The Fundamental Limits of Data & Metadata Privacy

Peter Kairouz

ECE Department

University of Illinois at Urbana-Champaign

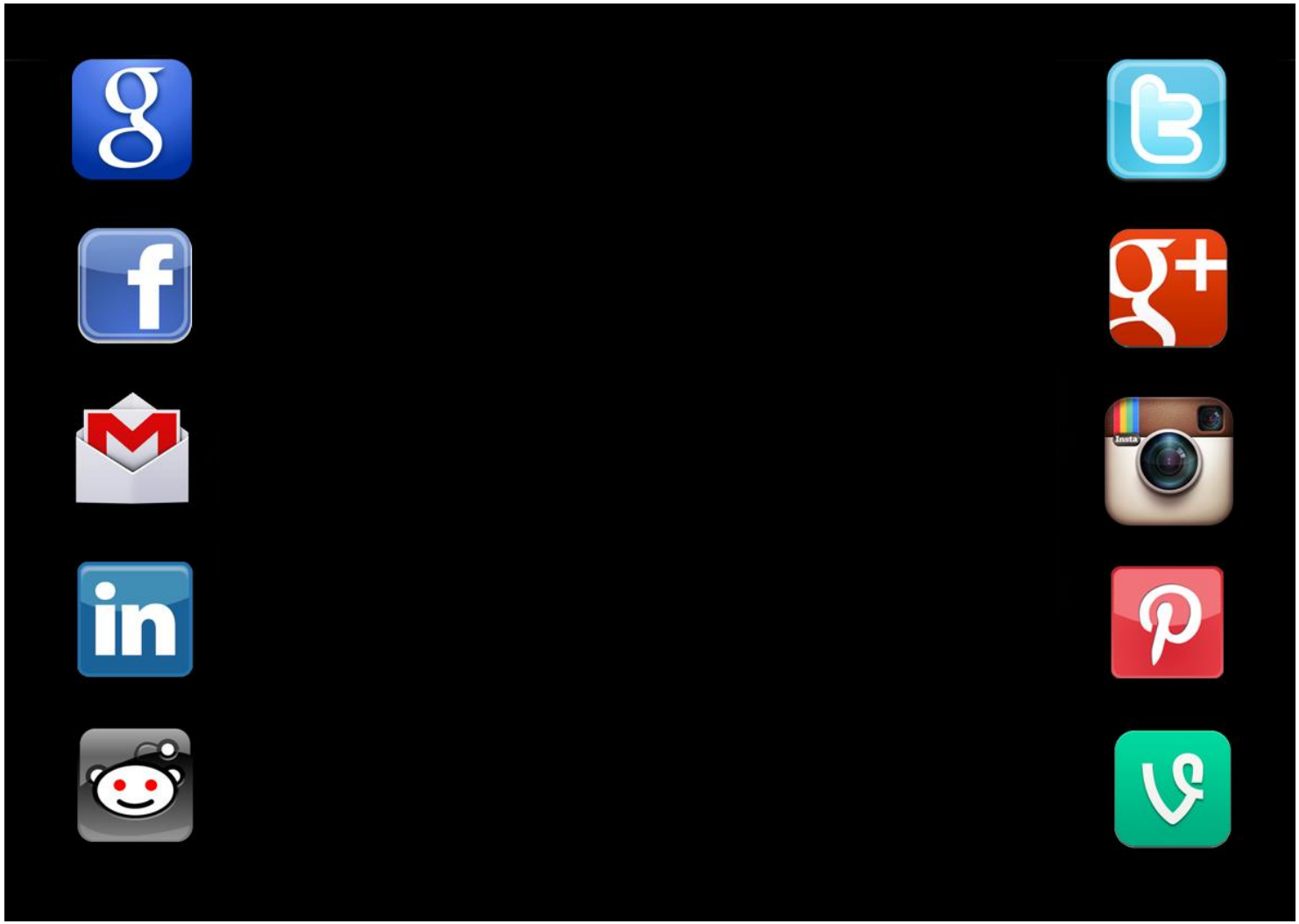
Communication

- transfer of information **from one point** in space-time **to the other**

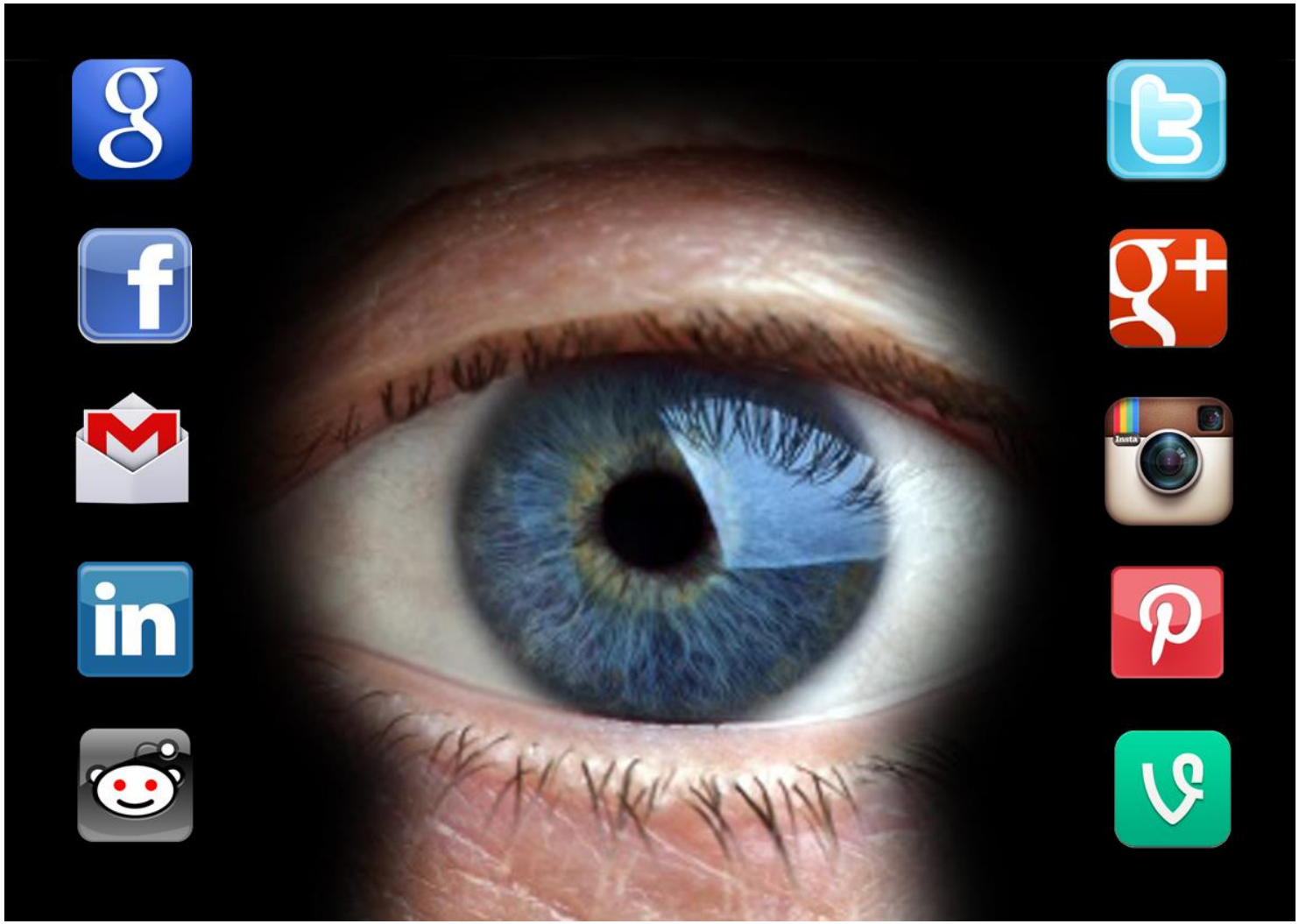
Wireless communication

- the **fundamental limits** of wireless communication are **well understood**

Rise of the planet of the apps!

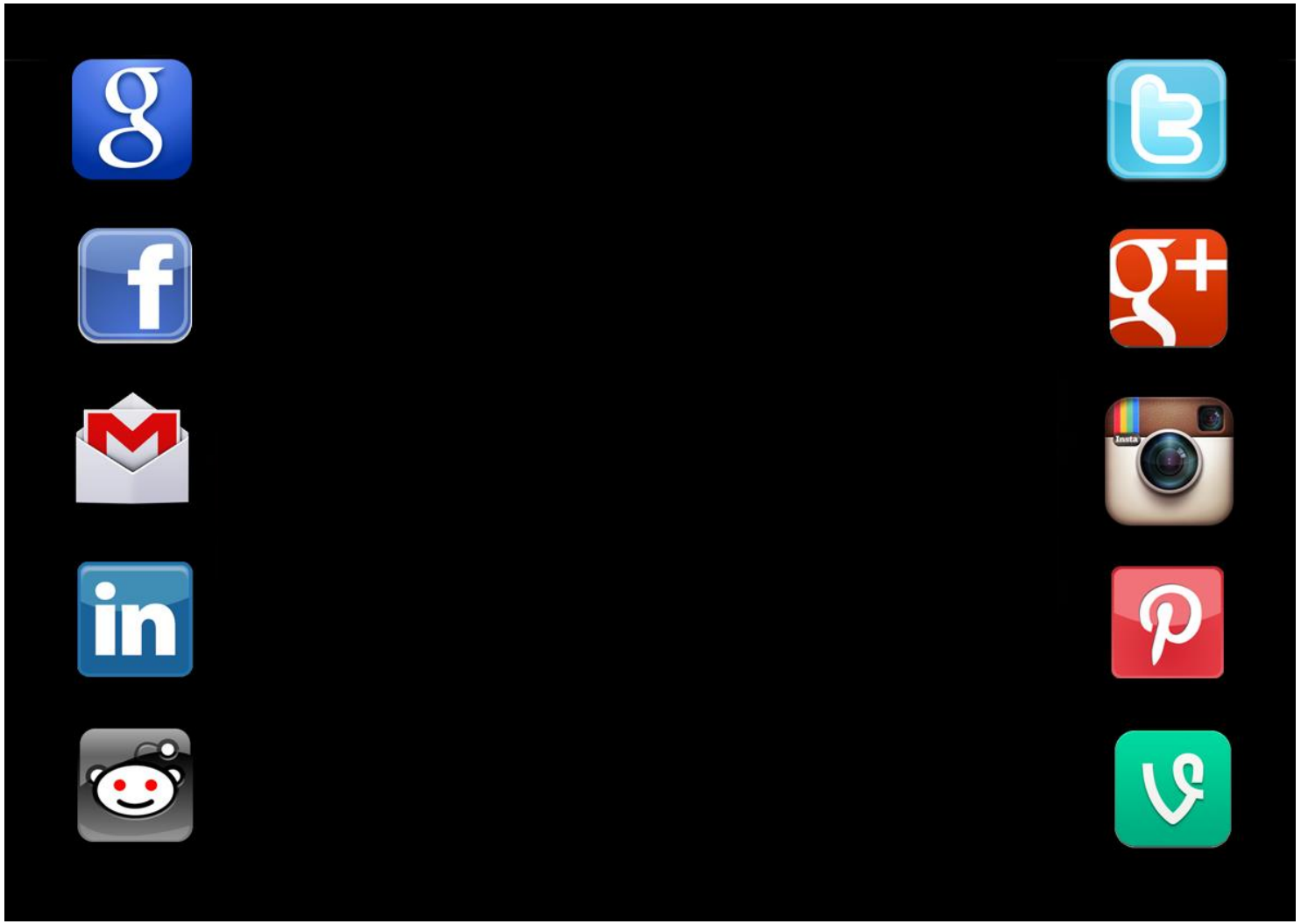


Rise of the planet of the apps!



Rise of the planet of the apps!

Rise of the planet of the apps!



can we communicate anonymously and privately?

Data and metadata privacy

Bob

Alice

Part 1:
Metadata Privacy

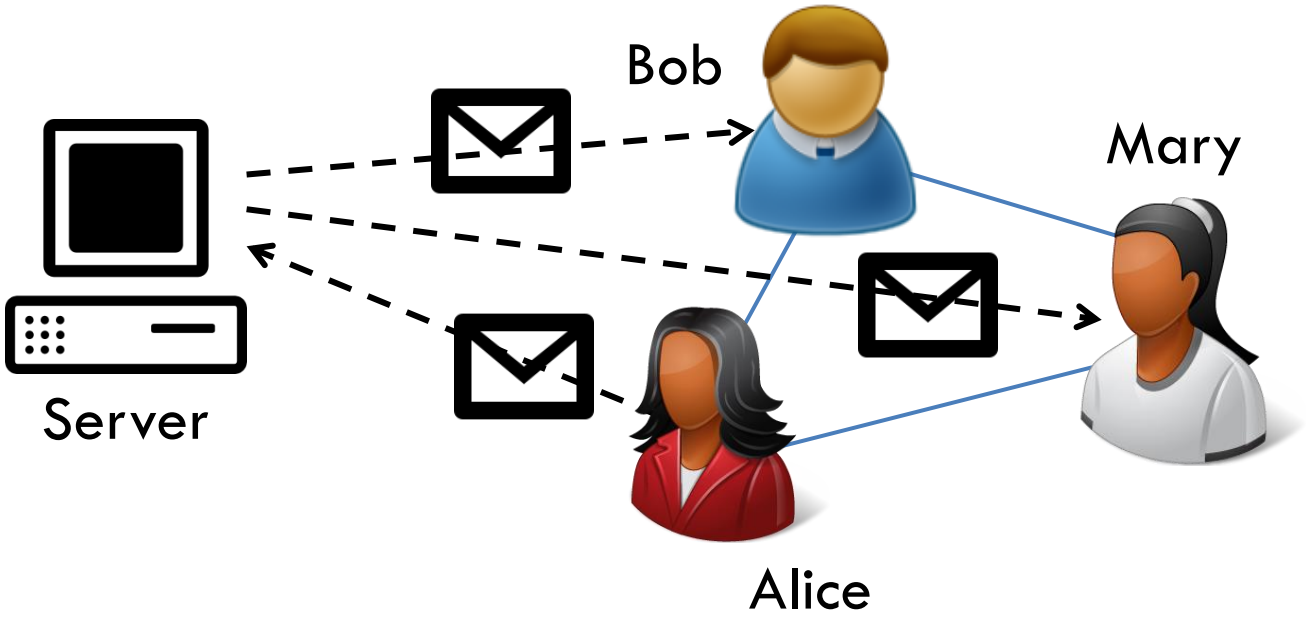
Political activism

some people have important,
sensitive things to say

Existing anonymous messaging apps

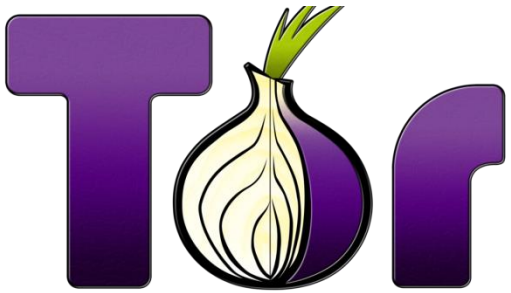
Yik Yak

whisper



centralized networks are not truly anonymous!

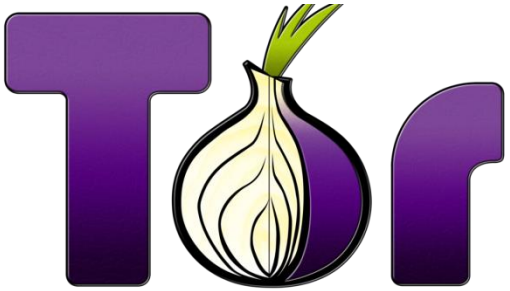
Anonymous communication



OneSwarm

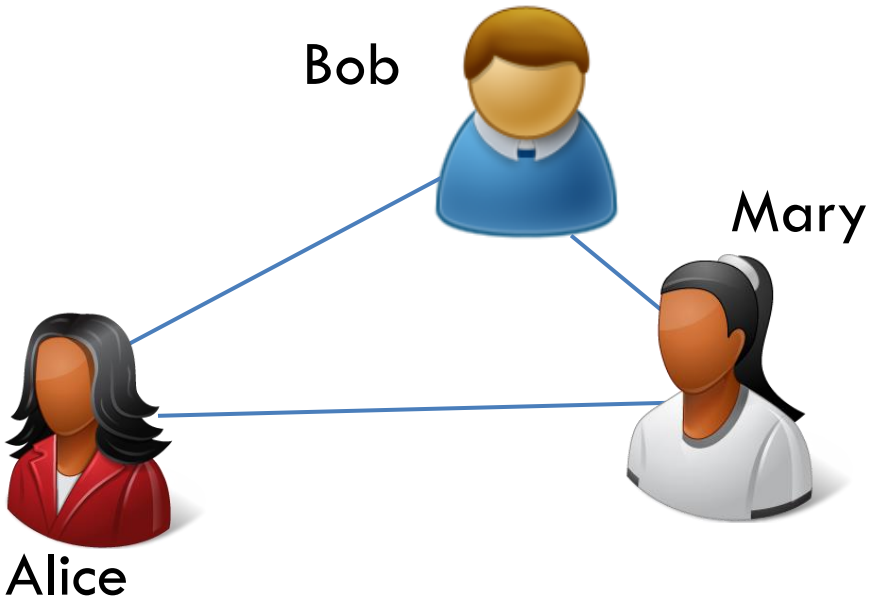
Privacy preserving peer-to-peer data sharing

Anonymous communication

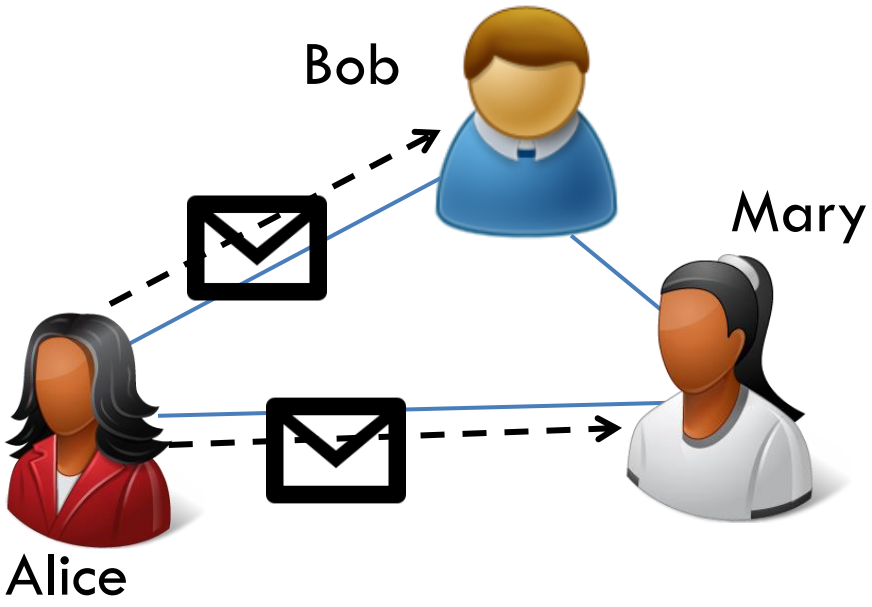


designed for point-to-point communication

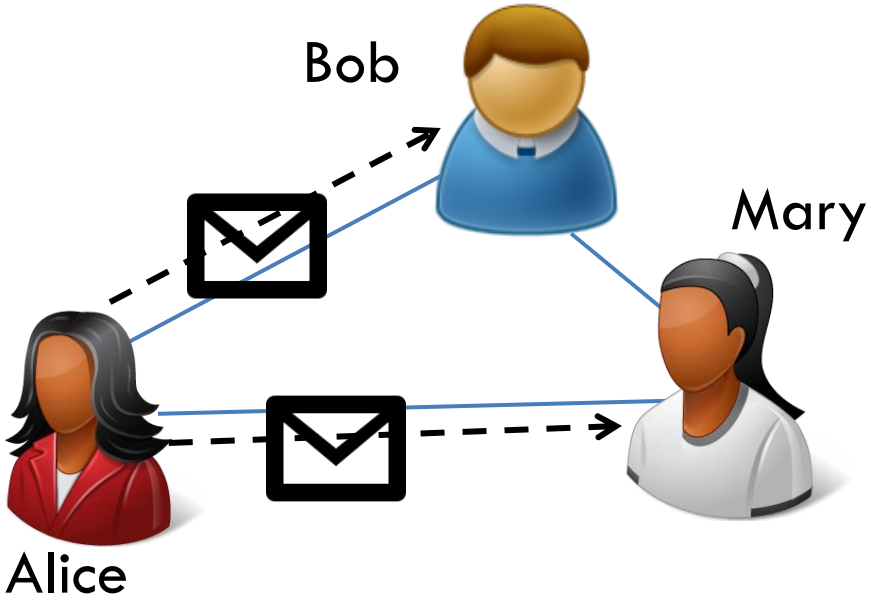
Distributed messaging



Distributed messaging

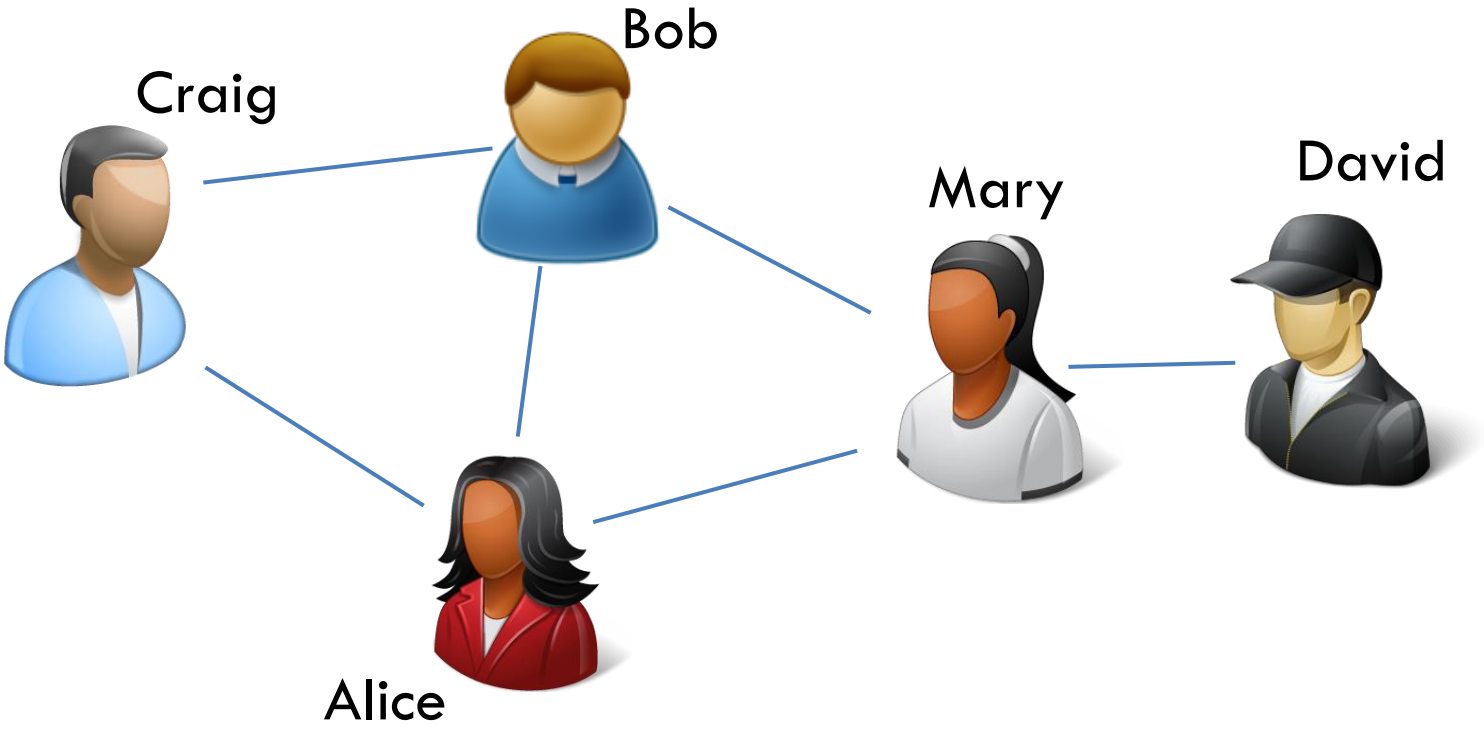


Distributed messaging

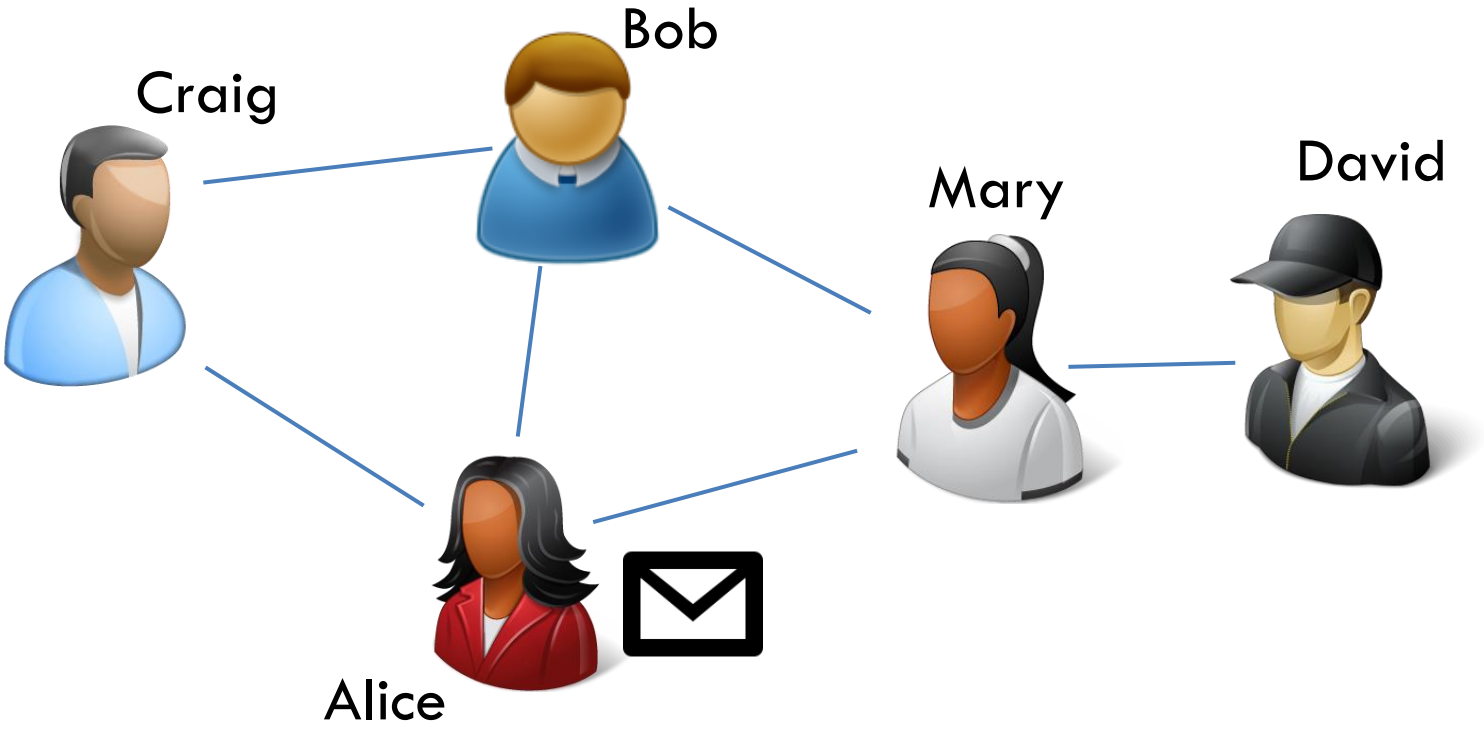


what can an adversary do?

