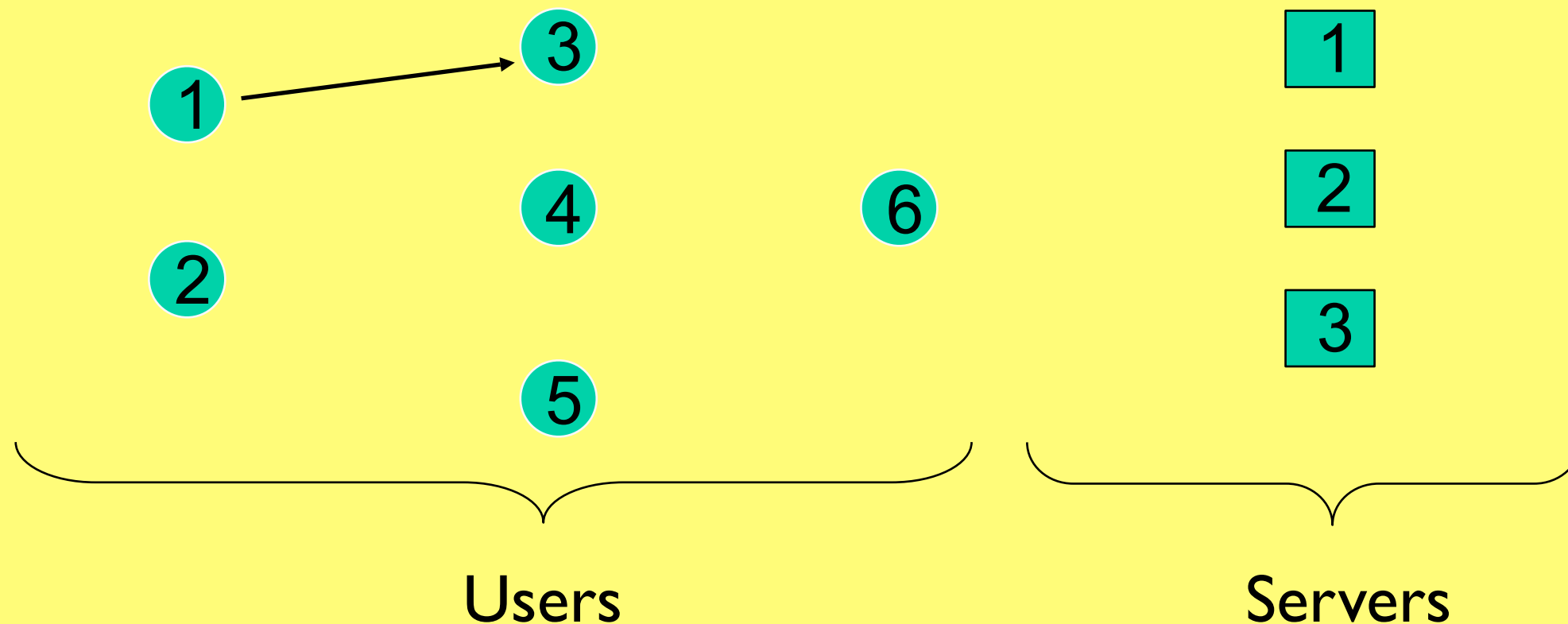


## The extended protocol

✦ **tCrowds** [here and now]: allows users anonymous web transactions in the presence of probabilistic principals' behaviours.

Initiator selects  
 $j$  with prob  $q_j$

Delivers to server with prob  $1 - p_f$   
Forwards to  $j$  with prob  $p_f \cdot q_j$