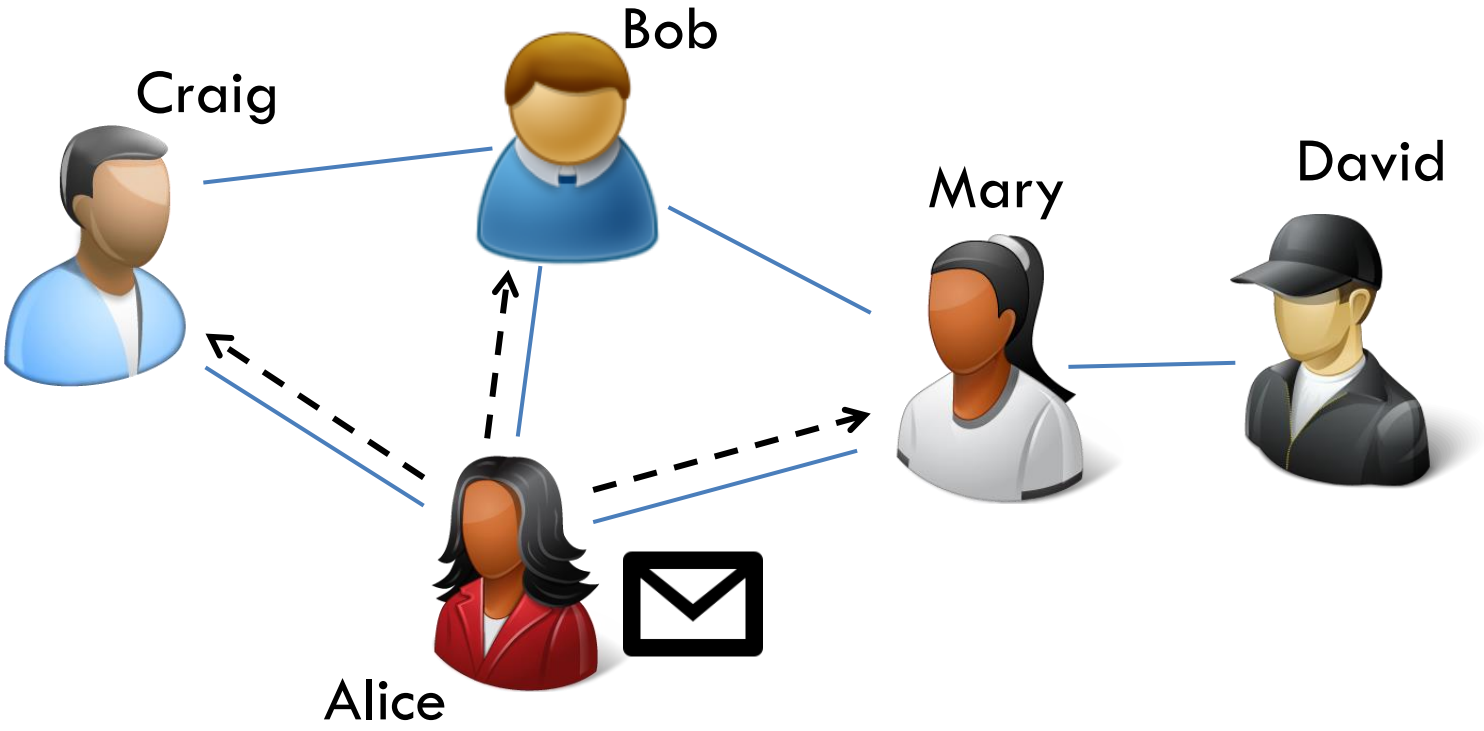
Adversary without timing information



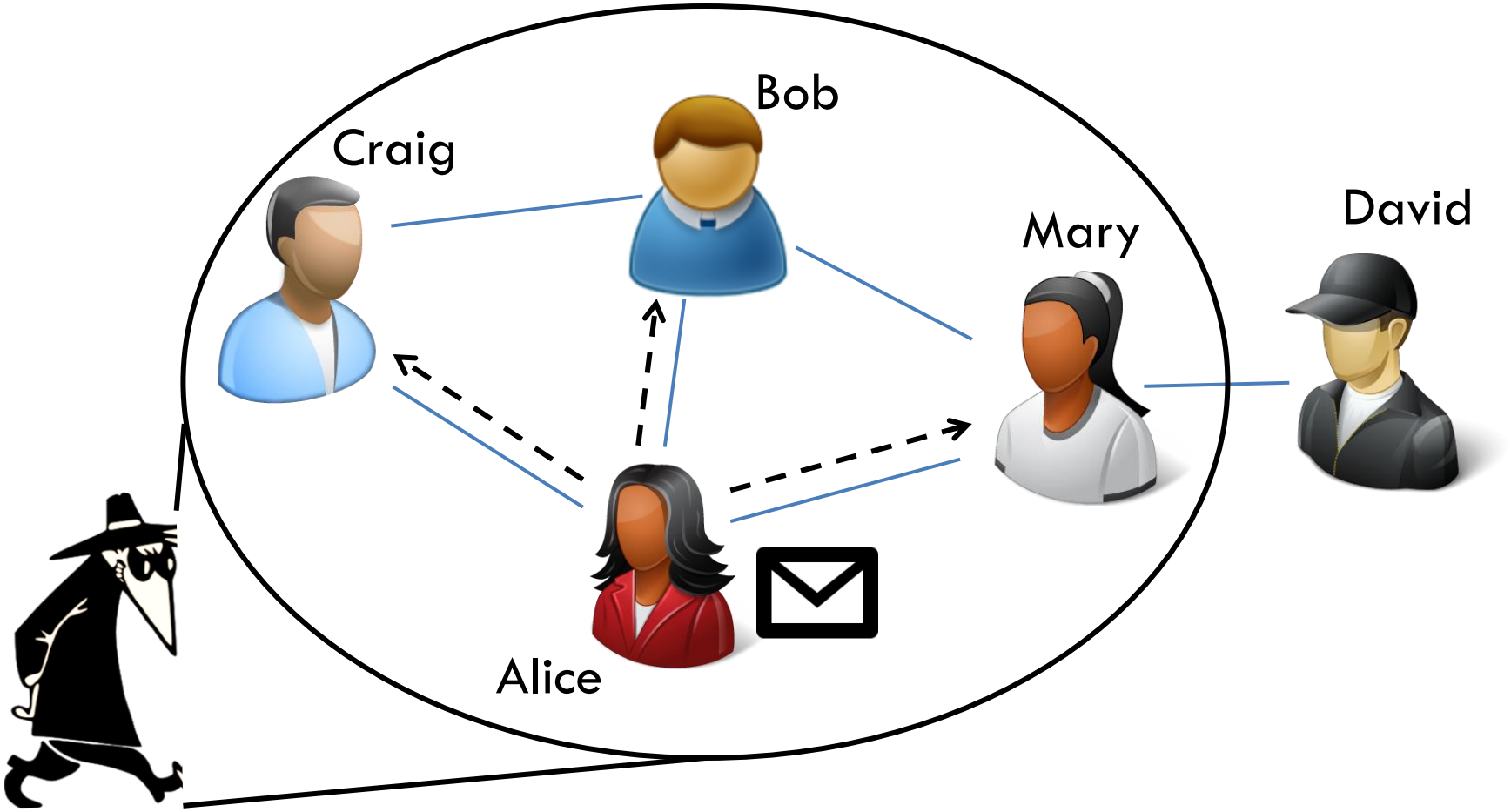
Adversary without timing information



Adversary without timing information

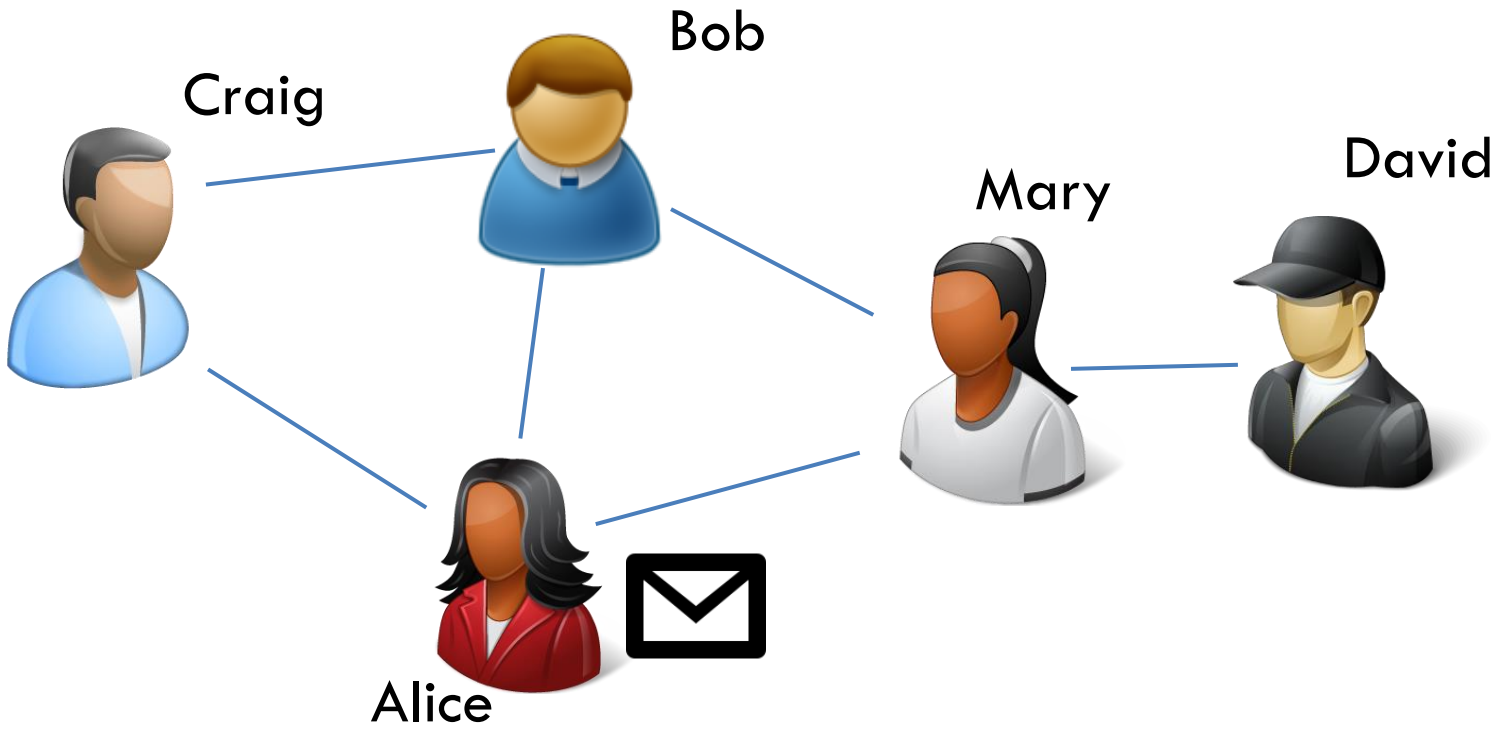


Adversary without timing information

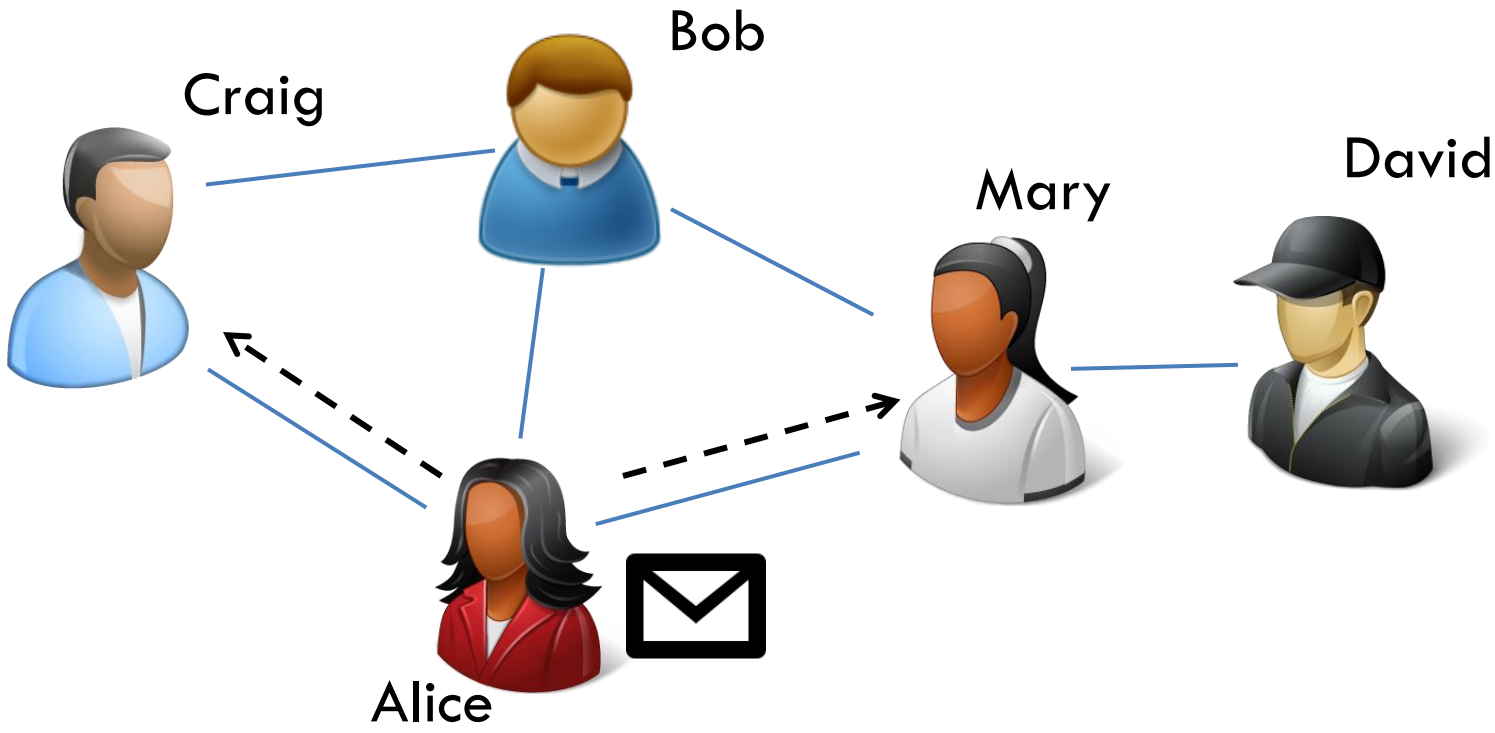


adversary can figure out who got the message

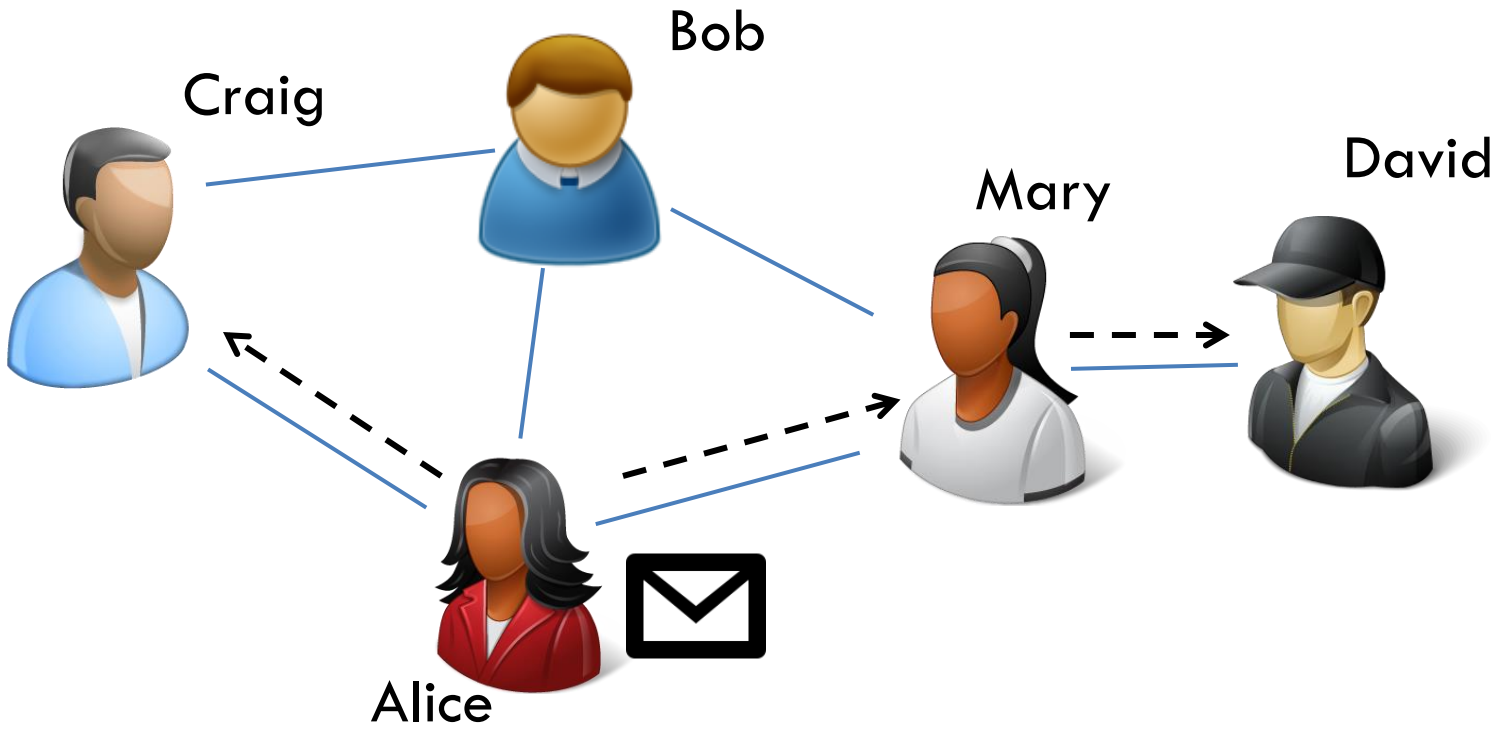
Adversary with timing information



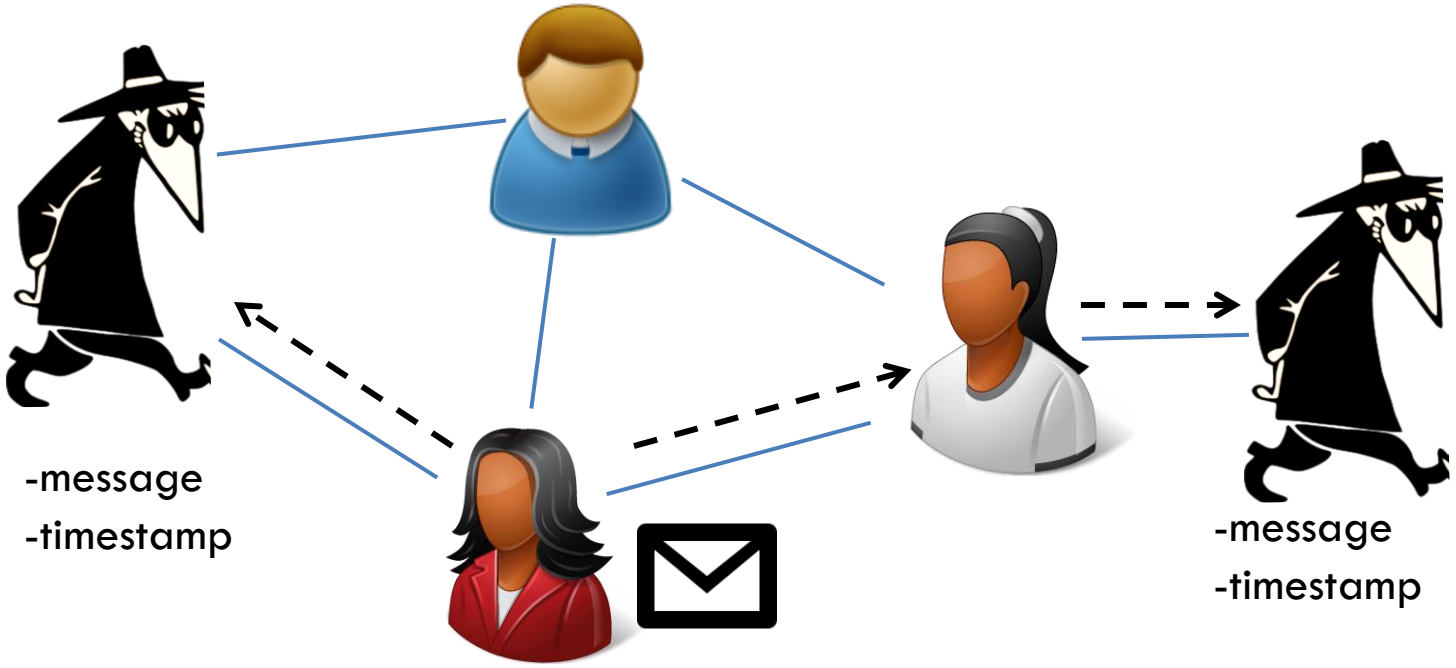
Adversary with timing information



Adversary with timing information

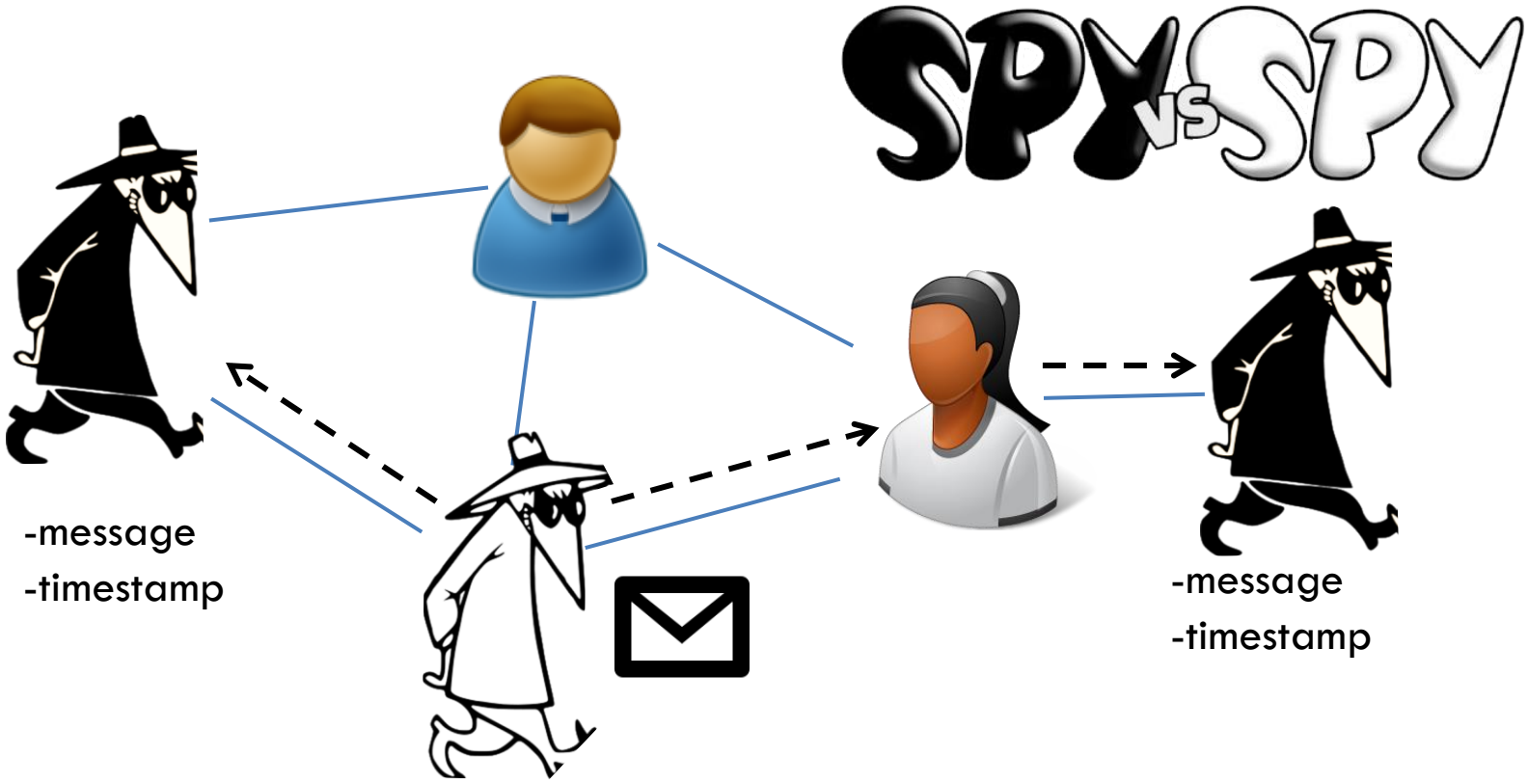


Adversary with timing information



adversary can collect timing information

Adversary with timing information

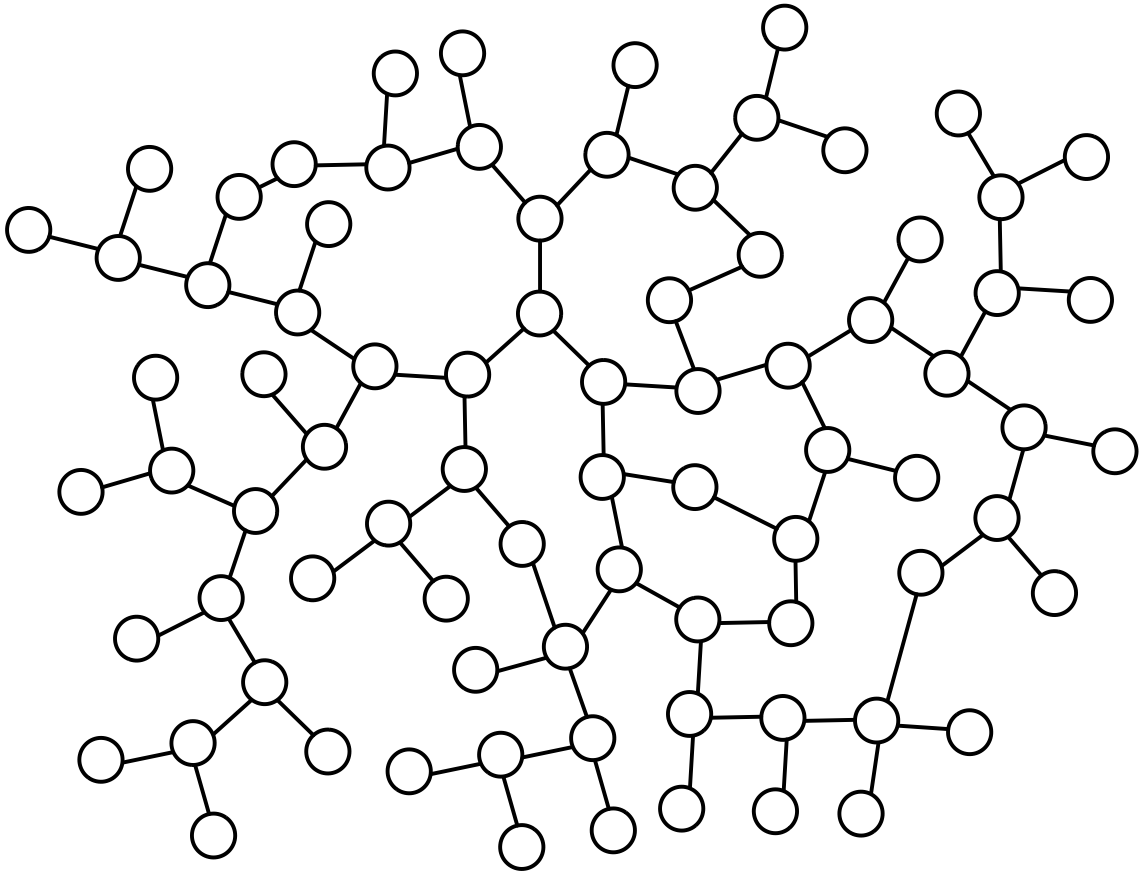


adversary can collect timing information

Distributed network forensics