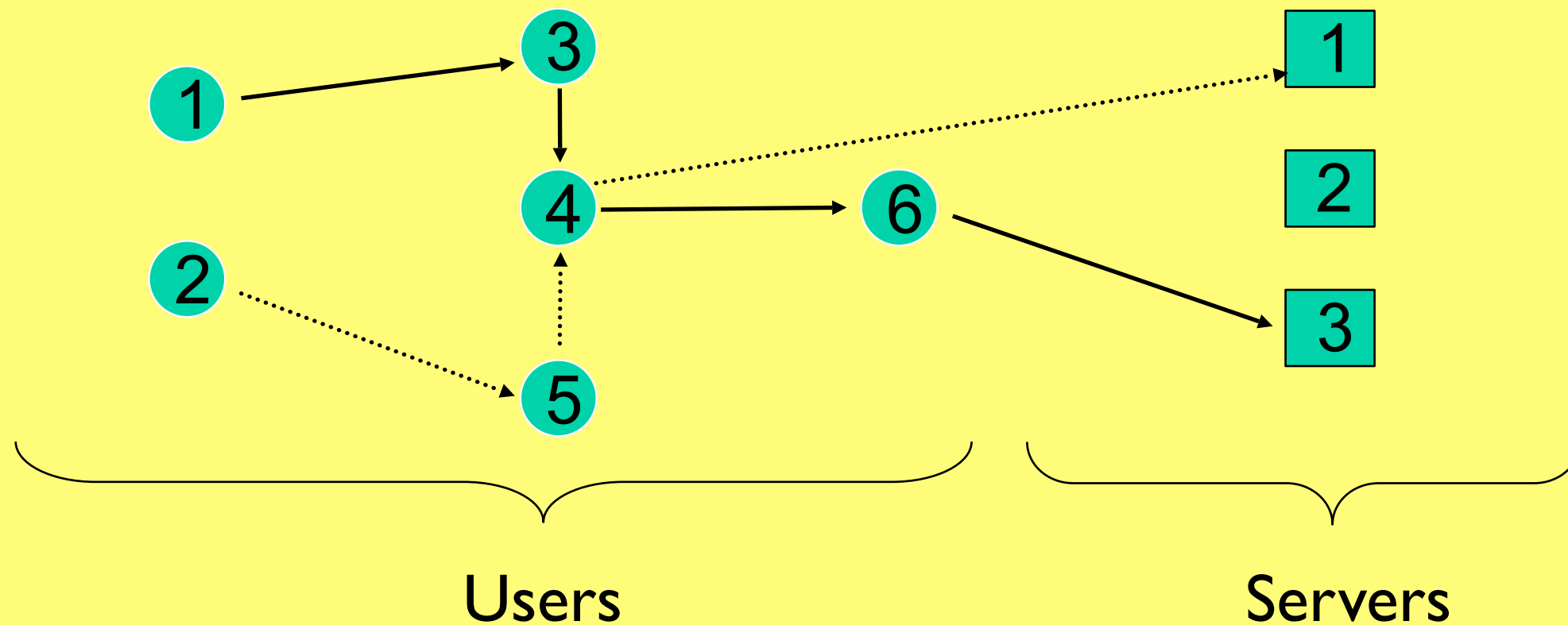


## The extended protocol

✧ **tCrowds** [here and now]: allows users anonymous web transactions in the presence of probabilistic principals' behaviours.

Initiator selects  
 $j$  with prob  $q_j$

Delivers to server with prob  $1 - p_f$   
Forwards to  $j$  with prob  $p_f \cdot q_j$



## The external observer

observe we assume transactions are short, otherwise users could become corrupt whilst answer from server travels back.