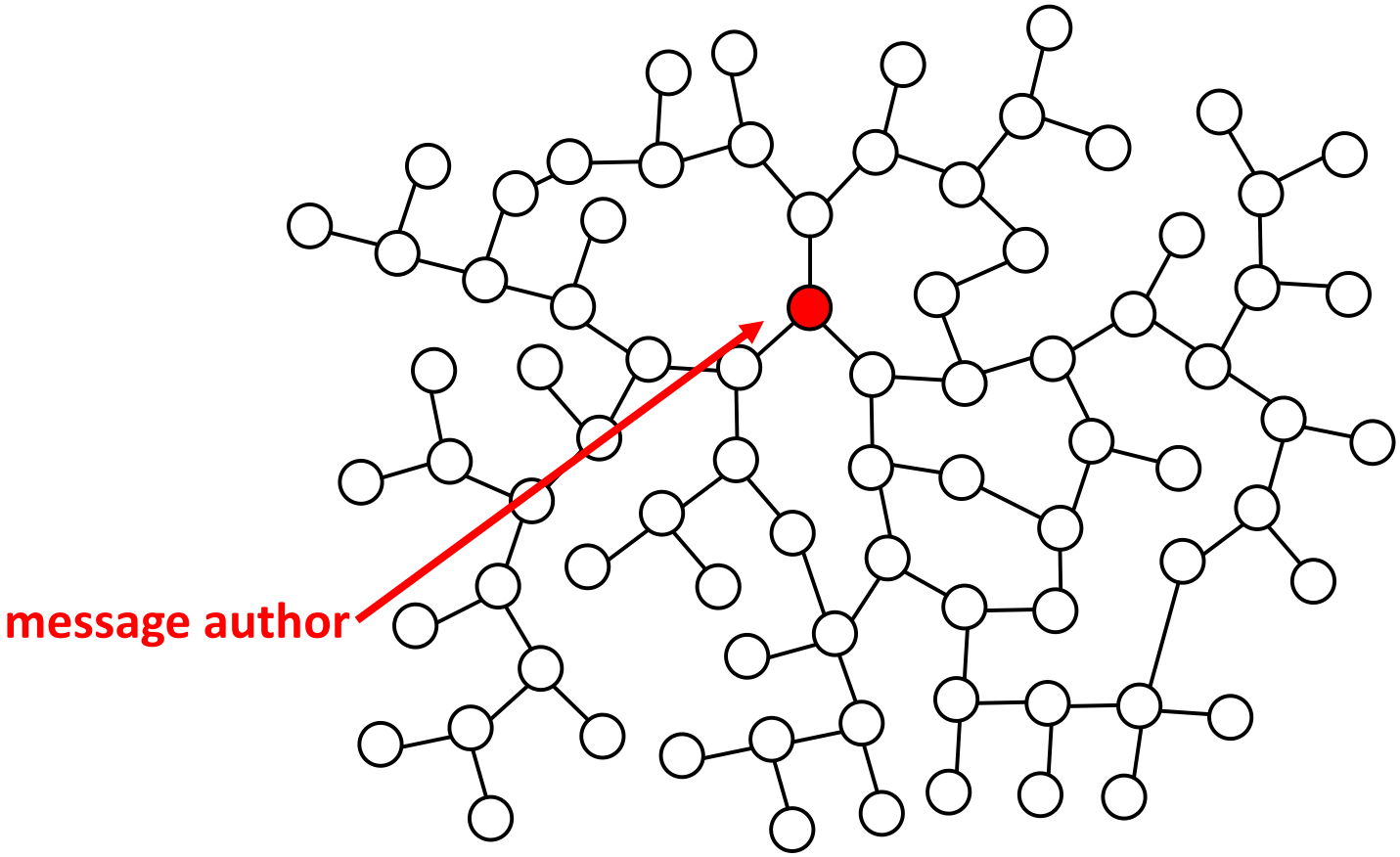
timing + who has the message = authorship

Information flow in social networks

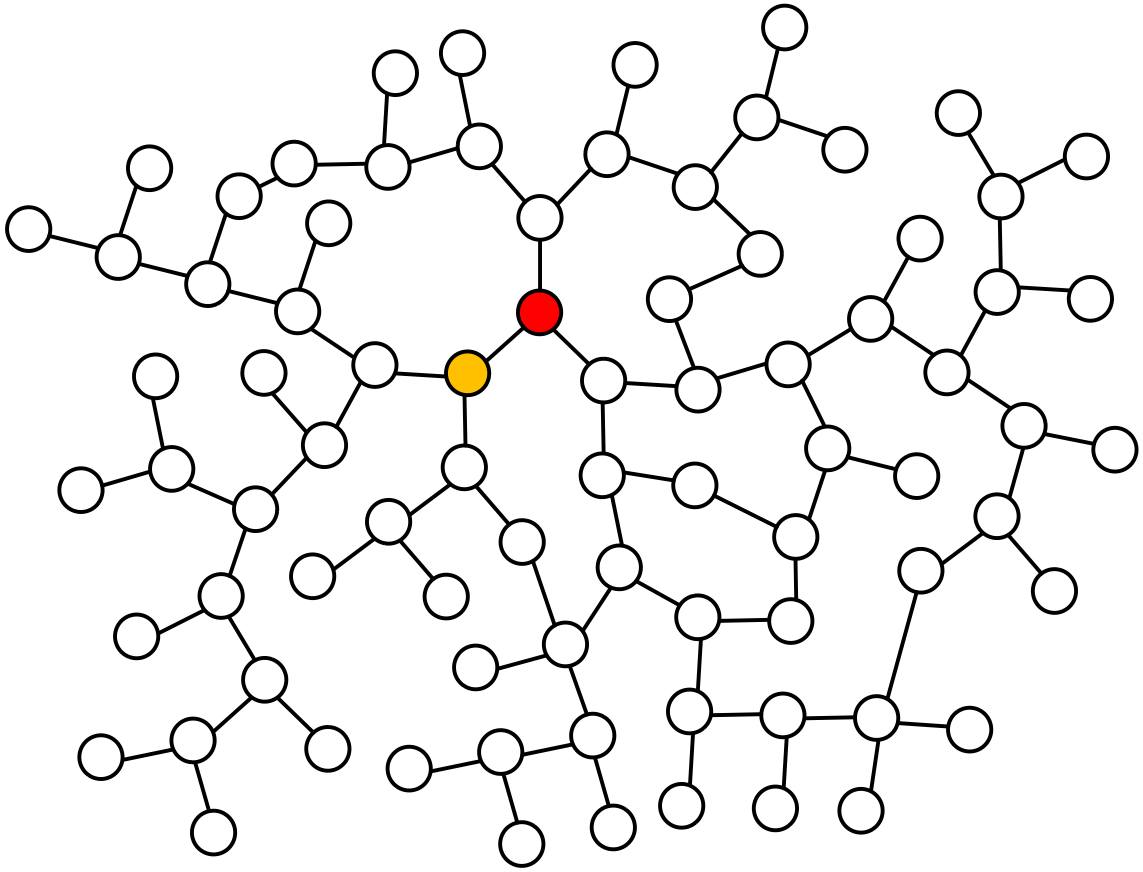


- G is the graph representing the social network

Information flow in social networks

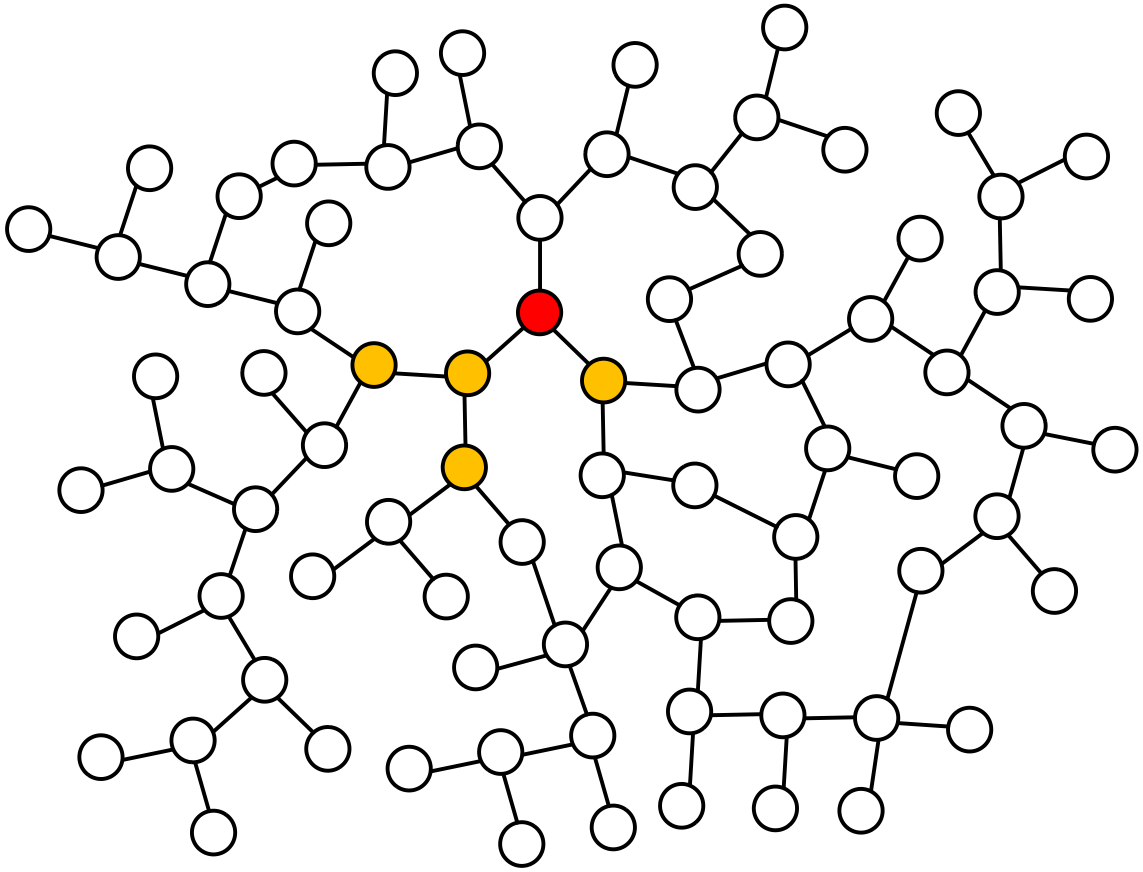


Information flow in social networks



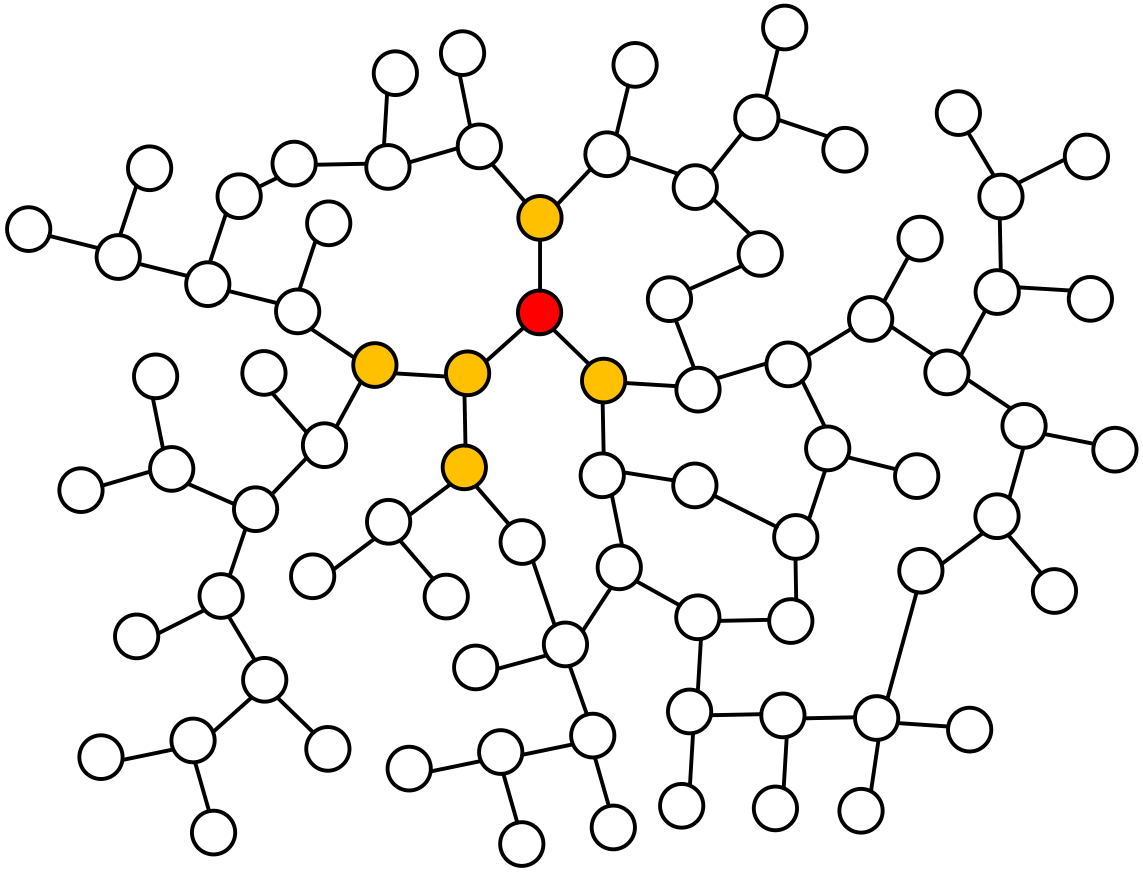
- the author passes the message to its neighbors

Information flow in social networks



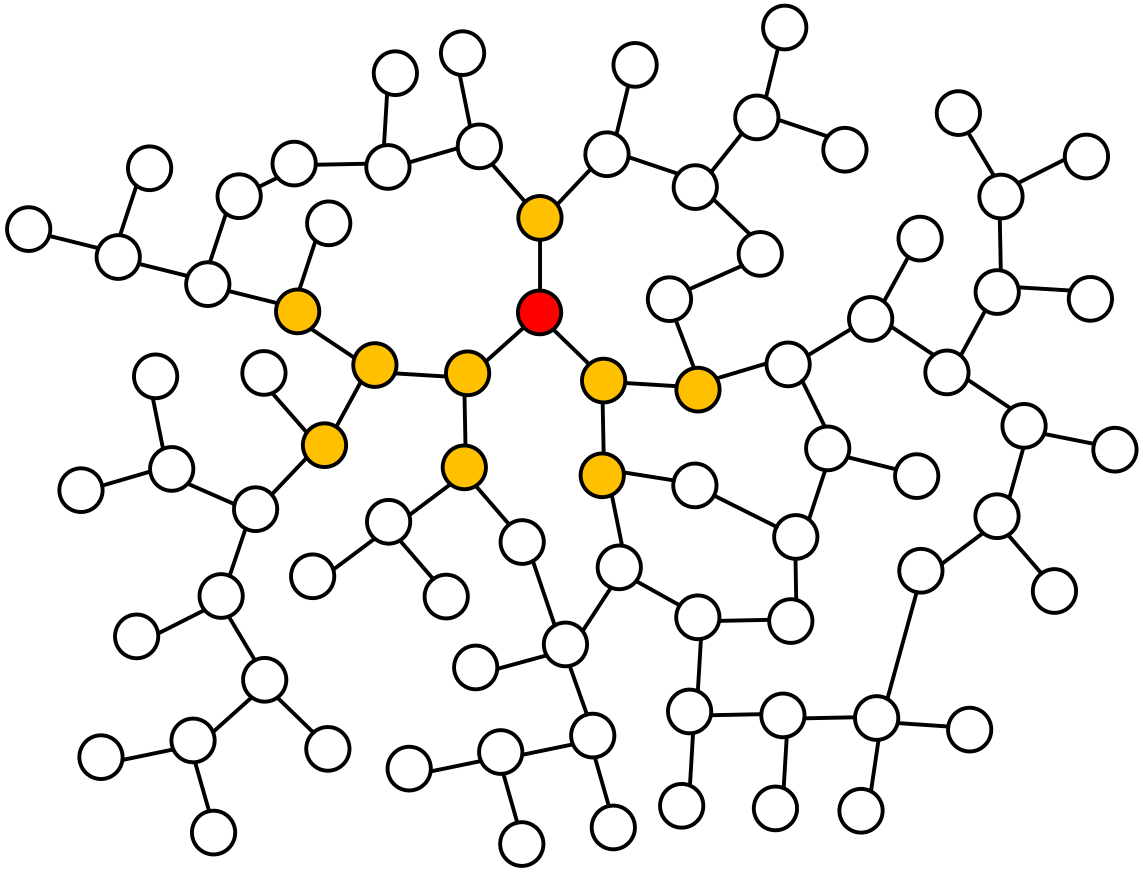
- its neighbors pass the message to theirs

Information flow in social networks



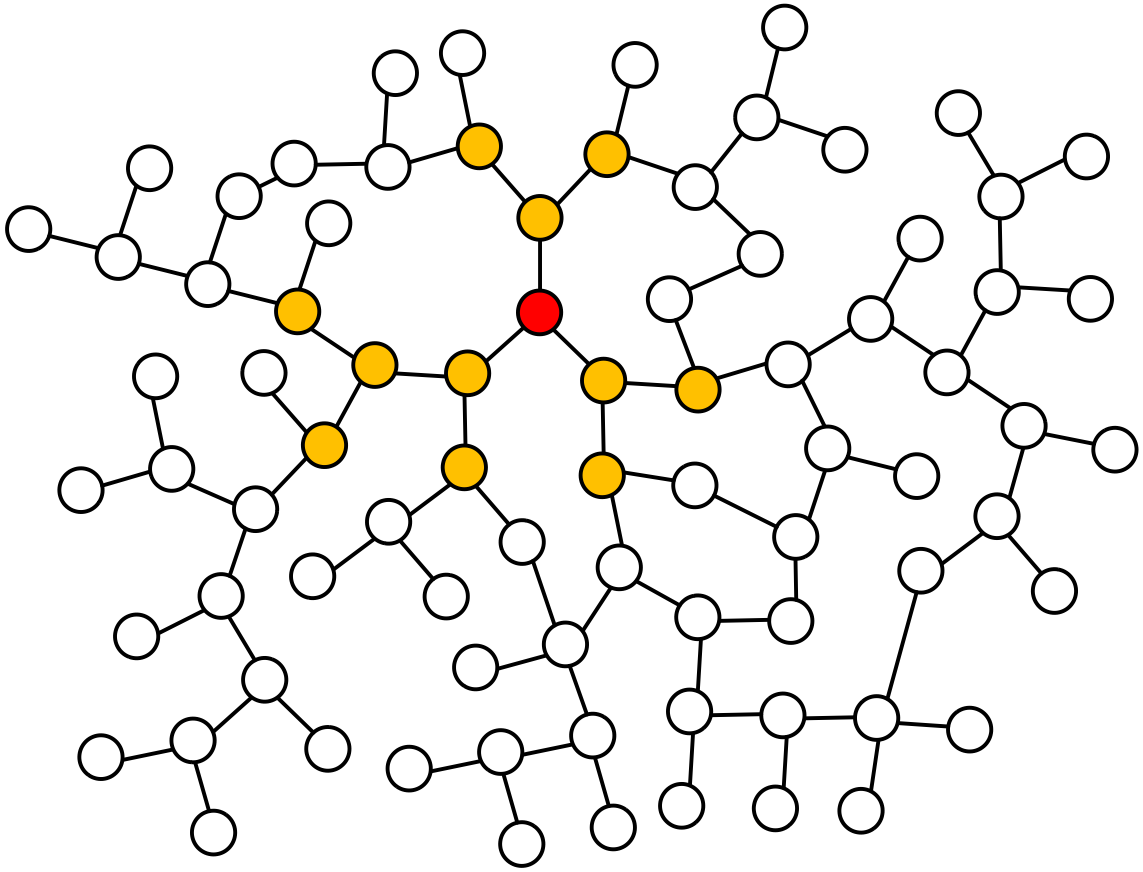
- the message spreads in **all directions** at the **same rate**

Information flow in social networks



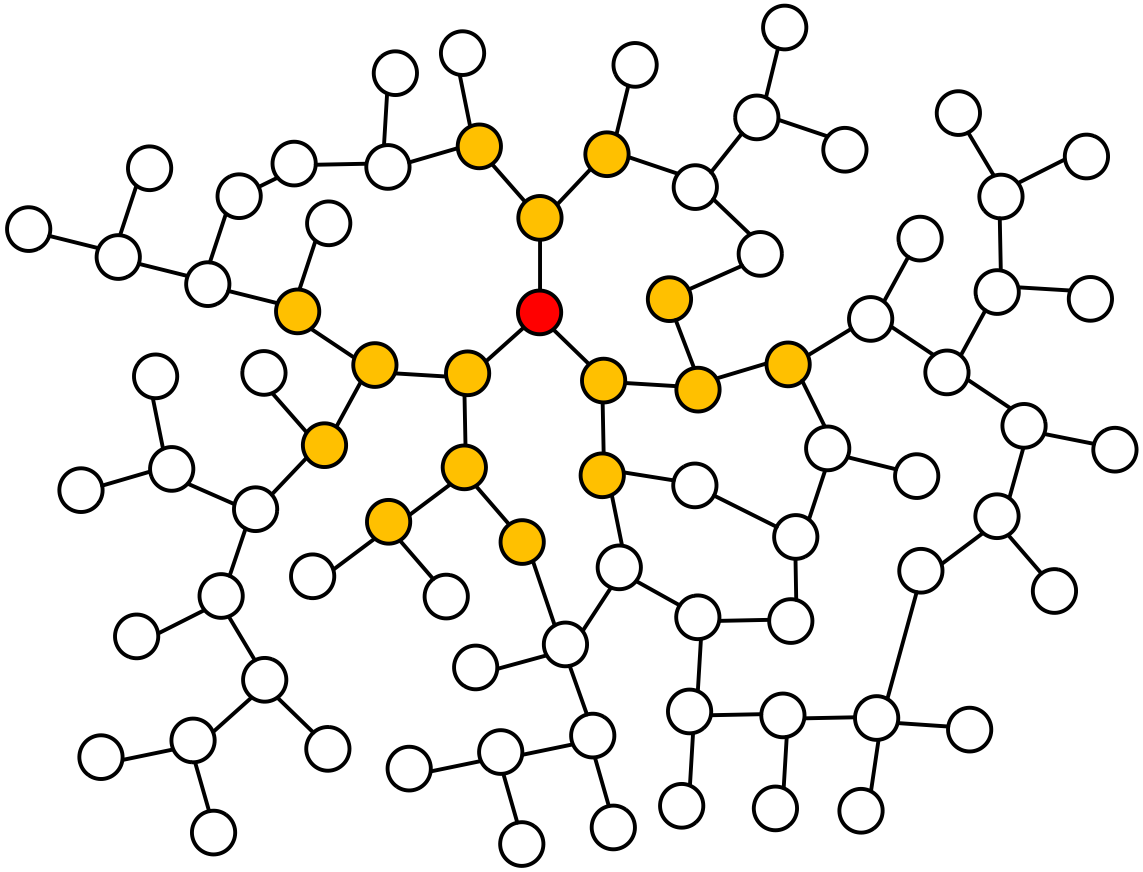
- the message spreads in **all directions** at the **same rate**