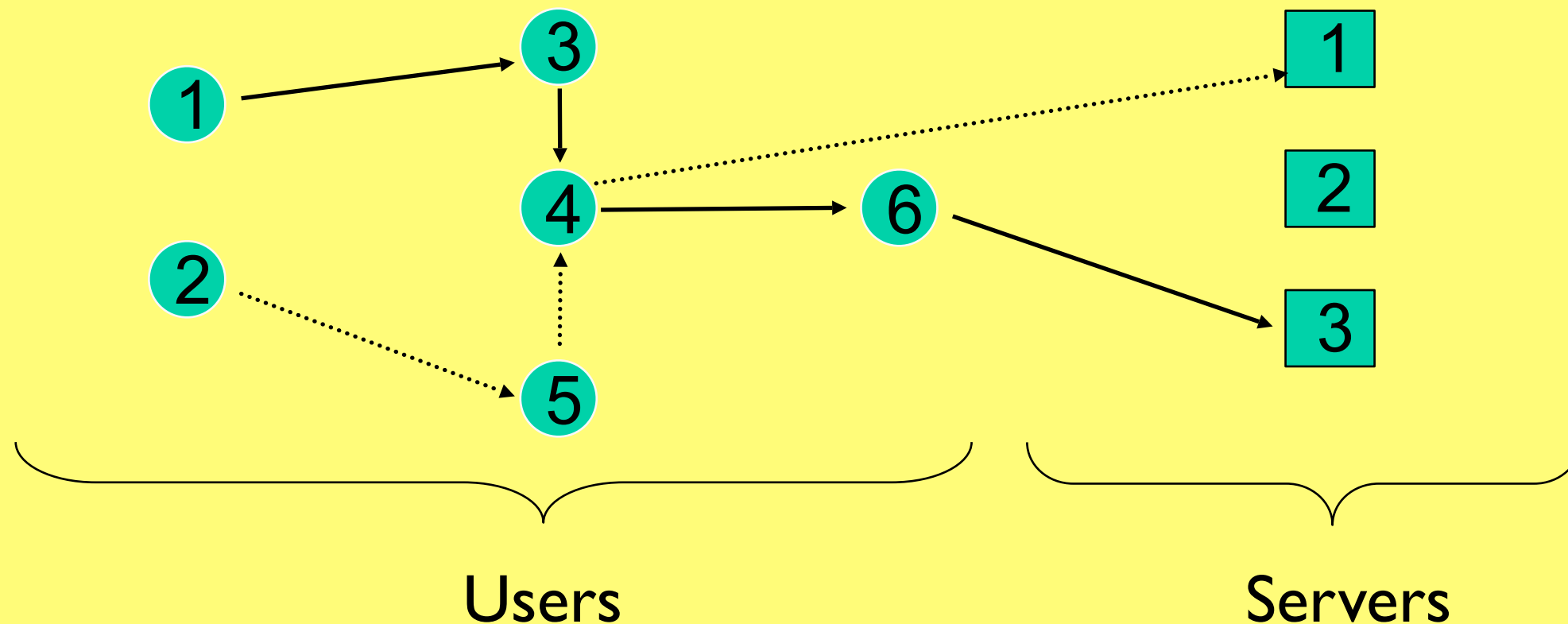
### ✦ tCrowds

web

transactions in the presence of probabilistic principals' behaviours.

Initiator selects  
 $j$  with prob  $q_j$

Delivers to server with prob  $1 - p_f$   
Forwards to  $j$  with prob  $p_f \cdot q_j$



# tCrowds

The external

observe we assume transactions are short,  
otherwise users could become corrupt  
whilst answer from server travels back.

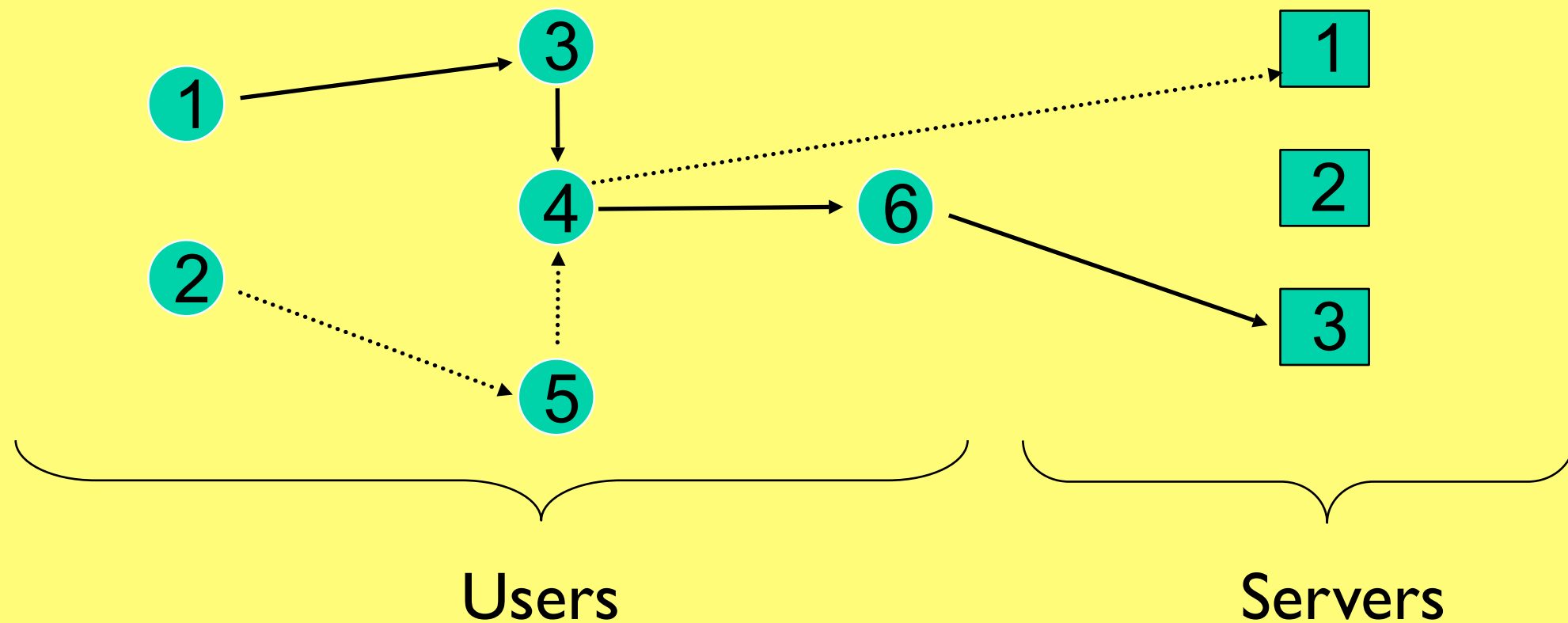
✦ tCrowds

transactions in the presence of malicious behaviours.

Initiator selects  
 $j$  with prob  $q_j$

Delivery  
Forwards to  $j$  with prob  $p_f \cdot q_j$

extension to the general  
case is work in progress



# Probable Innocence, again

→ Need to compute

$$P(a_i | o_i) = \frac{P(a_i, o_i)}{P(o_i)}$$

→ Start with:

# Probable Innocence, again

→ Need to compute

$$P(a_i | o_i) = \frac{P(a_i, o_i)}{P(o_i)}$$

→ Start with:

$$P(o_i, H_k) = \begin{cases} \frac{1}{n}(1 - t_i) & k = 0 \\ \frac{1}{n}t_i(1 - T) & k = 1 \\ \frac{1}{n}S T^{k-2} q_i t_i (1 - T) \cdot p_f^{k-1} & k \geq 2 \end{cases}$$

$$\text{with } S = \sum_{j=1}^n t_j \quad T = \sum_{j=1}^n q_j t_j$$



# Probable Innocence, again

→ Need to compute

$$P(a_i | o_i) = \frac{P(a_i, o_i)}{P(o_i)}$$

→ Start with:

$$P(o_i, H_k) = \begin{cases} \frac{1}{n}(1 - t_i) & k = 0 \\ \frac{1}{n}t_i(1 - T) & k = 1 \\ \frac{1}{n}S T^{k-2} q_i t_i (1 - T) \cdot p_f^{k-1} & k \geq 2 \end{cases}$$

1st attacker  
at position **k**

$$\text{with } S = \sum_{j=1}^n t_j \quad T = \sum_{j=1}^n q_j t_j$$

# Probable Innocence, again

→ Need to compute

$$P(a_i | o_i) = \frac{P(a_i, o_i)}{P(o_i)}$$

→ Start with:

$$P(o_i, H_k) = \begin{cases} \frac{1}{n}(1 - t_i) & k = 0 \\ \frac{1}{n}t_i(1 - T) & k = 1 \\ \frac{1}{n}S T^{k-2} q_i t_i (1 - T) \cdot p_f^{k-1} & k \geq 2 \end{cases}$$