Information flow in social networks



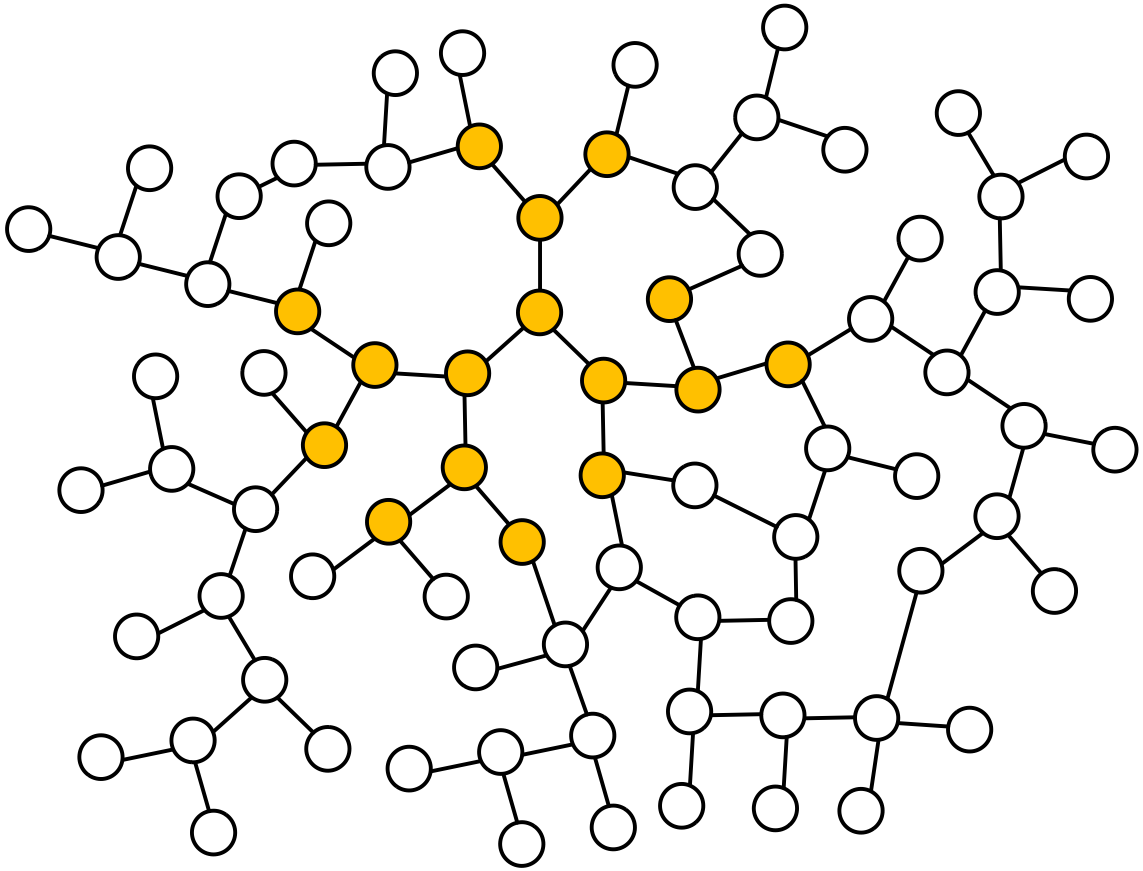
- the message spreads in **all directions** at the **same rate**

Information flow in social networks



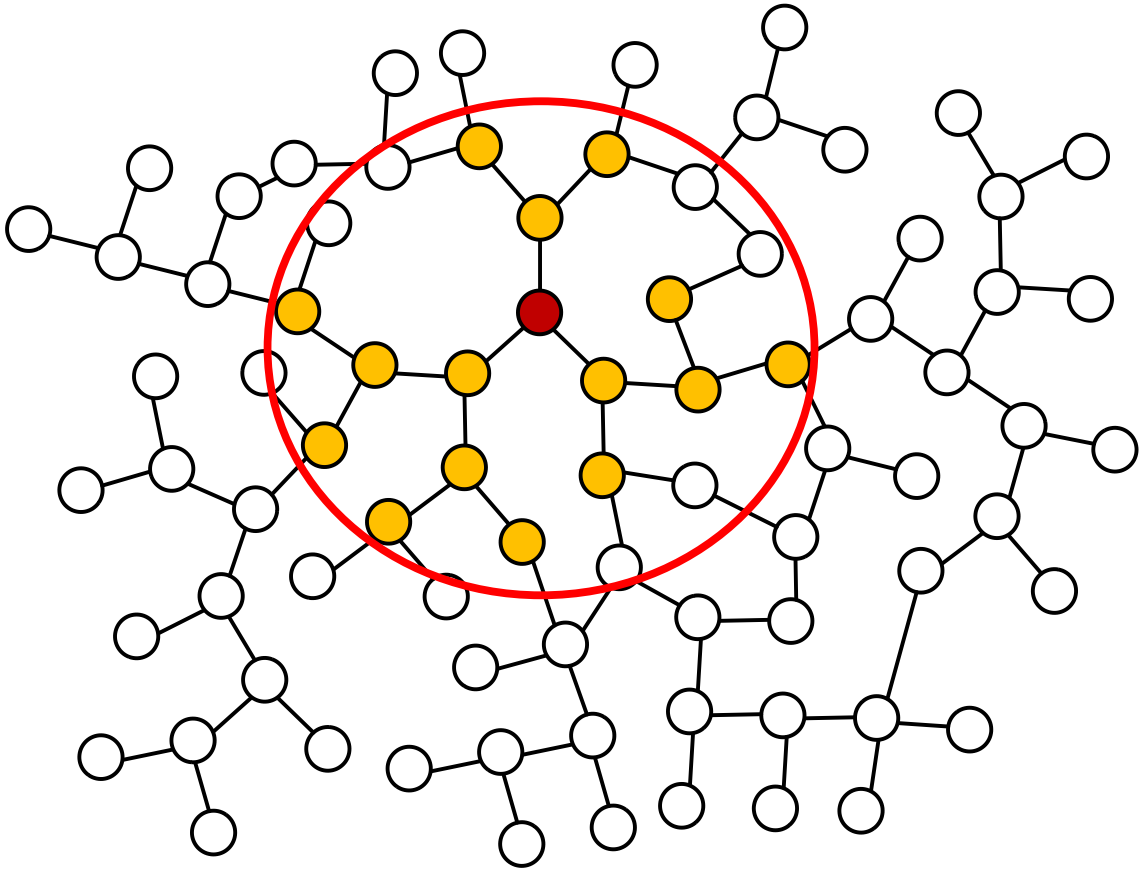
- this **spreading model** is known as the **diffusion model**

Adversary



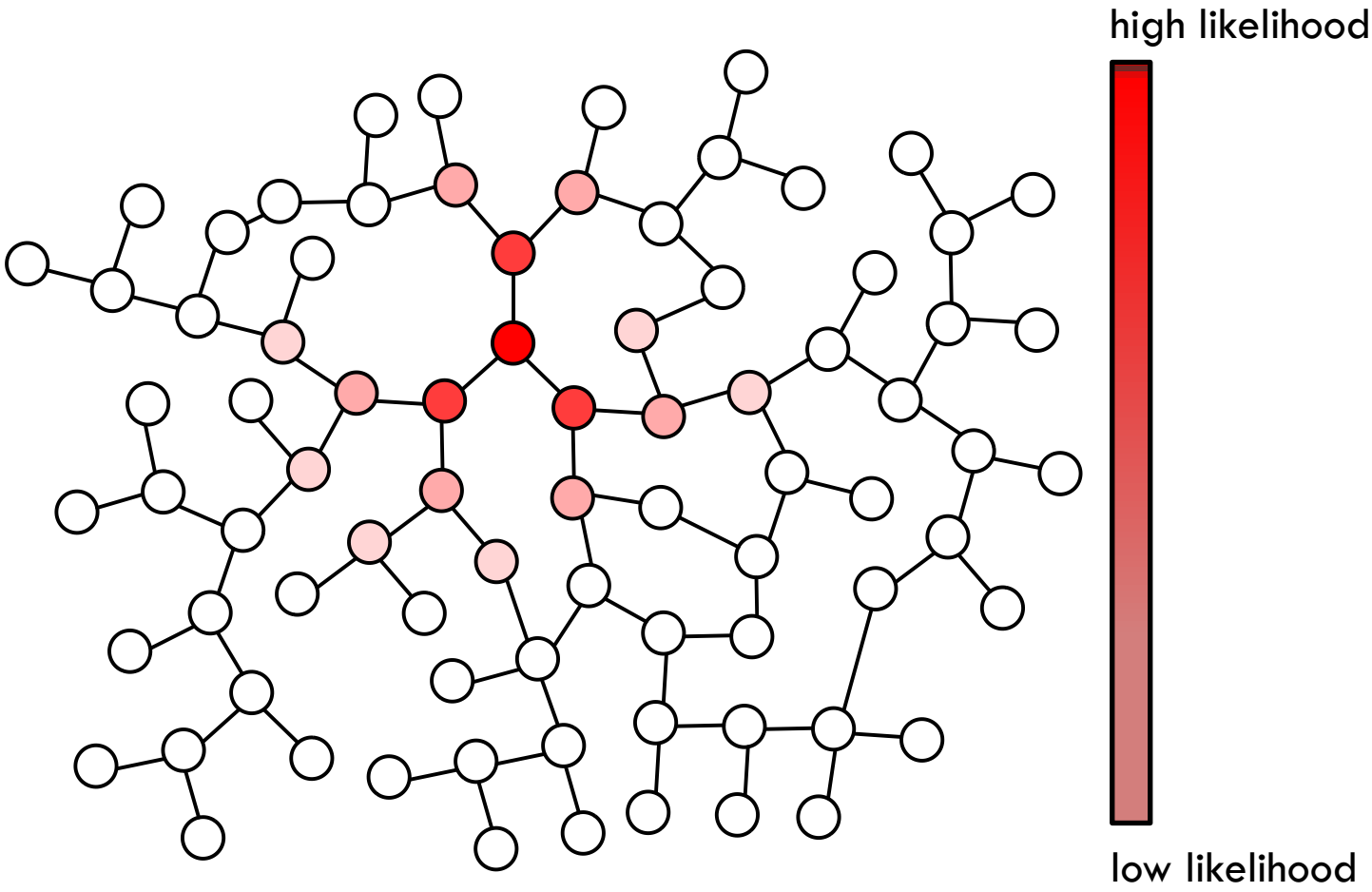
can we locate the message author?

Concentration around the center



- the **message author** is in the **"center"**

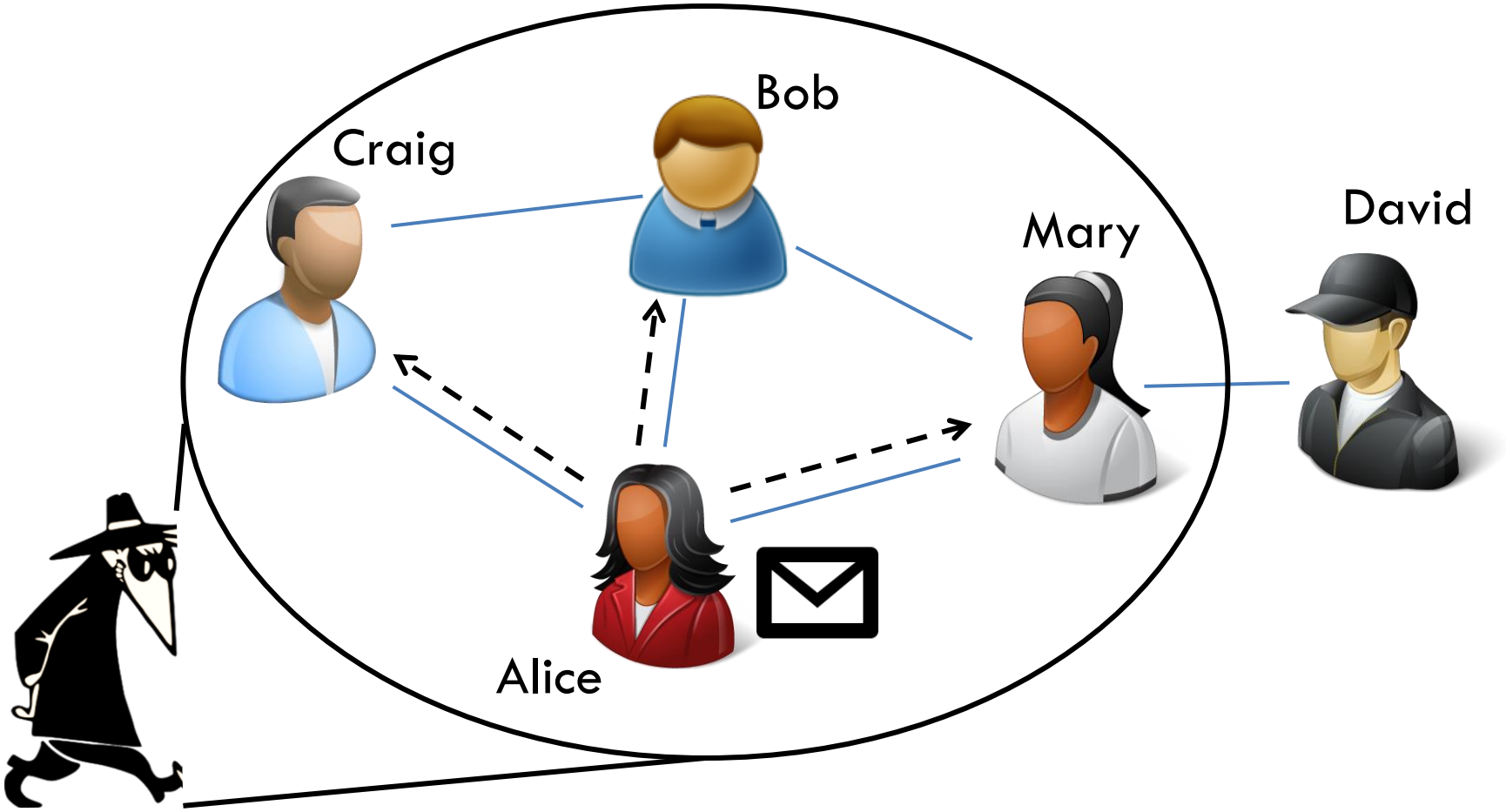
Maximum likelihood detection



diffusion spreading = deanonymization

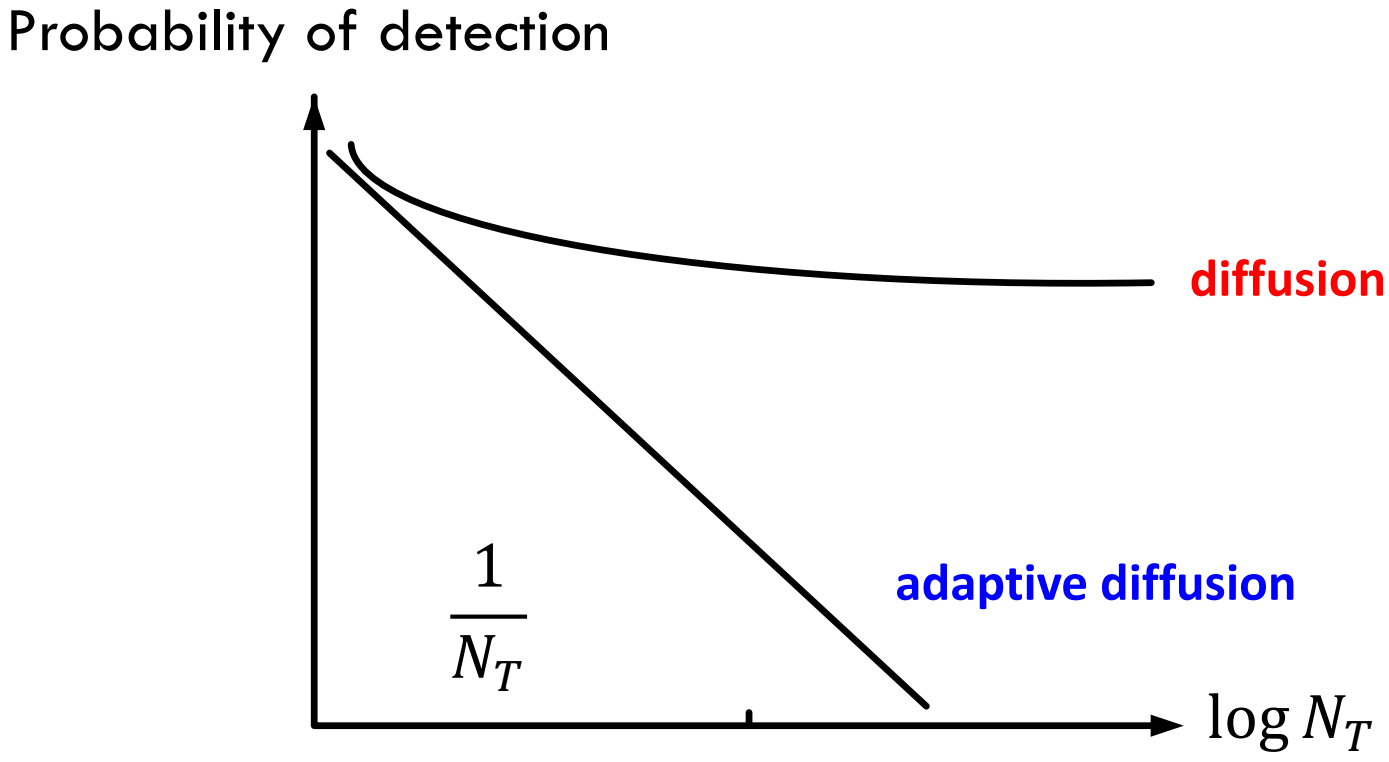
[*Shah, Zaman* 2011]