1st attacker  
at position **k**

$$\text{with } S = \sum_{j=1}^n t_j$$

$$T = \sum_{j=1}^n q_j t_j$$

prob to pick a  
honest principal

# Probable Innocence, again

→ Need to compute

$$P(a_i | o_i) = \frac{P(a_i, o_i)}{P(o_i)}$$

→ Continue with:

$$\begin{aligned} P(o_i) &= \sum_{k=0}^{\infty} P(o_i, H_k) \\ &= \frac{1}{n}(1 - t_i) + \frac{1}{n}t_i(1 - T) \\ &\quad + \sum_{k=2}^{\infty} \frac{1}{n} S T^{k-2} \cdot q_i t_i (1 - T) p_f^{k-1} \\ &= \frac{1}{n} \left( 1 - t_i T + S p_f q_i t_i \left( \frac{1 - T}{1 - p_f T} \right) \right) \end{aligned}$$

# Probable Innocence, again

→ Need to compute

$$P(a_i | o_i) = \frac{P(a_i, o_i)}{P(o_i)}$$

→ Continue with:

$$\begin{aligned} P(o_i) &= \sum_{k=0}^{\infty} P(o_i, H_k) \\ &= \frac{1}{n}(1 - t_i) + \frac{1}{n}t_i(1 - T) \\ &\quad + \sum_{k=2}^{\infty} \frac{1}{n} S T^{k-2} \cdot q_i t_i (1 - T) p_f^{k-1} \\ &= \frac{1}{n} \left( 1 - t_i T + S p_f q_i t_i \left( \frac{1 - T}{1 - p_f T} \right) \right) \end{aligned}$$

observe this is 0  
iff  $T=1$  and  $t_i=1$   
 $i$  is undetectable

# Probable Innocence, again

→ Need to compute

$$P(a_i | o_i) = \frac{P(a_i, o_i)}{P(o_i)}$$

→ Similarly:

$$\begin{aligned} P(a_i, o_i) &= \sum_{k=0}^{\infty} P(a_i, H_k, o_i) \\ &= \frac{1}{n}(1 - t_i) + \frac{1}{n}t_i(1 - T) \\ &\quad + \sum_{k=2}^{\infty} \frac{1}{n}t_i T^{k-2} \cdot q_i t_i (1 - T) p_f^{k-1} \\ &= \frac{1}{n} \left( 1 - t_i T + p_f q_i t_i^2 \left( \frac{1 - T}{1 - p_f T} \right) \right) \end{aligned}$$



# Probable Innocence, again

→ Need to compute

$$P(a_i | o_i) = \frac{P(a_i, o_i)}{P(o_i)}$$

→ And therefore:

$$P(a_i | o_i) = \frac{1 - t_i T + p_f q_i t_i^2 \left( \frac{1-T}{1-p_f T} \right)}{1 - t_i T + S p_f q_i t_i \left( \frac{1-T}{1-p_f T} \right)}$$

→ Observe that if  $i$  is detectable, this quantity is positive: ie, it can always be caught when is the initiator: Crowds never achieves “*absolute privacy*”

# Probable Innocence, again

→ Need to compute

$$P(a_i | o_i) = \frac{P(a_i, o_i)}{P(o_i)}$$

also observe that when  $T = 1 - c/n$  and  $S = n - c$ , which characterise the (standard) Crowds, then this formula simplifies to the standard one.

→ And therefore

$$P(a_i | o_i) = \frac{1 - t_i T + p_f q_i t_i^2 \left( \frac{1-T}{1-p_f T} \right)}{1 - t_i T + S p_f q_i t_i \left( \frac{1-T}{1-p_f T} \right)}$$

→ Observe that if  $i$  is detectable, this quantity is positive: ie, it can always be caught when is the initiator: Crowds never achieves “*absolute privacy*”

## **Proposition:** *(Provably Exposed Principals)*

For all users s.t.  $p(o_i) \neq 0$ , we have  $p(a_i | o_i) = 1$  iff one of the following holds.

1.  $p_f = 0$

2.  $t_i = 0$

3.  $q_i = 0$

4.  $T = 1$

5.  $S = t_i$

## **Proposition:** (*Provably Exposed Principals*)

For all users s.t.  $p(o_i) \neq 0$ , we have  $p(a_i | o_i) = 1$  iff one of the following holds.

1.  $p_f = 0$

2.  $t_i = 0$

3.  $q_i = 0$

4.  $T = 1$

5.  $S = t_i$

all paths  $\# \leq 2$

## **Proposition:** (*Provably Exposed Principals*)

For all users s.t.  $p(o_i) \neq 0$ , we have  $p(a_i | o_i) = 1$  iff one of the following holds.

1.  $p_f = 0$

2.  $t_i = 0$

3.  $q_i = 0$

4.  $T = 1$

5.  $S = t_i$

all paths  $\# \leq 2$

$i$  is corrupt!

## **Proposition:** (*Provably Exposed Principals*)

For all users s.t.  $p(o_i) \neq 0$ , we have  $p(a_i | o_i) = 1$  iff one of the following holds.

1.  $p_f = 0$

2.  $t_i = 0$

3.  $q_i = 0$

4.  $T = 1$

5.  $S = t_i$

all paths  $\# \leq 2$

$i$  is corrupt!

$i$  never picked  
as forwarder



## **Proposition:** (*Provably Exposed Principals*)

For all users s.t.  $p(o_i) \neq 0$ , we have  $p(a_i | o_i) = 1$  iff one of the following holds.

1.  $p_f = 0$

2.  $t_i = 0$

3.  $q_i = 0$

4.  $T = 1$

5.  $S = t_i$

all paths  $\# \leq 2$

$i$  is corrupt!

$i$  never picked  
as forwarder

all participants  
are honest!

## **Proposition:** (*Provably Exposed Principals*)

For all users s.t.  $p(o_i) \neq 0$ , we have  $p(a_i | o_i) = 1$  iff one of the following holds.

1.  $p_f = 0$

2.  $t_i = 0$

3.  $q_i = 0$

4.  $T = 1$

5.  $S = t_i$

all paths  $\# \leq 2$

$i$  is corrupt!

$i$  never picked  
as forwarder

all participants  
are honest!

all but  $i$  are  
corrupt!

**Theorem:** *(Monotonicity in forwarding)*

$p(a_i | o_i)$  is a decreasing function of  $p_f$

**Corollary:** *(Anonymity range)*

$$\forall i. P(a_i | o_i) \geq 1 - \frac{q_i t_i \sum_{j \neq i}^n t_j}{1 - t_i \sum_{j \neq i}^n q_j t_j + q_i t_i \sum_{j \neq i}^n t_j}$$

**Theorem:** (*Monotonicity in forwarding*)

$p(a_i | o_i)$  is a decreasing function of  $p_f$

tells us that high values of  $p_f$  enhance privacy. Yet, they slow the protocol down

**Corollary:** (*Anonymity range*)

$$\forall i. P(a_i | o_i) \geq 1 - \frac{q_i t_i \sum_{j \neq i}^n t_j}{1 - t_i \sum_{j \neq i}^n q_j t_j + q_i t_i \sum_{j \neq i}^n t_j}$$

**Theorem:** (*Monotonicity in forwarding*)

$p(a_i | o_i)$  is a decreasing function of  $p_f$

tells us that high values of  $p_f$  enhance privacy. Yet, they slow the protocol down

**Corollary:** (*Anonymity range*)

$$\forall i. P(a_i | o_i) \geq 1 - \frac{q_i t_i \sum_{j \neq i}^n t_j}{1 - t_i \sum_{j \neq i}^n q_j t_j + q_i t_i \sum_{j \neq i}^n t_j}$$

tells us that  $p_f = 1$  minimises  $p(a_i | o_i)$ . But then the message never reaches...

**Theorem:** ( *$\alpha$ -Probable Innocence*)

For all  $\alpha \in [0, 1]$ , the extended protocol guarantees  $\alpha$ -probable innocence to all its participants if

$$\forall i. \frac{q_i t_i \sum_{j \neq i}^n t_j}{1 - t_i \sum_{j \neq i}^n q_j t_j + q_i t_i \sum_{j \neq i}^n t_j} \geq 1 - \alpha$$



## **Theorem:** ( $\alpha$ -Probable Innocence)

For all  $\alpha \in [0, 1]$ , the extended protocol guarantees  $\alpha$ -probable innocence to all its participants if

$$\forall i. \frac{q_i t_i \sum_{j \neq i}^n t_j}{1 - t_i \sum_{j \neq i}^n q_j t_j + q_i t_i \sum_{j \neq i}^n t_j} \geq 1 - \alpha$$

observe that this provides a *system of linear inequalities* that can be solved in  $\mathbf{q}_i$  to try and achieve  $\alpha$ -probable innocence

## Achieving $\alpha$ -Probable Innocence

Maintain the lower bound on  $p(a_i | o_i) = I$  below  $\alpha$  by manipulating the forwarding distribution (*social policy*), or by excluding untrustworthy participants (*rational policy*).