Our goal



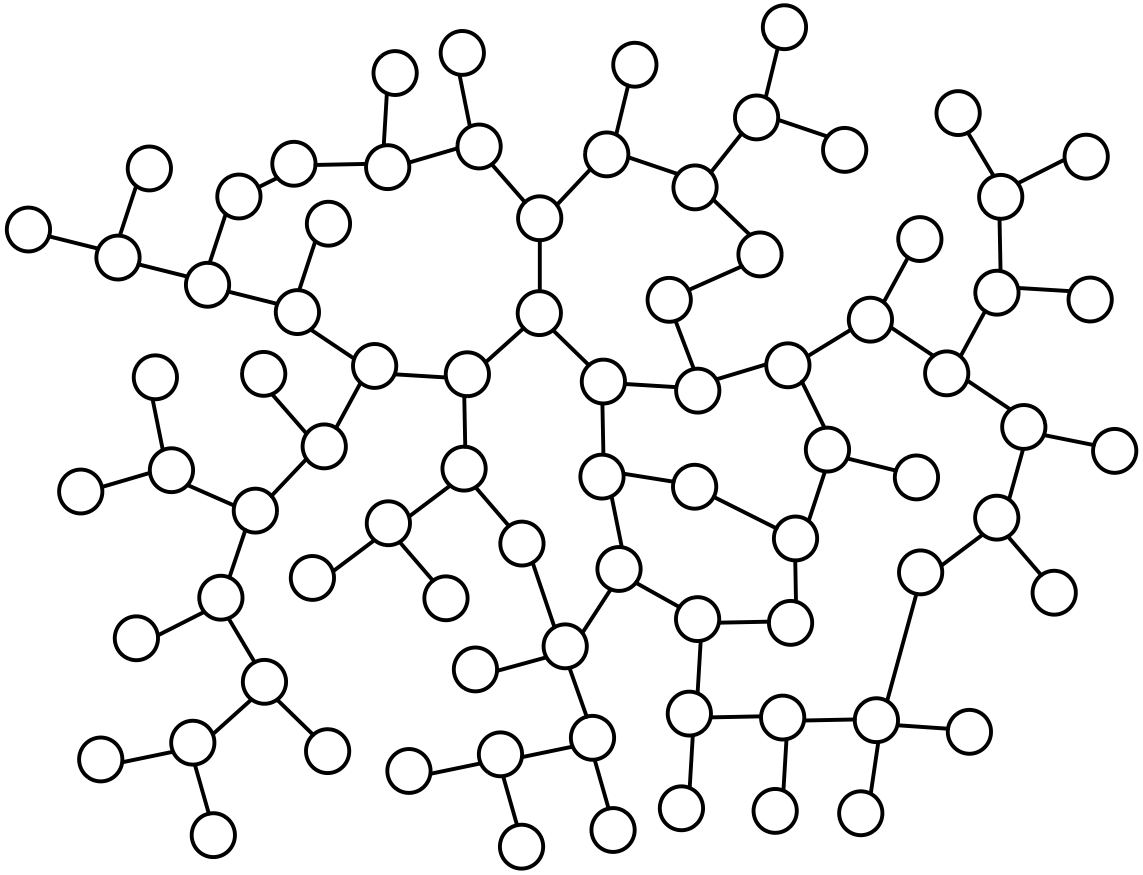
engineer the spread to hide authorship

Our goal

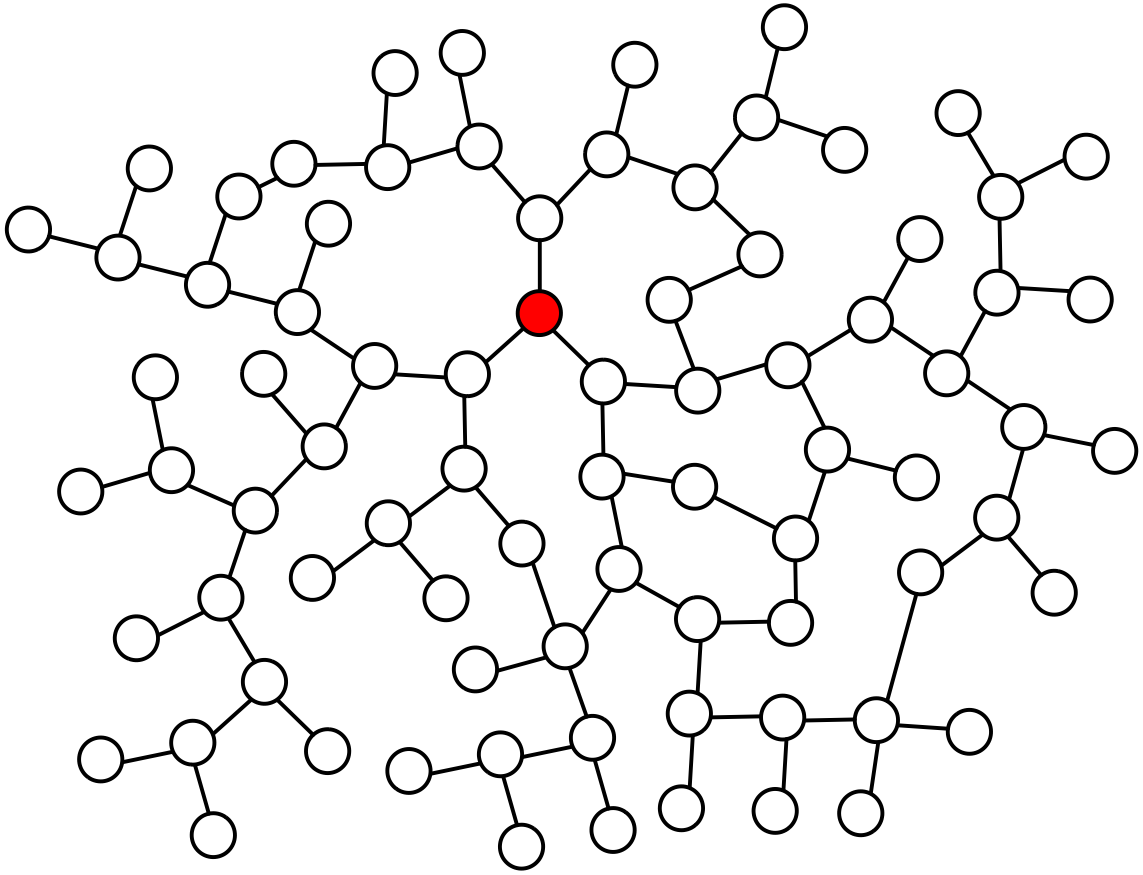


- N_T: **expected number** of nodes with the message at time T

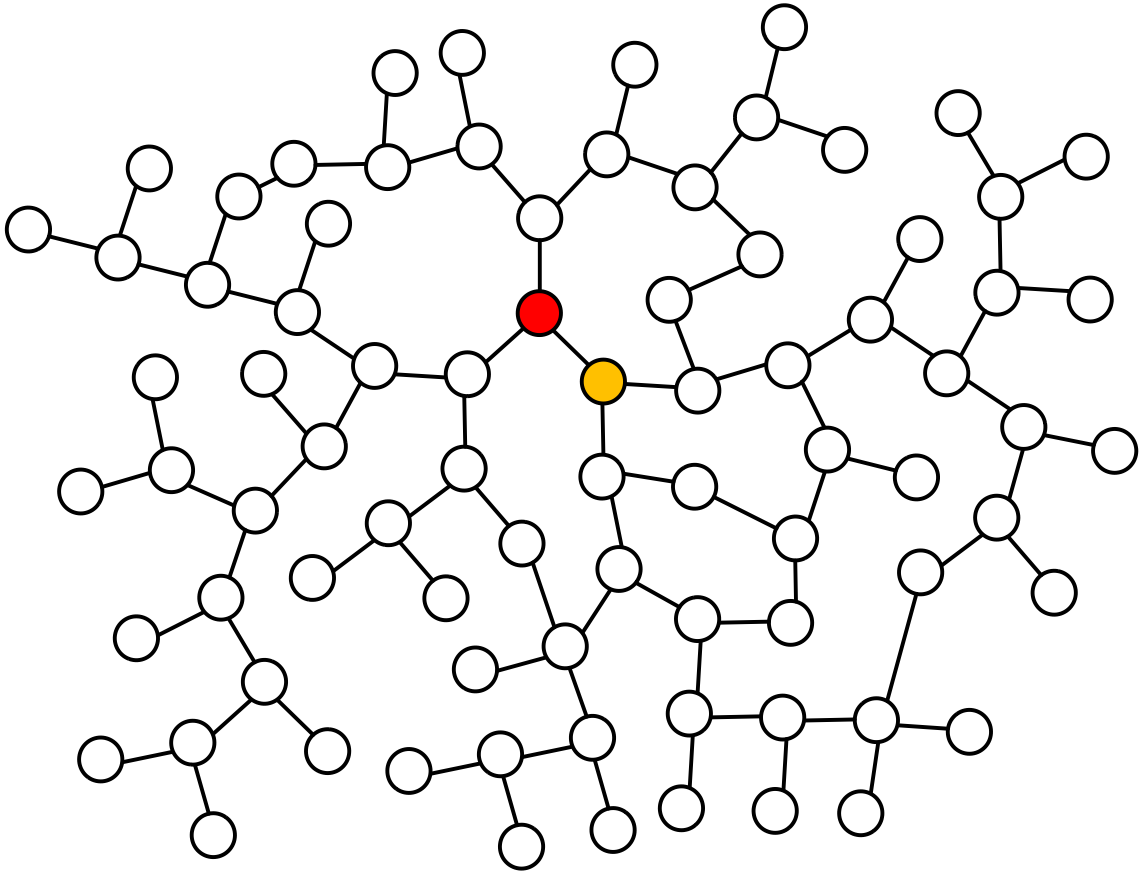
Main Result: Adaptive diffusion



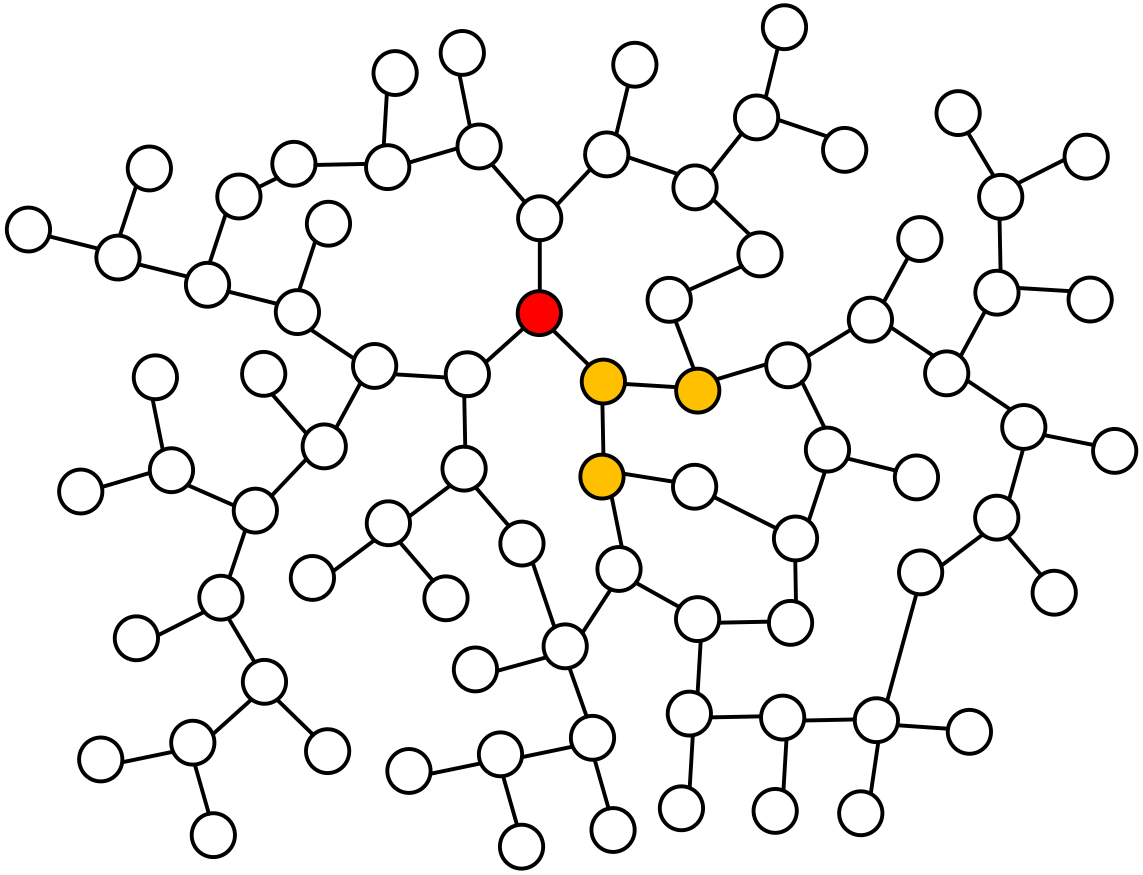
Main Result: Adaptive diffusion



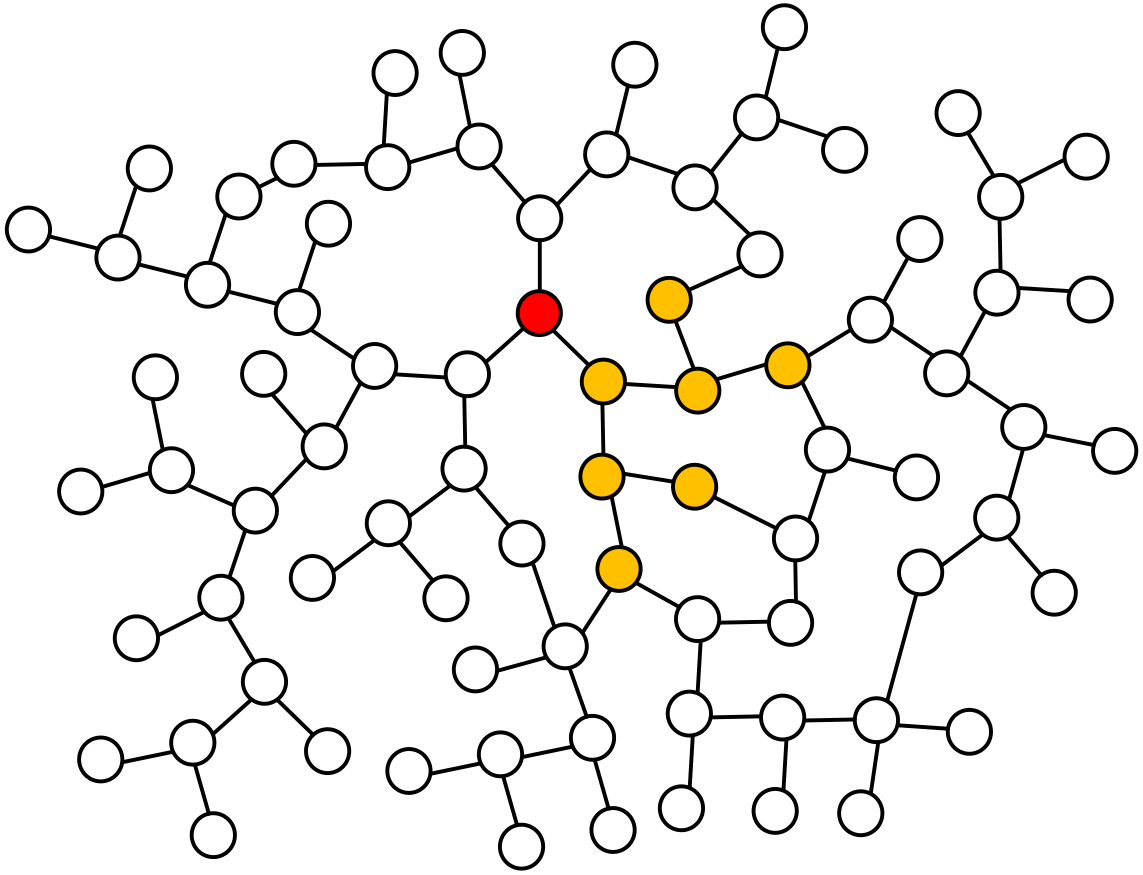
Main Result: Adaptive diffusion



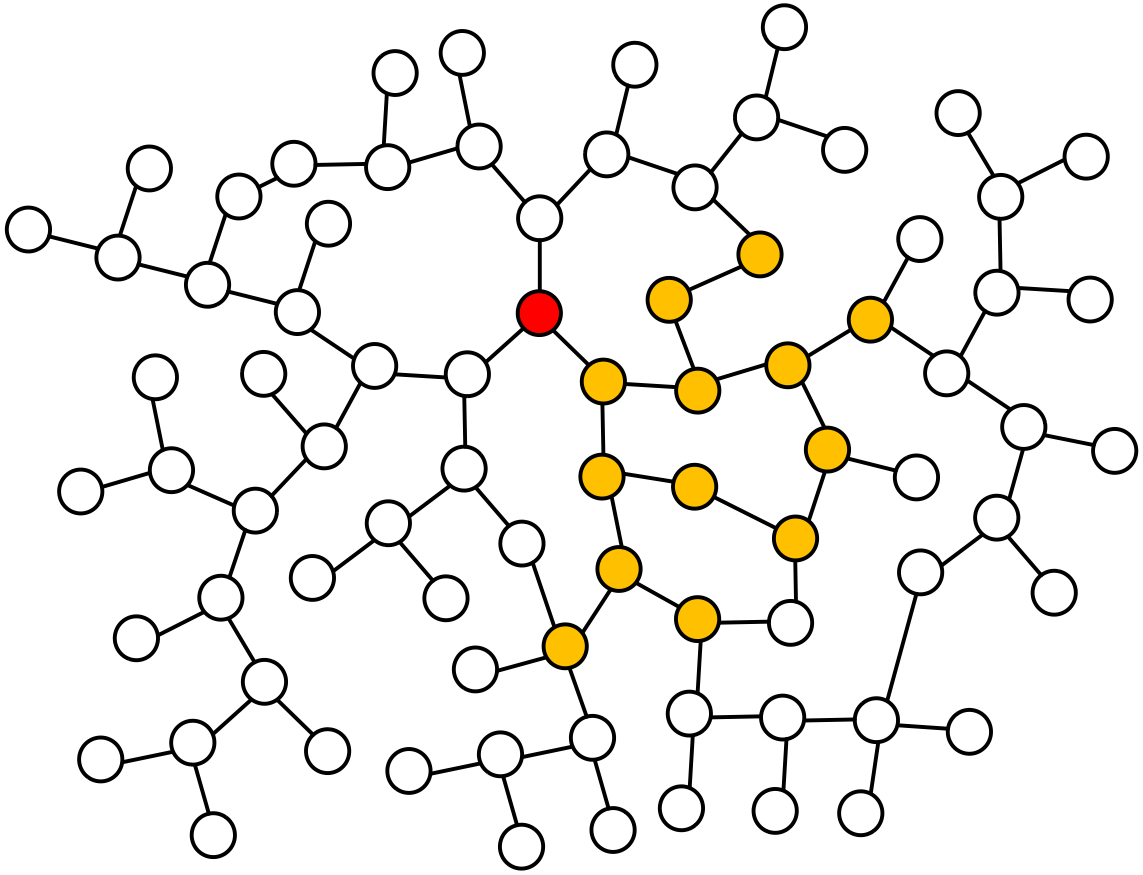
Main Result: Adaptive diffusion



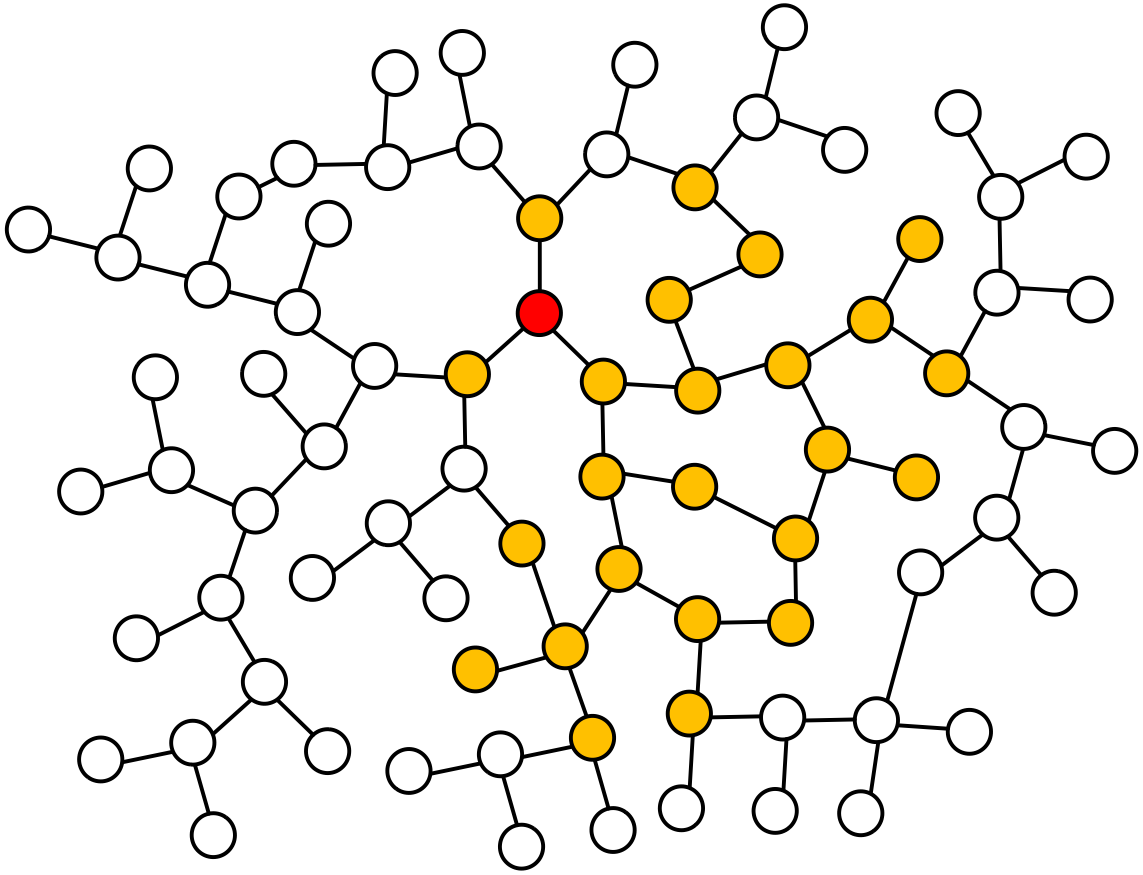
Main Result: Adaptive diffusion



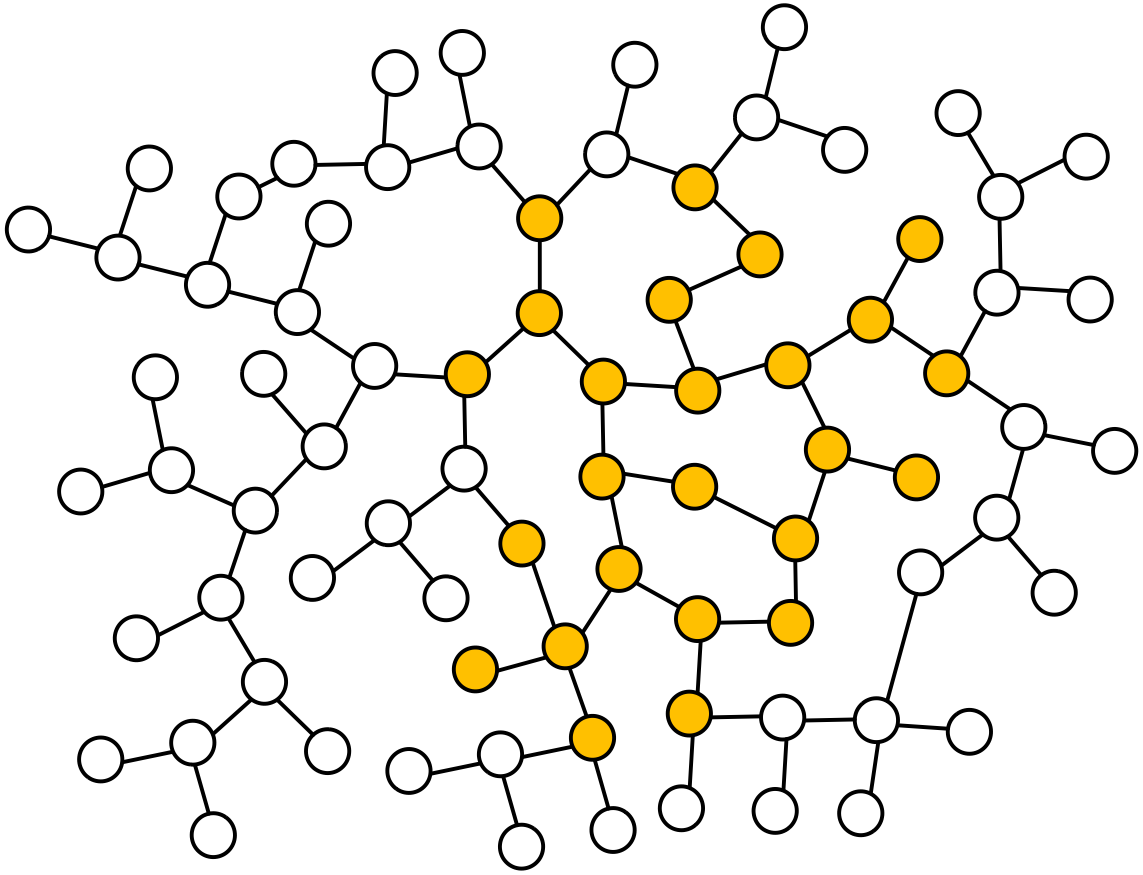
Main Result: Adaptive diffusion



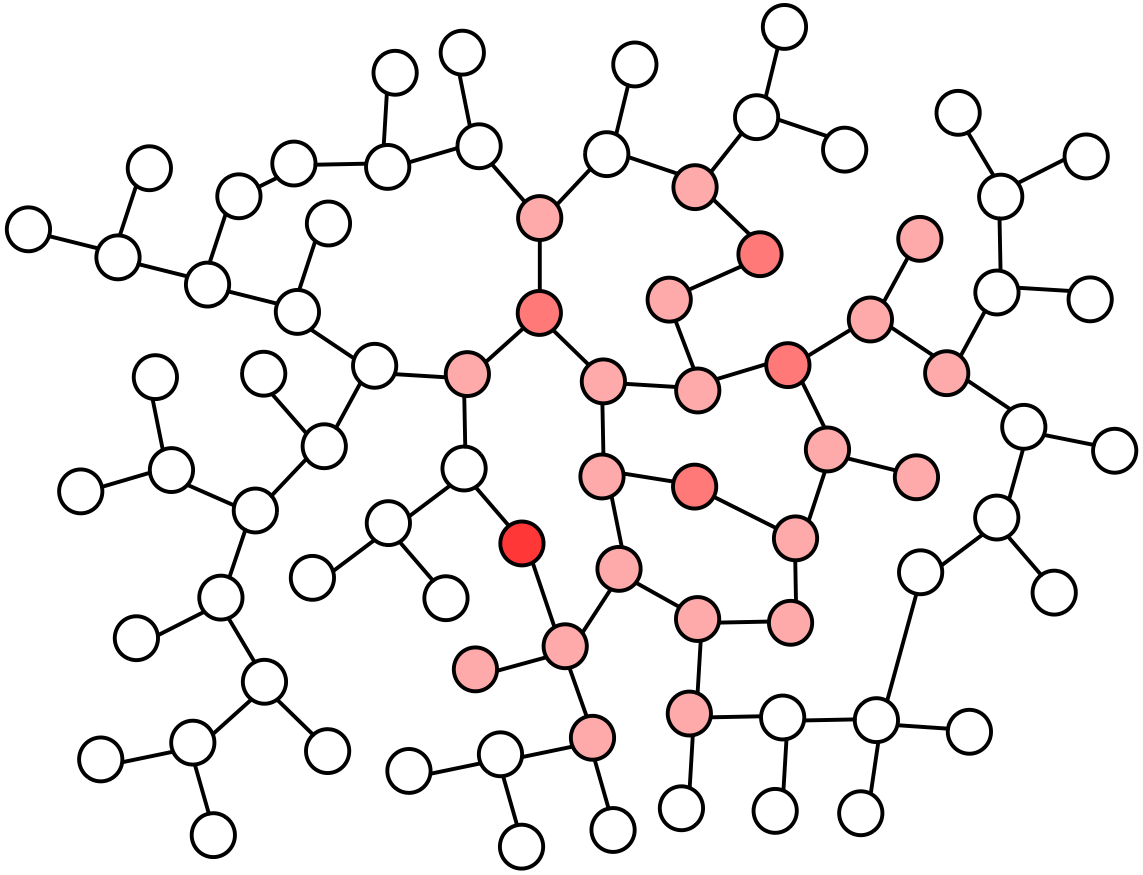
Main Result: Adaptive diffusion



Main Result: Adaptive diffusion



Main Result: Adaptive diffusion

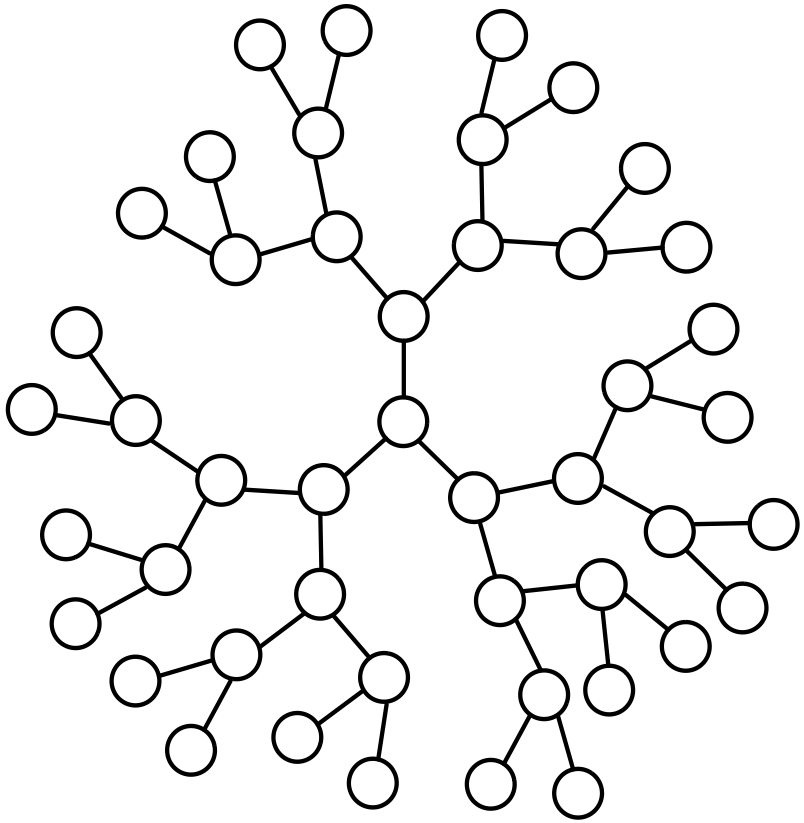


provides provable anonymity guarantees

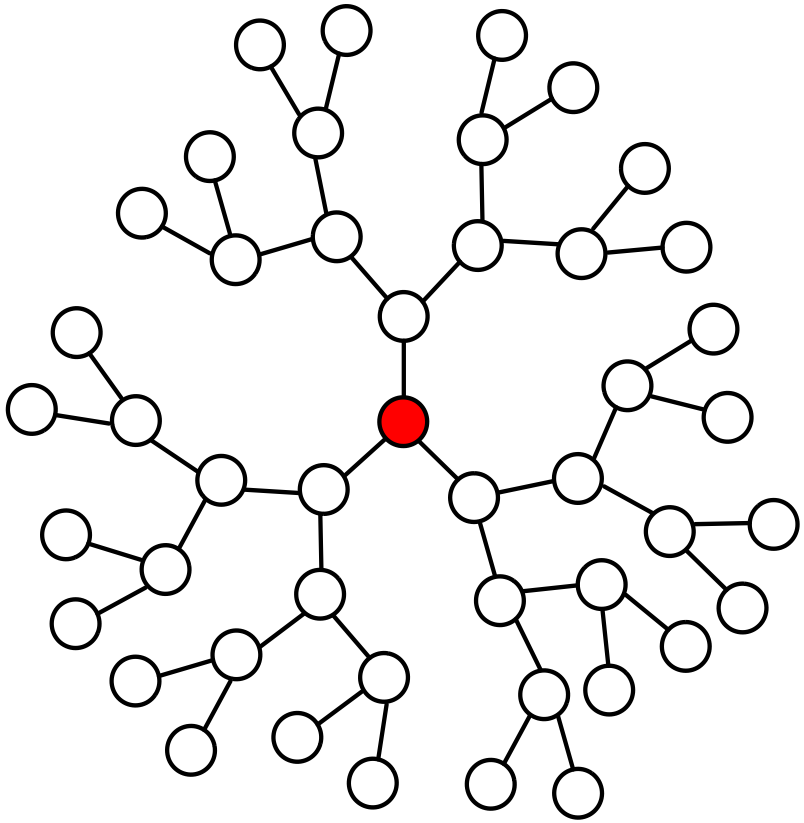
[Best Paper Award at SIGMETRICS 15]