**Example:** Suppose  $t_1 = 0.70$ ,  $t_2 = 0.97$ ,  $t_3 = 0.99$

For  $\alpha = 1/2$  the system admits two solutions, eg

$$q_1 = 0.4575, \quad q_2 = 0.2620, \quad q_3 = 0.2805 .$$

Observe how user **1** is helped (at the others' risk!) to offset its higher tendency to corruption. Indeed, probable innocence in (standard) Crowds cannot be achieved.

The alternative, is for **2** and **3** to exclude **1** and yield higher overall security.

- We have extended *Crowds* to take into account that principals are not usually either honest or malicious, but are liable to become *corrupt* (and again uncorrupt). Ours is the first attempt to cope with such probabilistic behaviour.
- Our forwarding policies can be used to make the protocol more secure (either *socially* or *rationally*) once an estimation of trust is available. A lot more work on integrating trust estimation is to be done.
- A deeper analysis of trust is likely to be possible on advanced anonymity protocols such as *Tarzan* and *ToR*.
- We are in the process of complete this analysis by *dropping* the hypothesis of short transactions.

# Related Work

## Crowds & External knowledge

- ✦ **Real world:** attackers usually gather additional information correlated to the anonymous agents before attacking the protocol.
- ✦ **Example:** two agents voting by “yes” or “no” and the result of the vote is {yes, no}
  - ✦ Agents used different colours but the adversary does not know the correlation between the colors and the agents:  
$$\{\text{yes}, \text{no}\} \equiv \{\text{yes}, \text{no}\}$$
  - ✦ The adversary knows the correlation:  $\{\text{yes}, \text{no}\} \neq \{\text{yes}, \text{no}\}$

# Related Work

## Crowds & External knowledge

- ✦ **Real world:** attackers usually gather additional information correlated to the anonymous agents before attacking the protocol.
- ✦ **Example:** two agents voting by “yes” or “no” and the result of the vote is {yes, no}
  - ✦ Agents used different colours but the adversary does not know the correlation between the colors and the agents:  
 $\{\text{yes}, \text{no}\} \equiv \{\text{yes}, \text{no}\}$
  - ✦ The adversary knows the correlation:  $\{\text{yes}, \text{no}\} \neq \{\text{yes}, \text{no}\}$

analysis of the impact of attackers' extra knowledge on the security of information hiding protocols.

in *FAST 2009*  
with C. Palamidessi

# Related Work

## Crowds & Beliefs & Vulnerability

- ✦ Open problem: measure and account for the **accuracy** of the adversary extra knowledge.
- ✦ Integrate the notion of adversary's beliefs:
  - ✦ Assume both actual a priori distribution of the hidden input and its correlation to the extra information unknown to adversary.
  - ✦ Generalise the approach to information flow systems.
- ✦ Results:
  - ✦ New metric for quantitative information flow based on the concept of vulnerability that takes into account the adversary's beliefs.
  - ✦ Model allows to identify the levels of accuracy for the adversary's beliefs which are compatible with the security of a given program or protocol.

# Related Work

## Crowds & Beliefs & Vulnerability

- ✦ Open problem: measure and account for the **accuracy** of the adversary extra knowledge.
- ✦ Integrate the notion of adversary's beliefs:
  - ✦ Assume both actual a priori distribution of the hidden input and its correlation to the extra information unknown to adversary.
  - ✦ Generalise the approach to information flow systems.
- ✦ Results:
  - ✦ New metric for quantitative information flow based on the concept of vulnerability that takes into account the adversary's beliefs.
  - ✦ Model allows to identify the levels of accuracy for the adversary's beliefs which are compatible with the security of a given program or protocol.

in *IEEE Symp on Security & Privacy 2010*  
with C. Palamidessi