Part 1.1:
Adversary without timing info

d-regular trees : adaptive diffusion

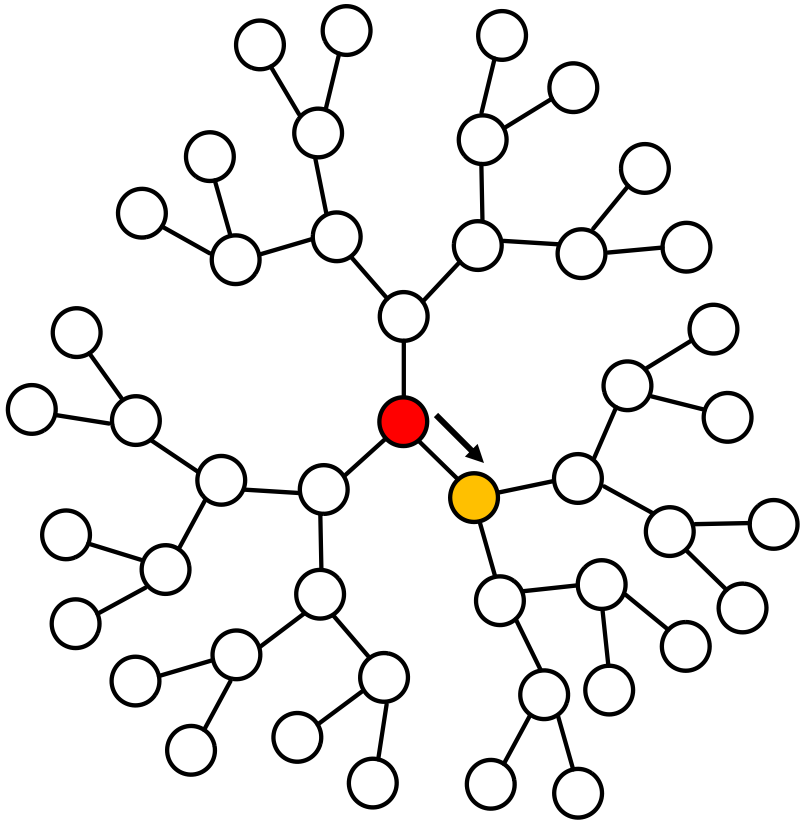


d-regular trees : adaptive diffusion



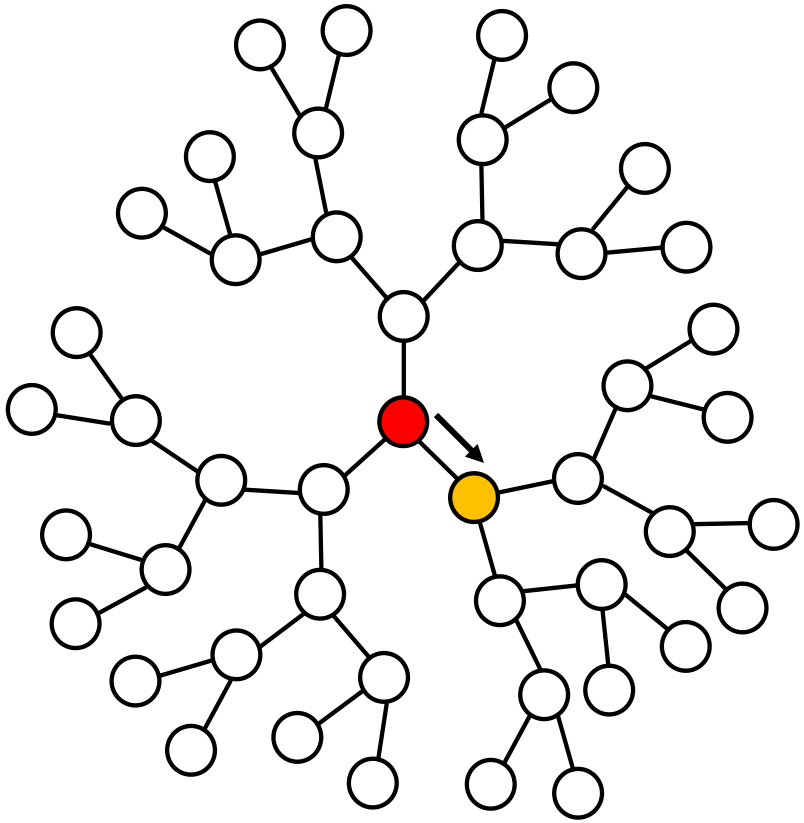
- initially, the author is also the **"virtual source"**

d-regular trees : adaptive diffusion



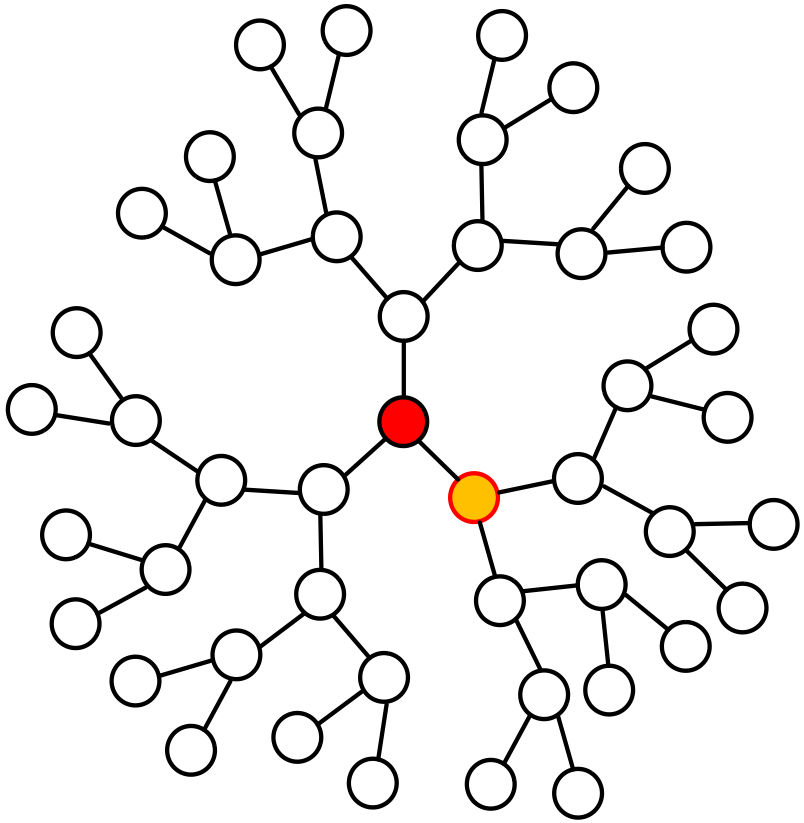
- at $T = 1$, the author selects one neighbor at random

d-regular trees : adaptive diffusion



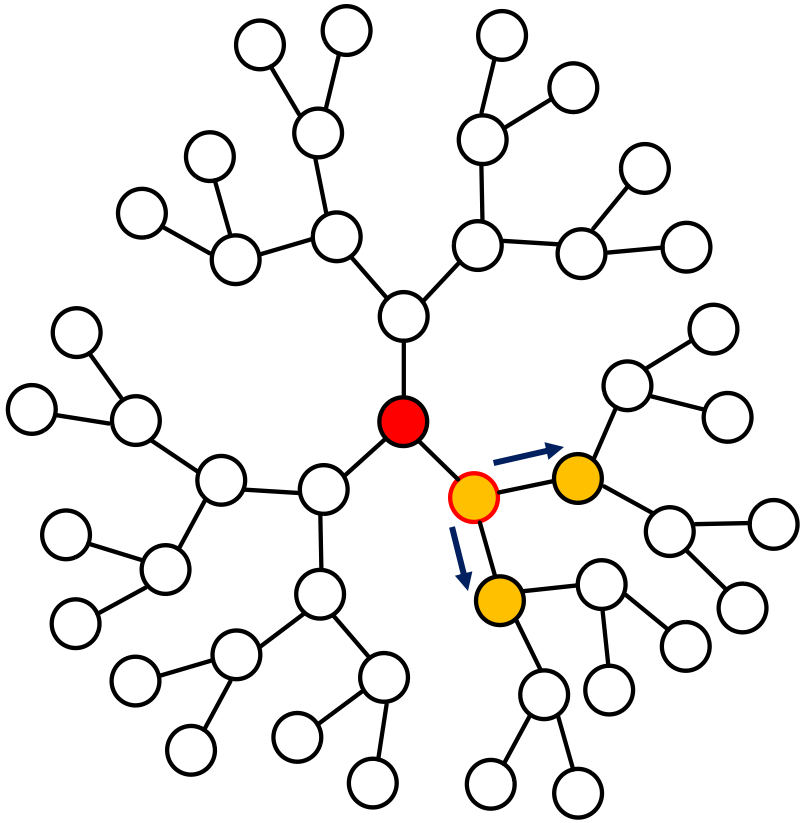
- at $T = 1$, the author selects one neighbor at random

d-regular trees : adaptive diffusion



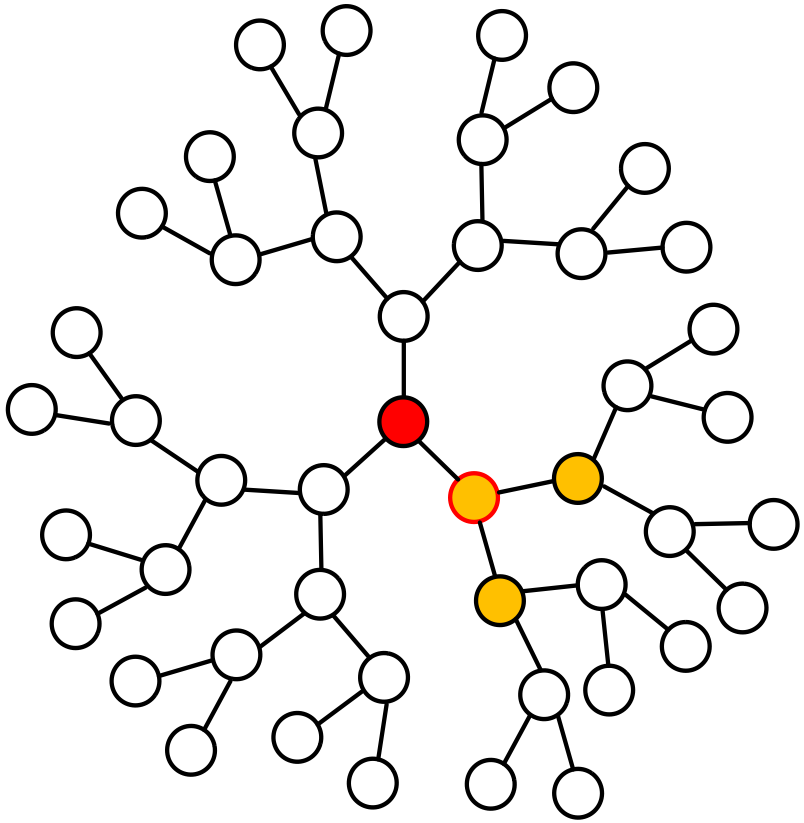
- the chosen neighbor becomes the **new virtual source**

d-regular trees : adaptive diffusion



- at $T = 2$, the **virtual source** passes the message to all its neighbors

d-regular trees : adaptive diffusion

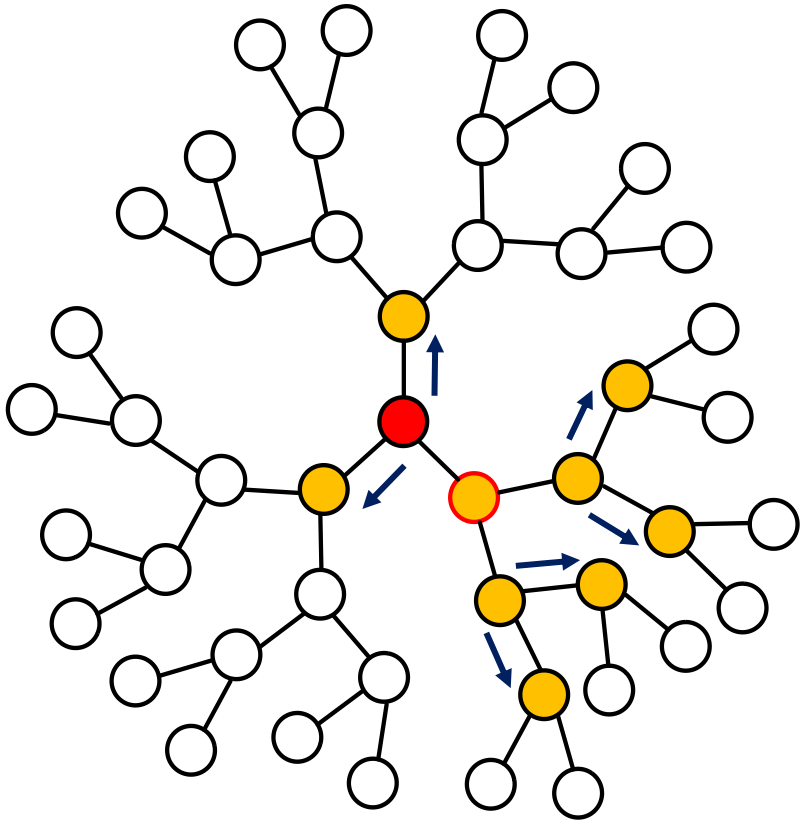


- the new virtual source has two options:

keeping the virtual source token

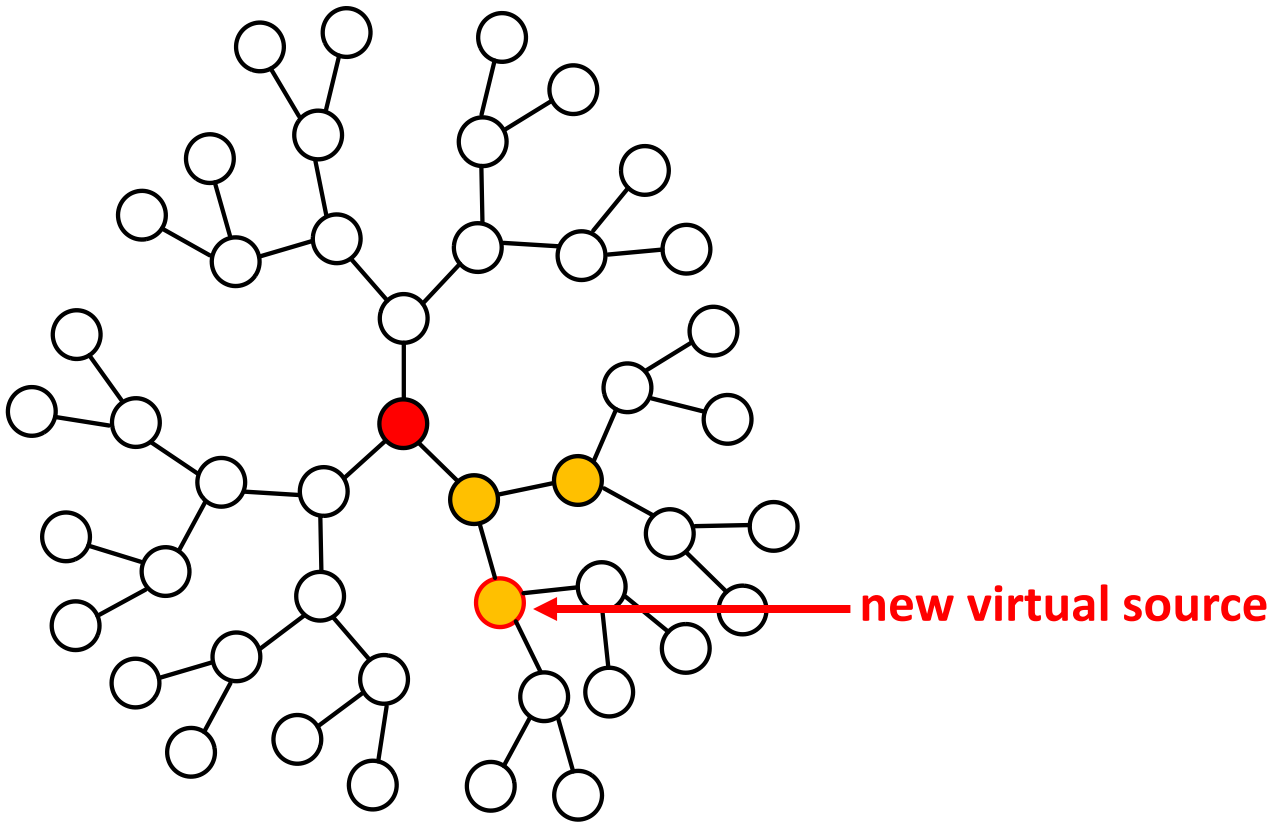
passing the virtual source token

Keeping the virtual source token



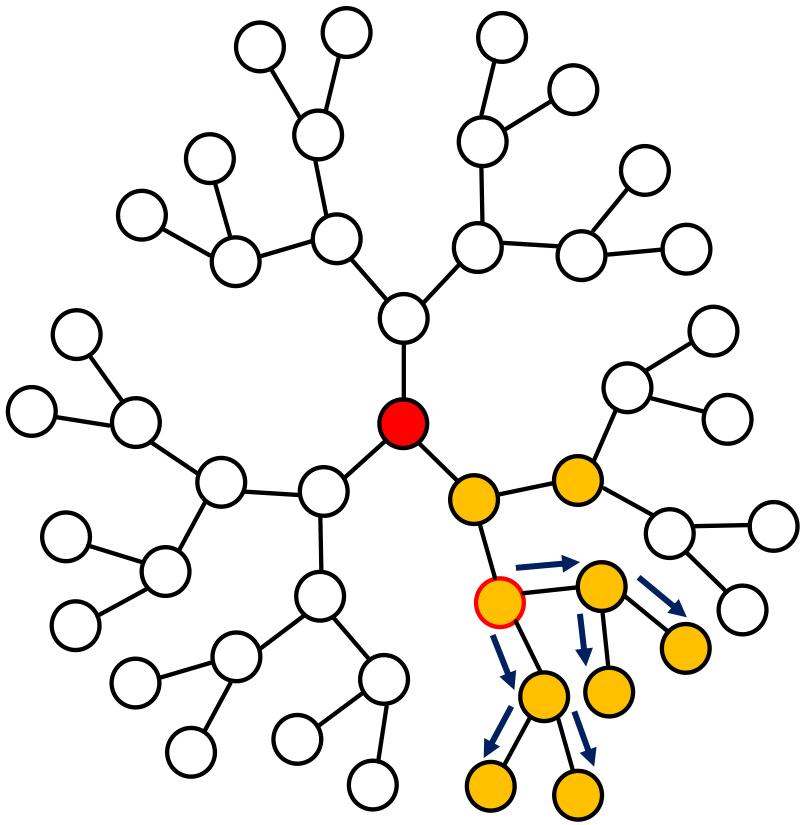
- all leaf nodes with the message pass it to their neighbors

Passing the virtual source token



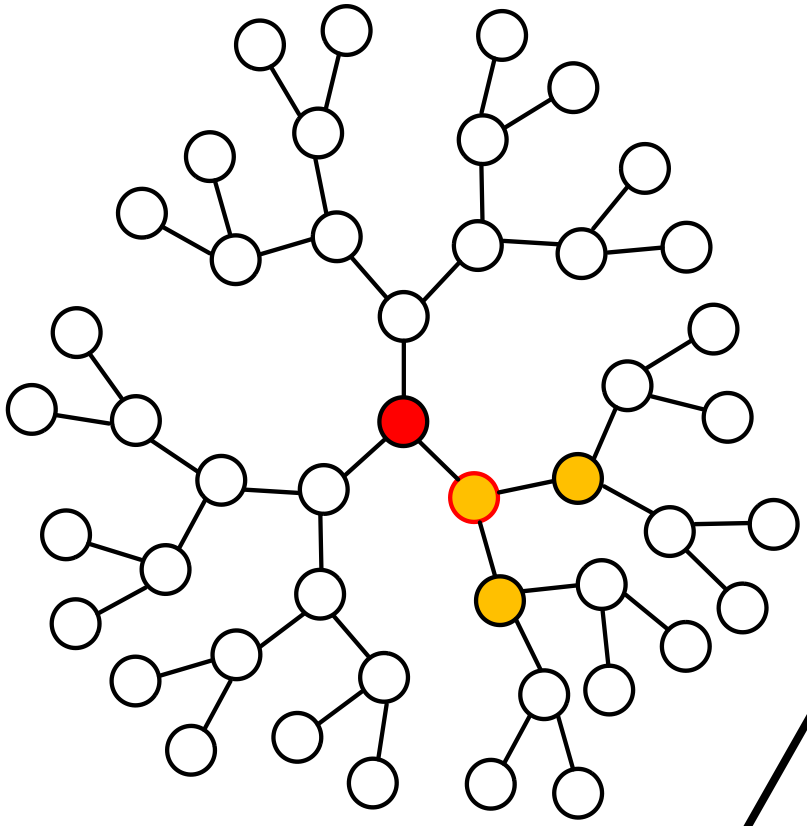
- current virtual source selects one of its neighbors at random

Passing the virtual source token



- new virtual source passes the message to its neighbors which in turn pass it to their neighbors

When to keep the virtual source?



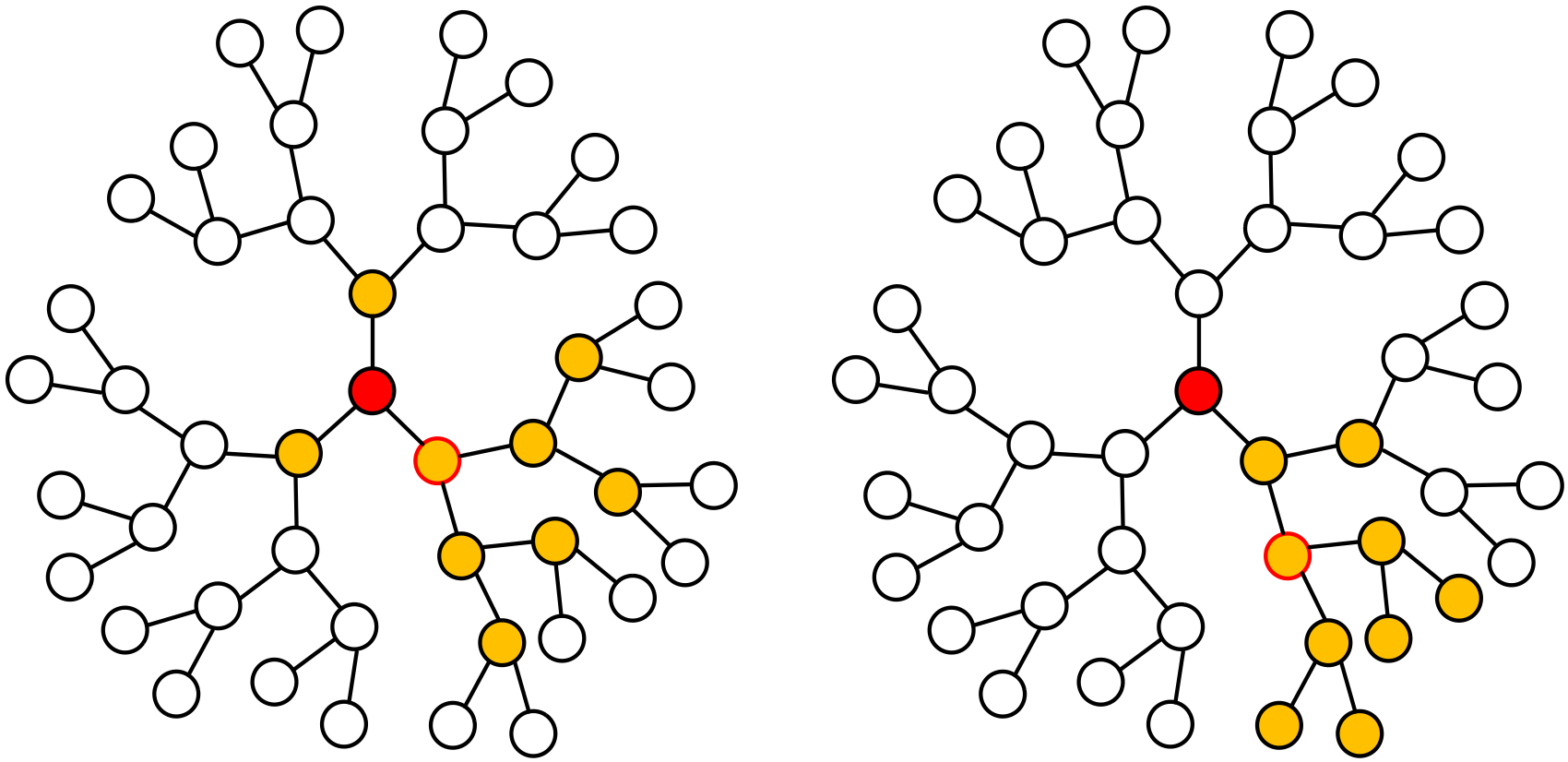
tree degree

time index

distance to source

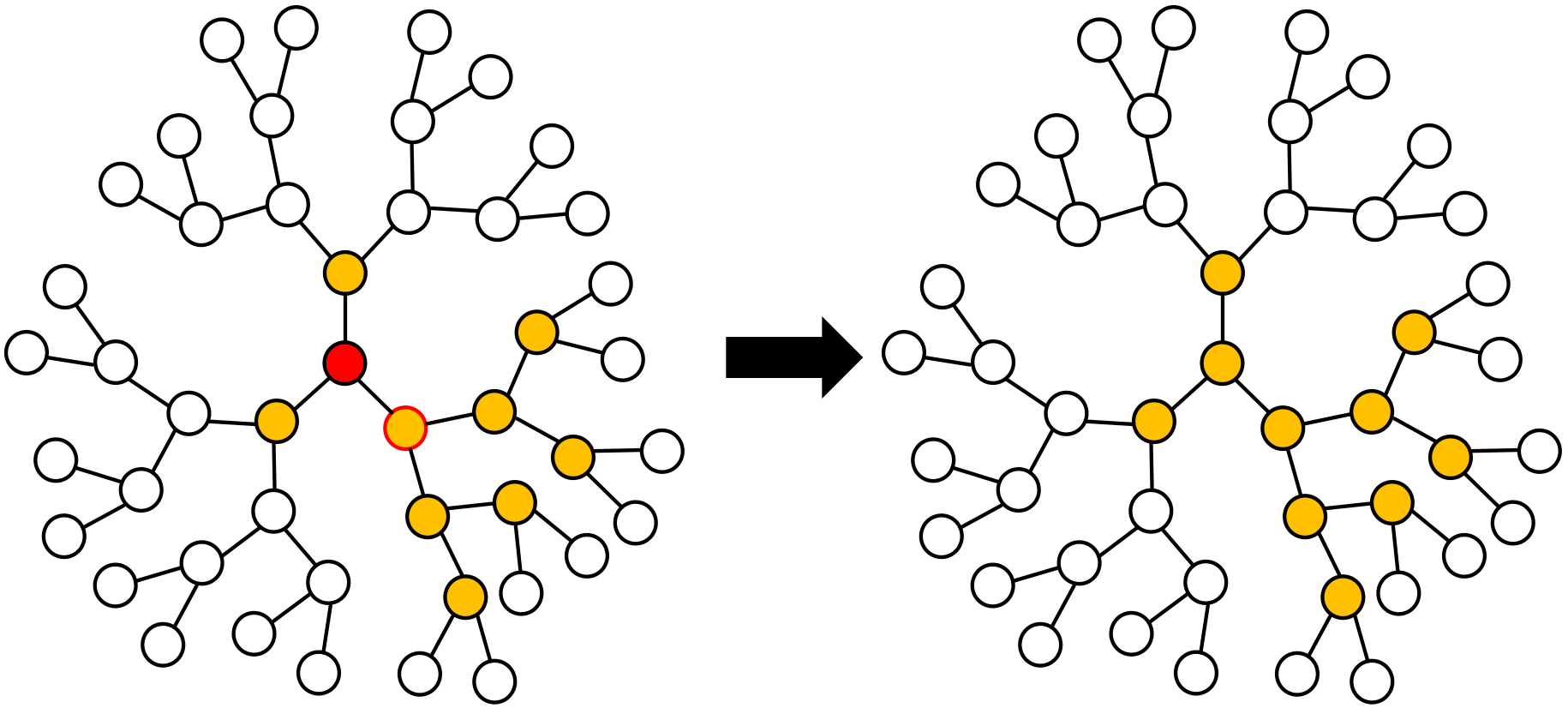
- virtual source token is kept with probability $\frac{(d-1)^{\frac{T}{2}-h-1}-1}{(d-1)^{\frac{T}{2}+1}-1}$

Symmetry properties



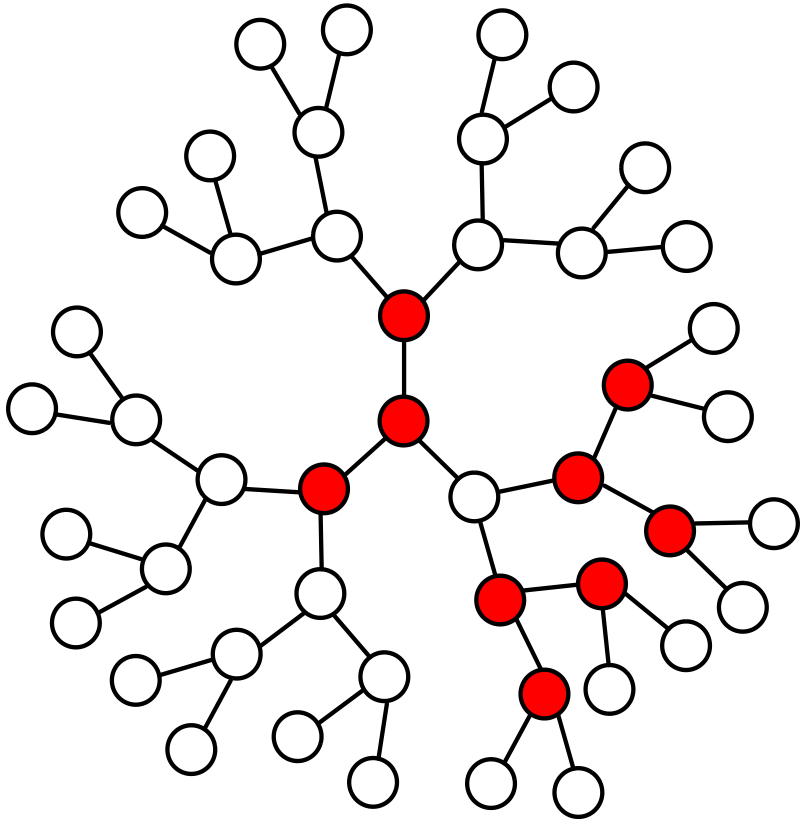
- the graph is **always symmetric** around the **virtual source**

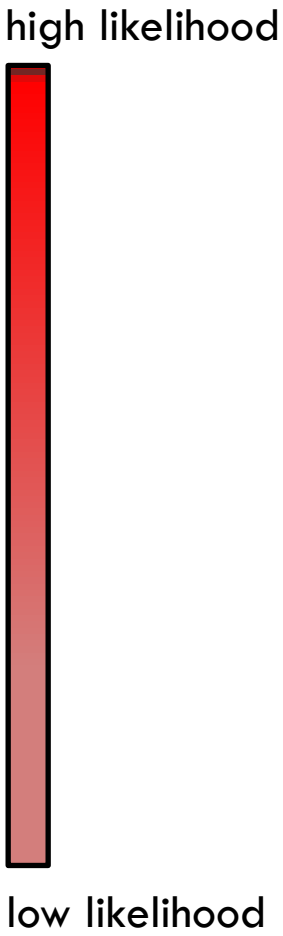
Adversary



can we locate the message author?

Maximum likelihood detection





- **all nodes** except for the final virtual source **are equally likely**

Main result: adaptive diffusion

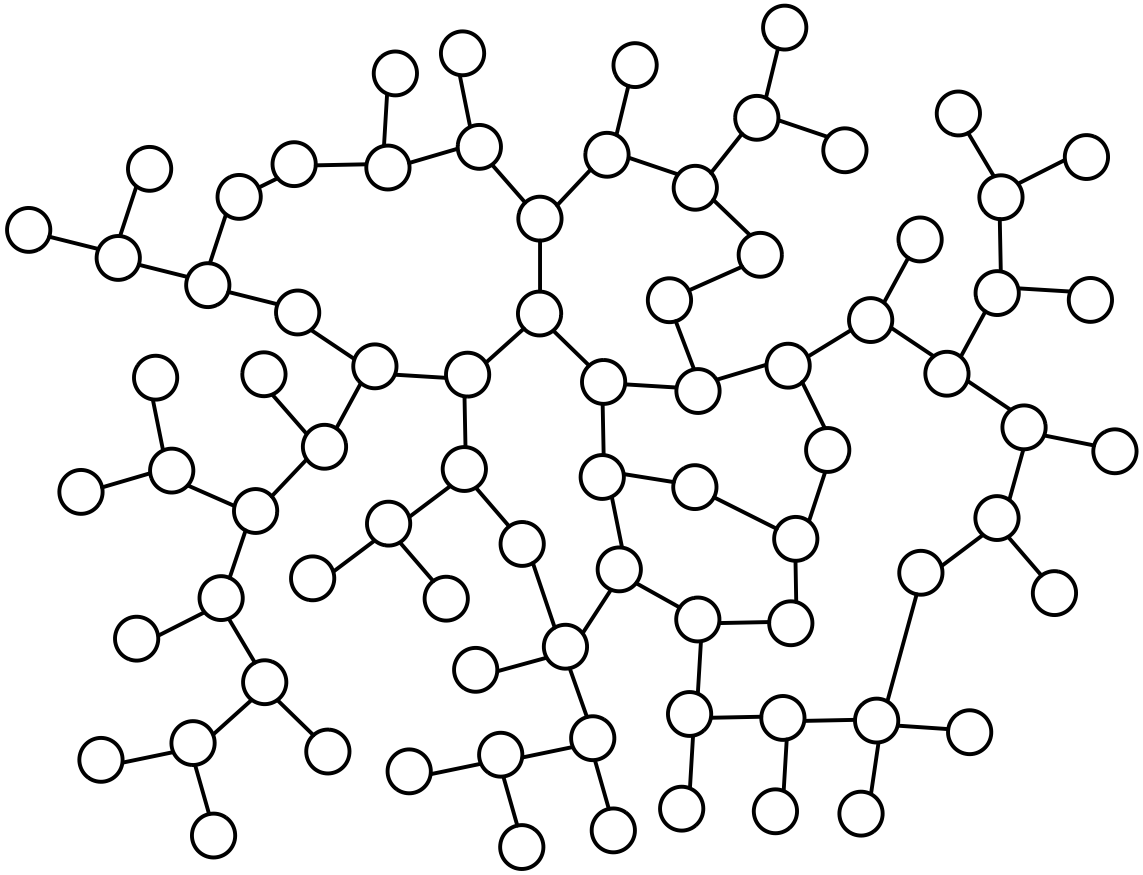
1. *We spread fast:* $N_T \approx (d-1)^{\frac{T}{2}}$
2. *All nodes except for the final virtual source are equally likely to be the source, hence*

$$P(\hat{v}_{ML} = v^*) = \frac{1}{N_T - 1}$$

3. *The expected distance between the estimated and true source is at least* $\frac{T}{2}$.

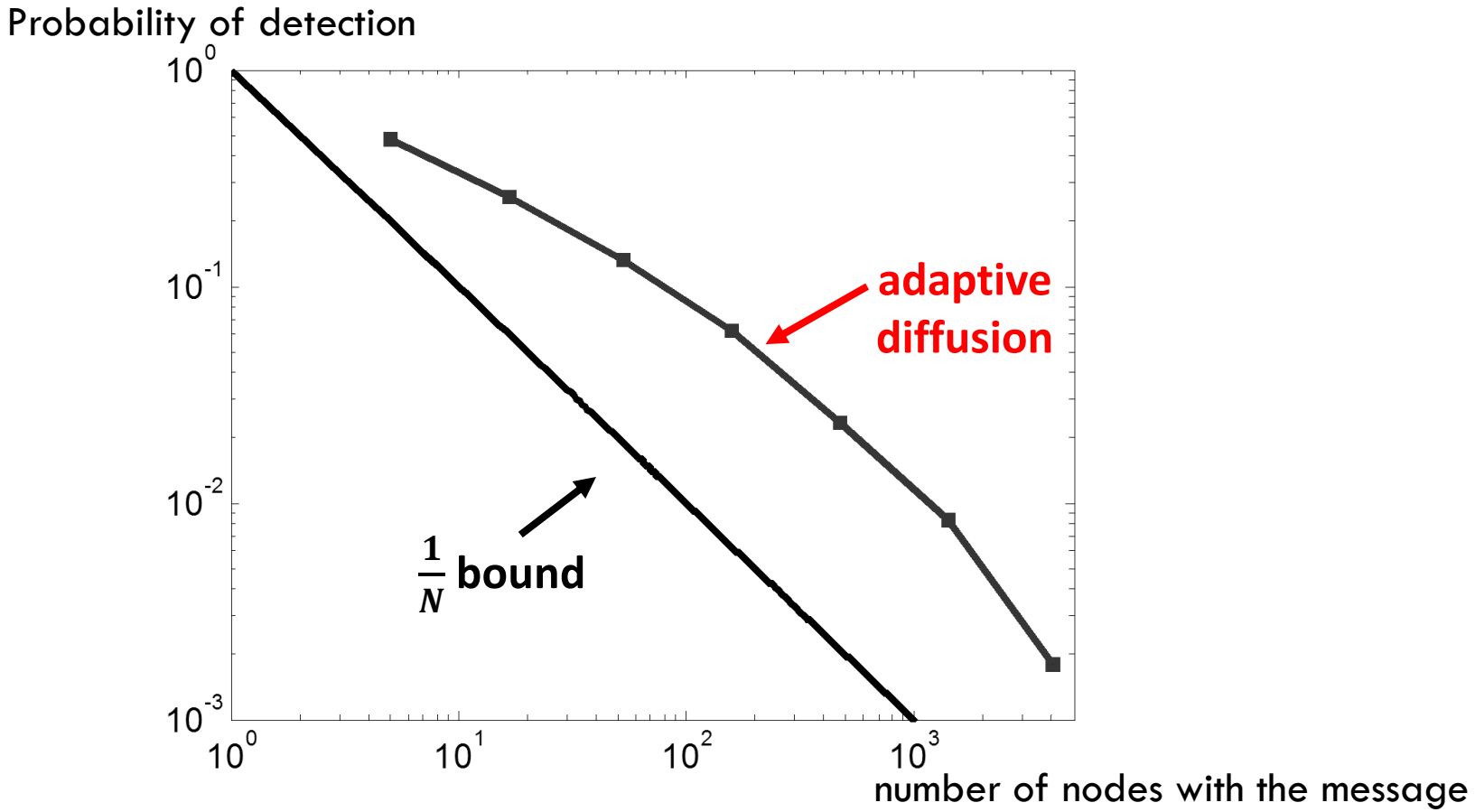
[Best Paper Award at SIGMETRICS 15]

General graphs



can we extend adaptive diffusion for general graphs?

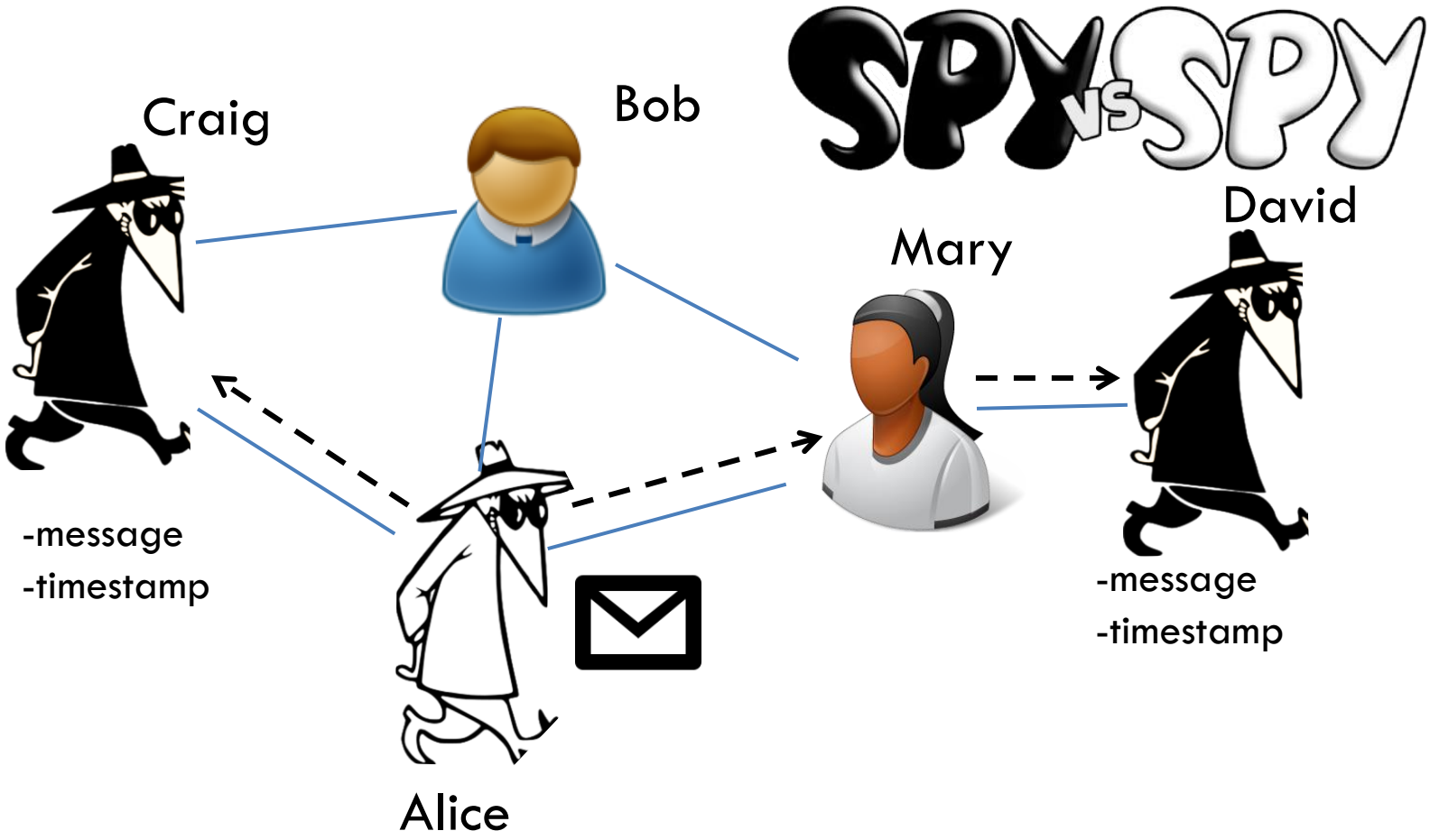
Simulation results: Facebook graph



- likelihoods can be **approximated** numerically

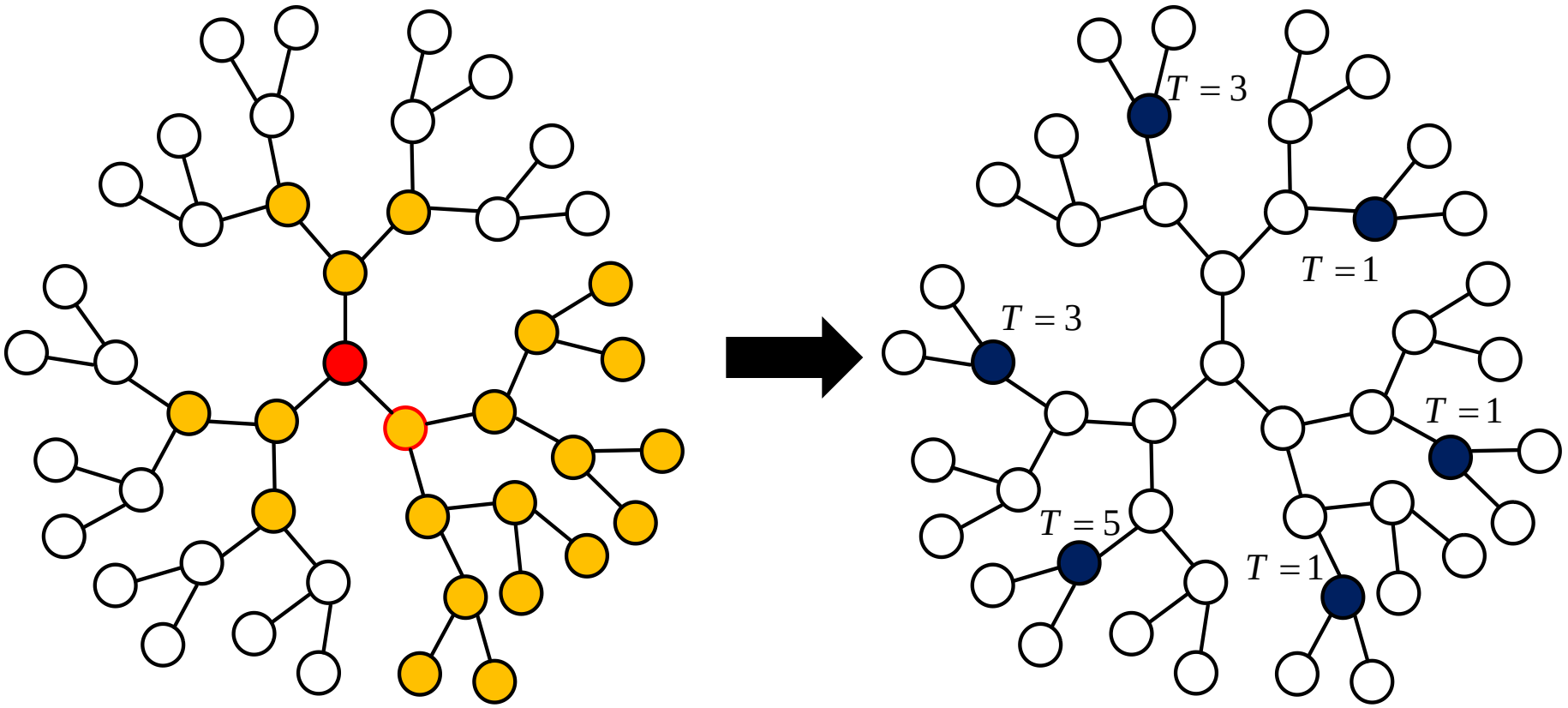
Part 1.2:
Adversary with timing info

Spy adversarial model

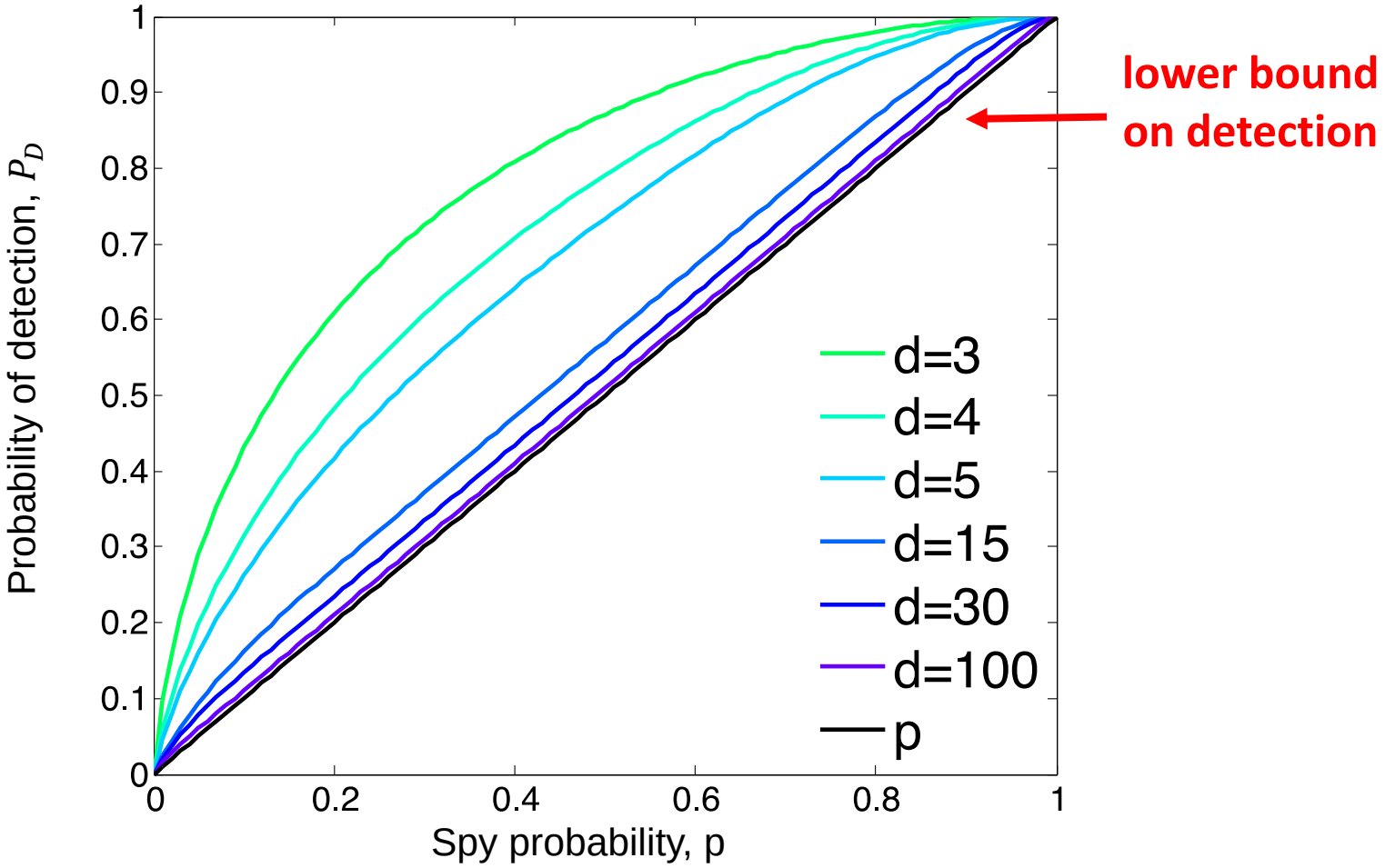


adversary can collect timing information

Adversary with timing

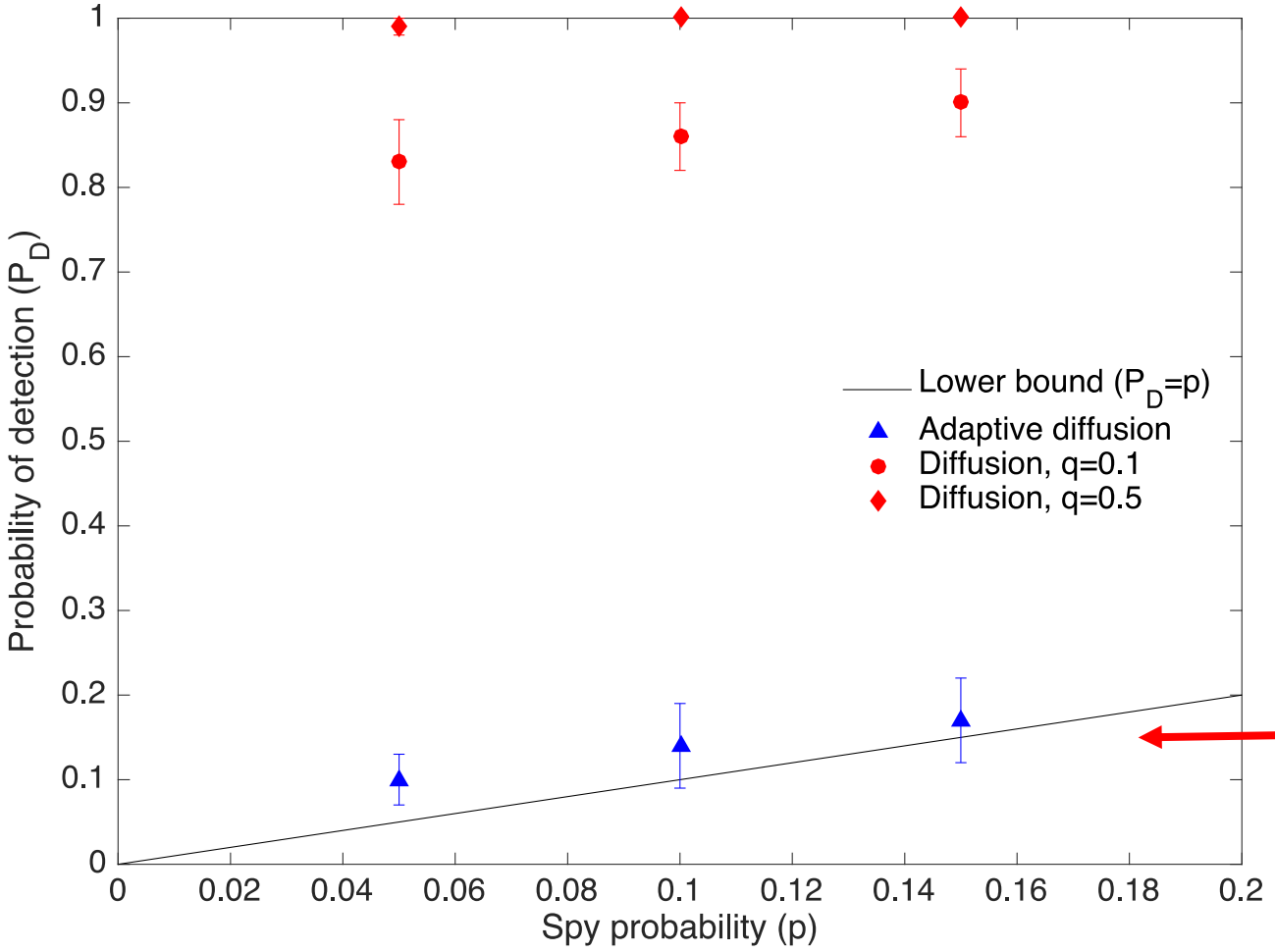


Results on d-regular trees



MAIN RESULT: Probability of detection $= p + o(p)$

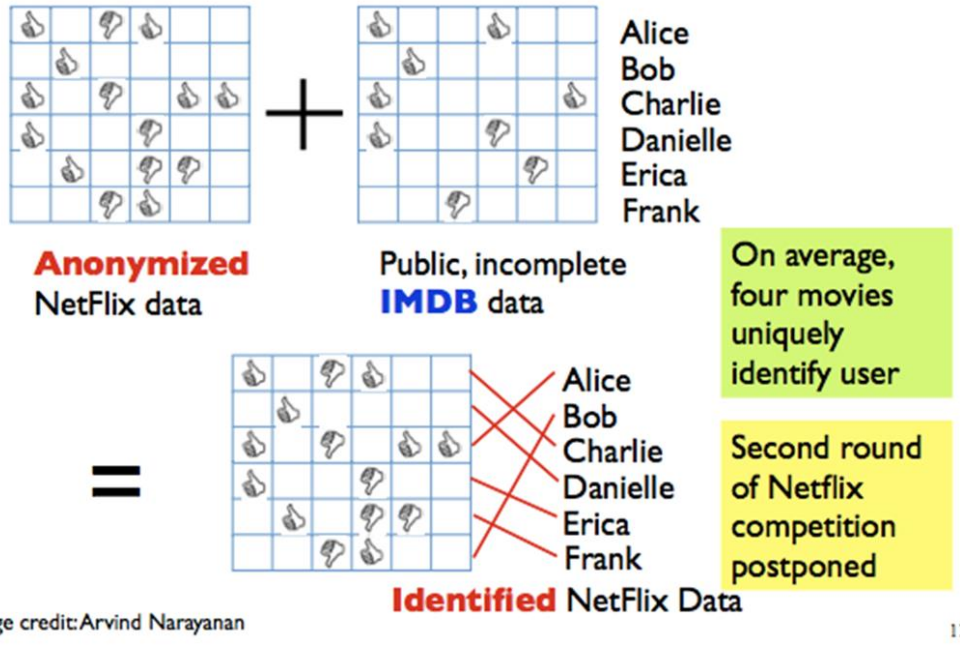
What about general graphs?

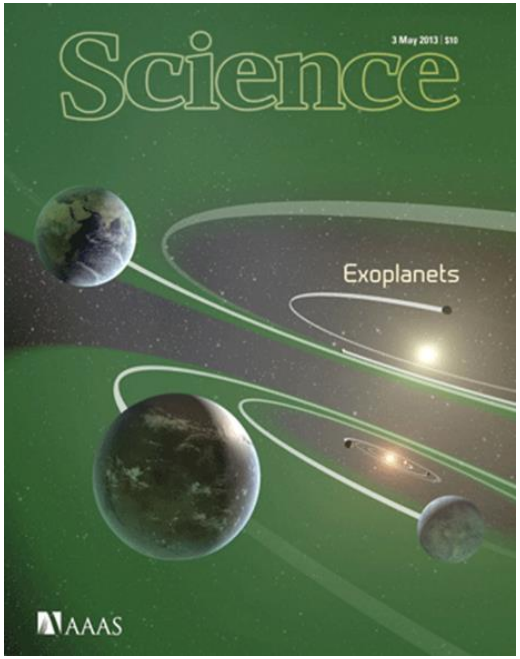


lower bound on detection

Part 2:
Data Privacy

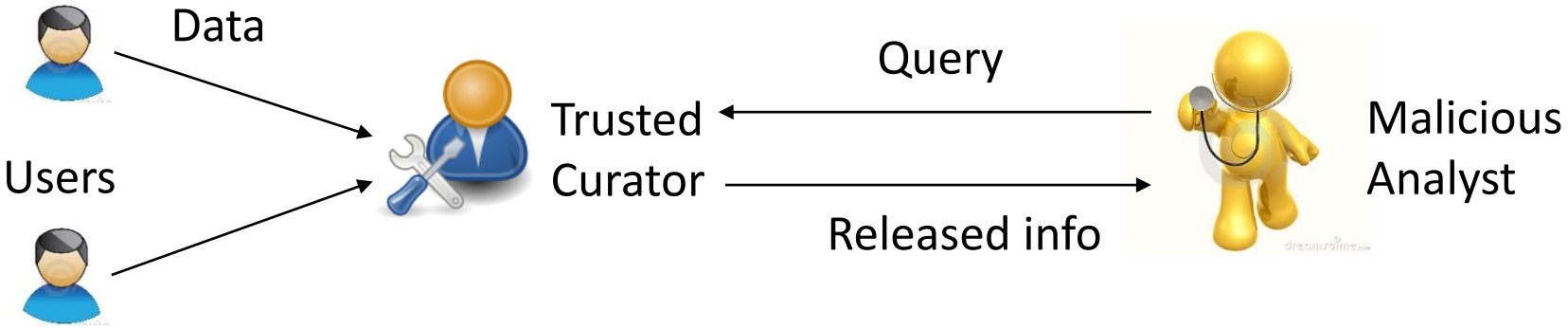
Recent data privacy leaks

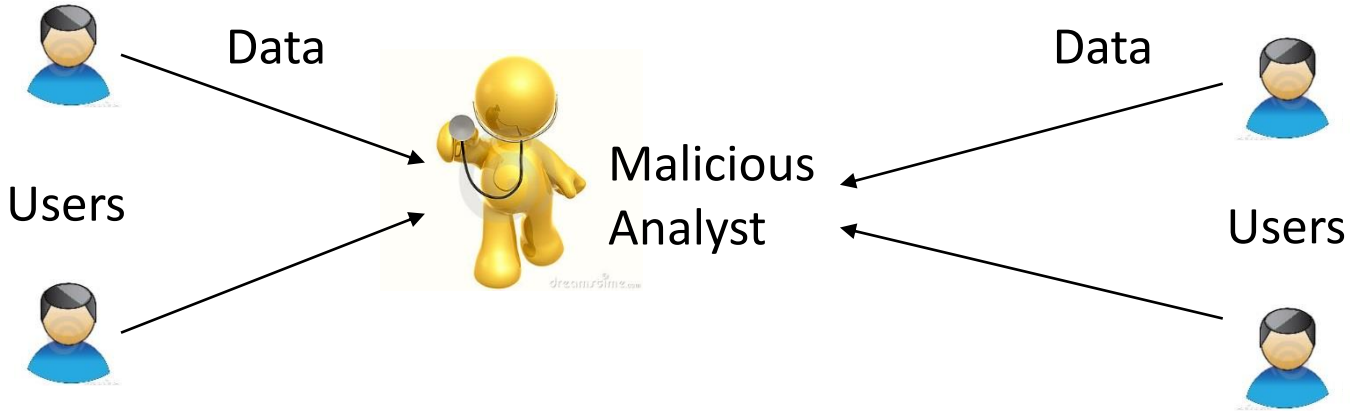




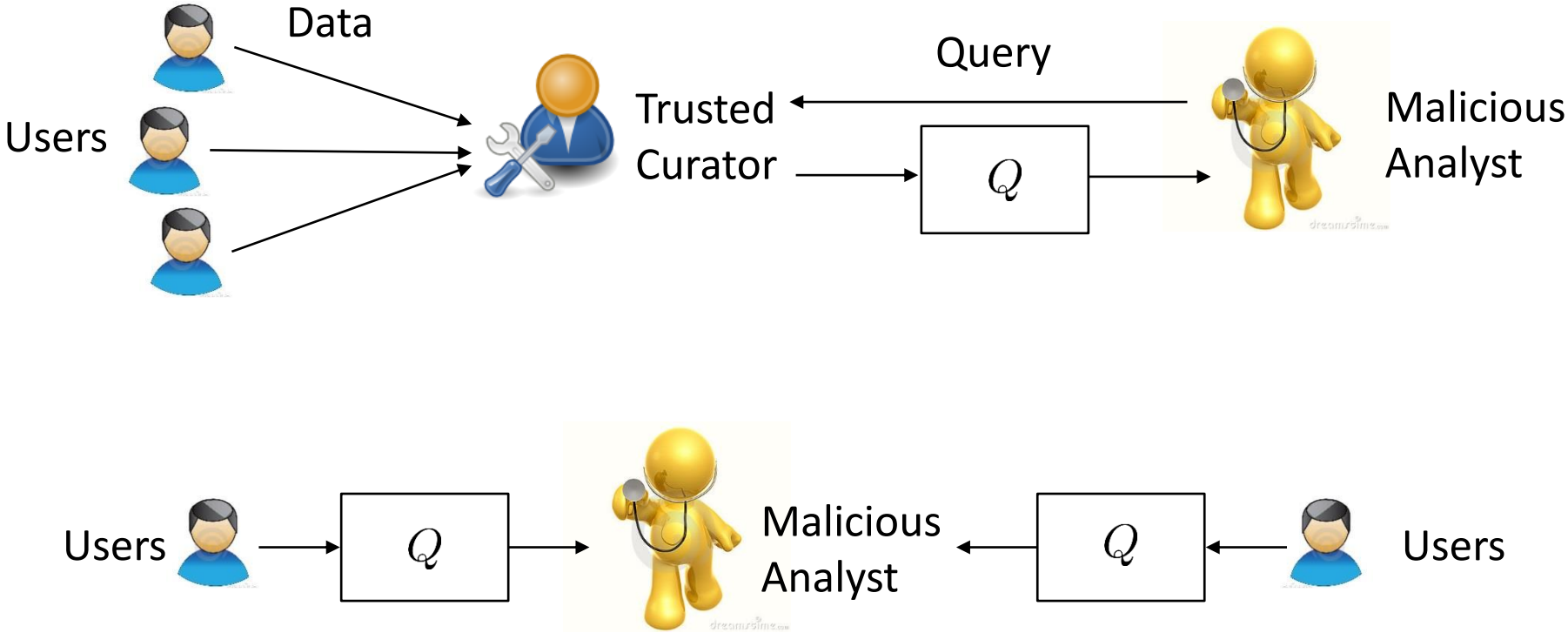
deanonymizing Netflix data, **identifying** personal genomes, etc.

Global vs. local models





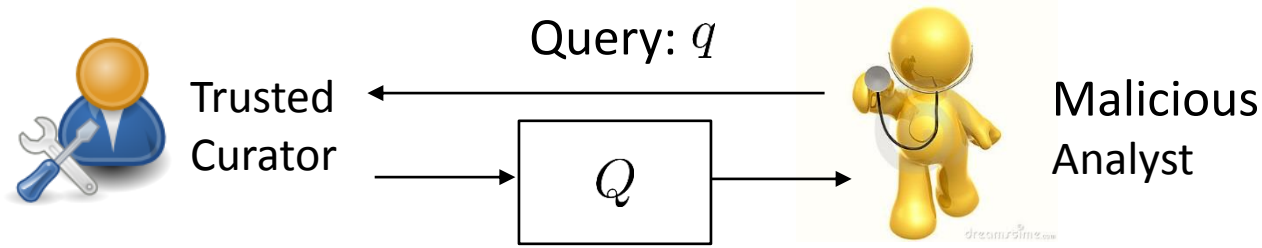
Global vs. local privacy



- Q is a privacy mechanism

Part 2.1:
Global Privacy

Global privacy



Database

Age
A: 20
B: 35
C: 43
D: 30

Query function: q

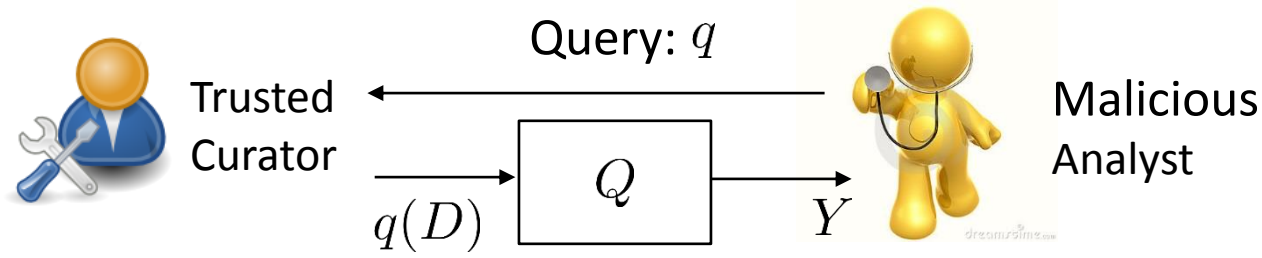
How many people are older than 32?

$q(D) = 2$

Privacy preserving query

say 0 w.p. $1/10$
say 1 w.p. $2/10$
say 2 w.p. $4/10$
say 3 w.p. $2/10$
say 4 w.p. $1/10$

Global privacy



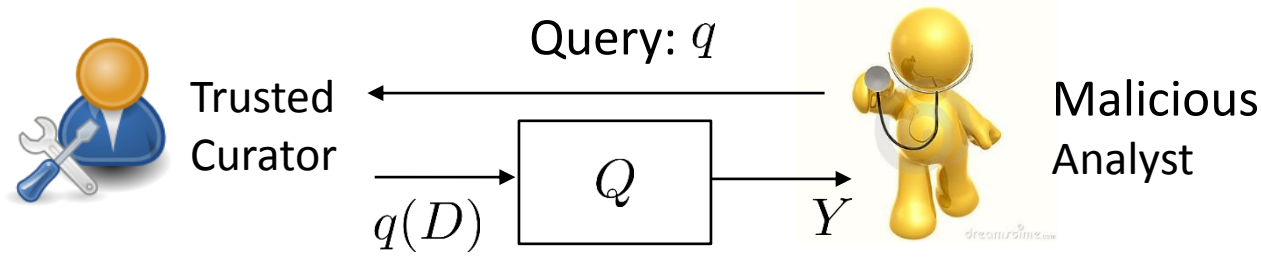
Database

Age
A: 20
B: 35
C: 43
D: 30

Privacy preserving mechanism

Given that the query output is $q(D)$,
say Y with probability $Q\left(Y|q(D)\right)$

Global differential privacy



Dataset: D_1

Age
A: 20
B: 35
C: 43
D: 30

Dataset: D_2

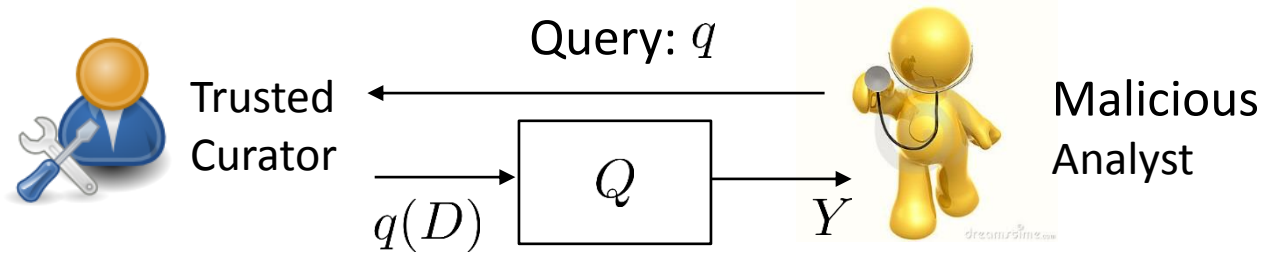
Age
A: 20
B: 35
D: 30

← **neighboring datasets**

$$Q\left(Y|q(D_1)\right) \approx Q\left(Y|q(D_2)\right)$$

- **presence** or **absence** of any individual should not have a significant impact on Y

Global differential privacy



Q is ε differentially private if $\forall$ neighboring datasets D_1, D_2 and all Y

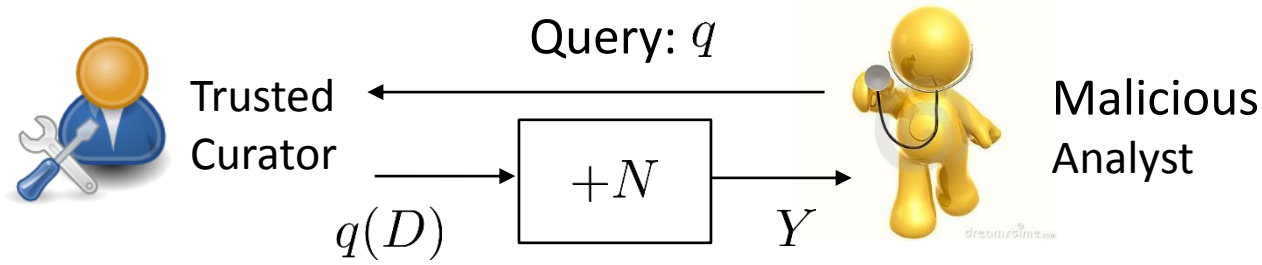
$$e^{-\varepsilon} \leq \frac{Q(Y|q(D_1))}{Q(Y|q(D_2))} \leq e^{+\varepsilon}$$

ε controls the level of privacy

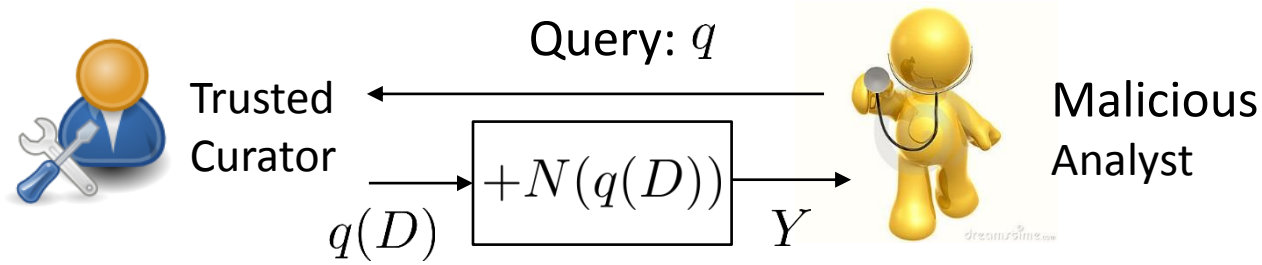
large ε, low privacy

small ε, high privacy

Two types of privacy mechanisms

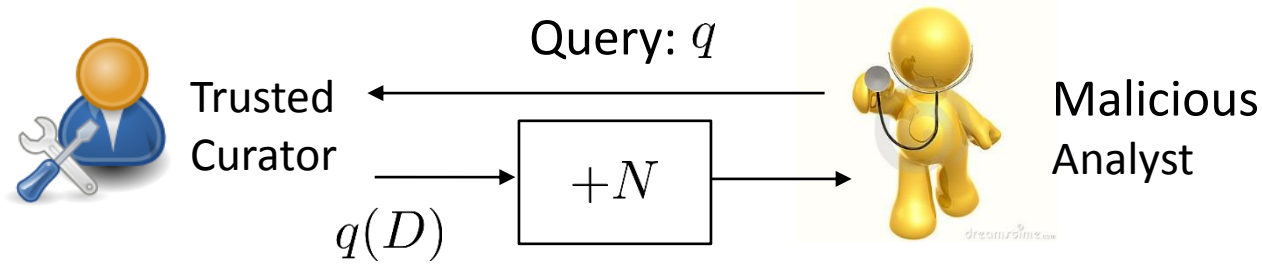


data independent privacy mechanisms : $Y = q(D) + N$

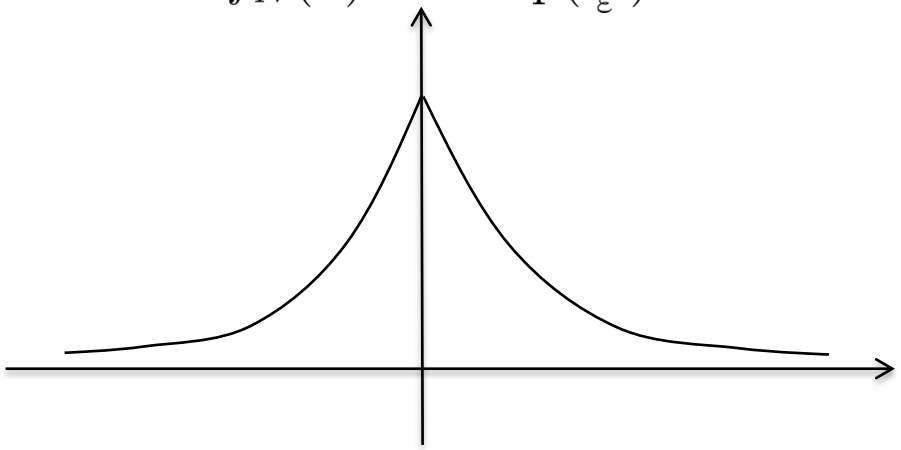


data dependent privacy mechanisms : $Y = q(D) + N(q(D))$

The Laplace mechanism



$$f_N(x) = \mathrm{Lap}(\tfrac{\Delta}{\varepsilon})$$



$$\Delta = \max_{D_1, D_1 \text{neighbors}} |q(D_1) - q(D_2)|$$

Privacy-utility tradeoff

- measuring the performance of a privacy mechanism

loss functions

$$L(x-y) = |y-x|$$

$$L(x-y) = (y-x)^2$$

true query output $q(D) = x$

noised query output $q(D) + N(q(D)) = y$

average loss when $q(D) = x$ **:** $C(Q,x) = \int L(x-y)Q(y|x)dy$

- loss functions considered are symmetric and increase with $|x-y|$

Privacy-utility tradeoff

- the **more private** you want to be, the **less utility** you get
- there is a **fundamental tradeoff** between **privacy** and **utility**

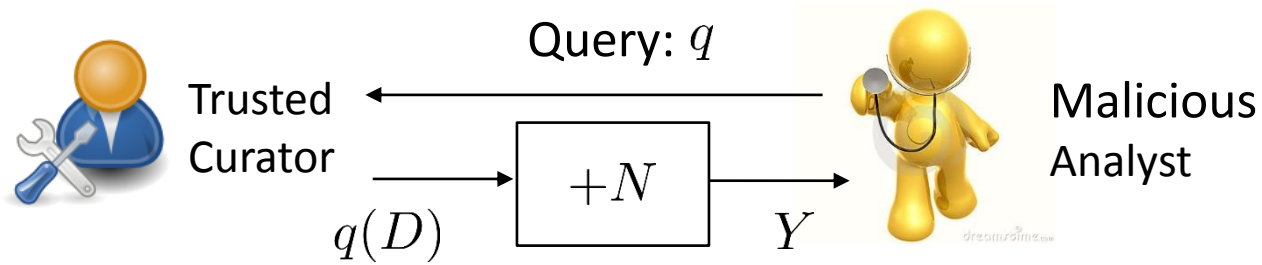
worst case loss

$$\begin{aligned} \underset{Q}{\text{minimize}} \quad & \sup_x C(Q, x) \\ \text{subject to} \quad & Q \in \mathcal{D}_\varepsilon \end{aligned}$$

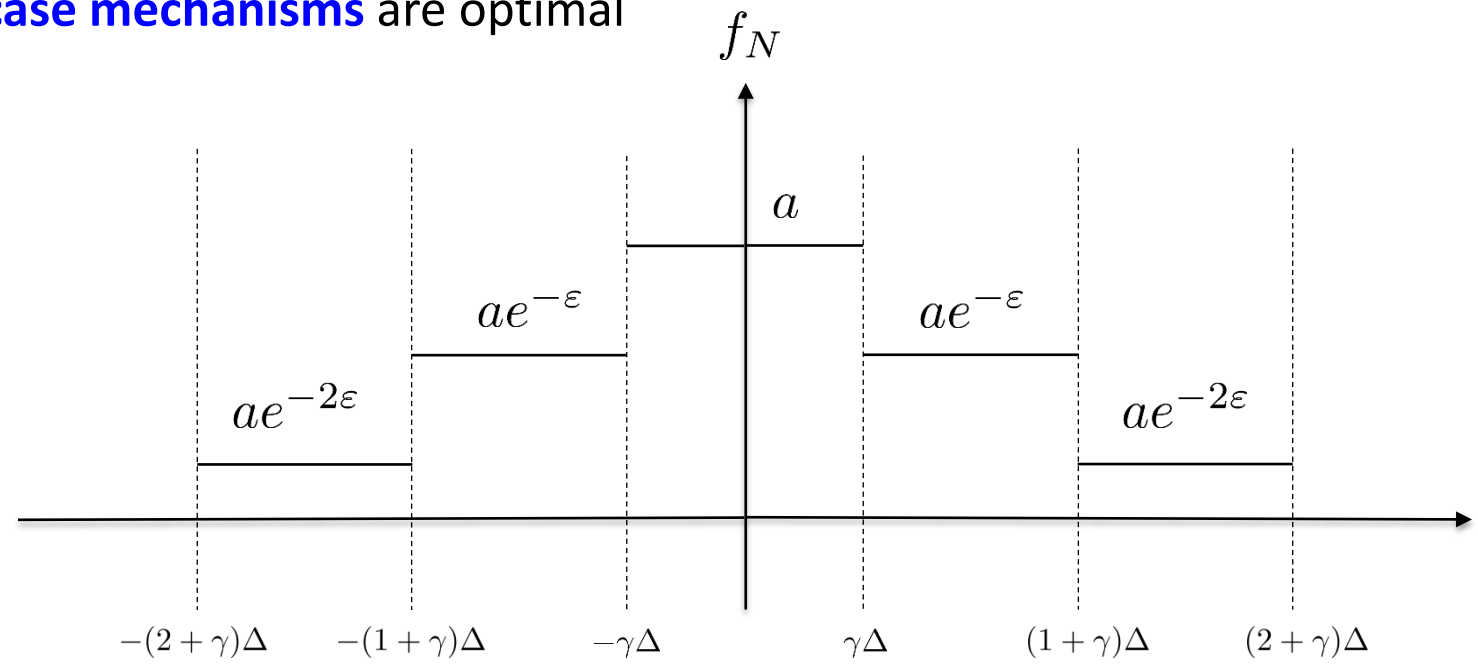
set of all differentially private mechanisms

Optimality of staircase mechanisms

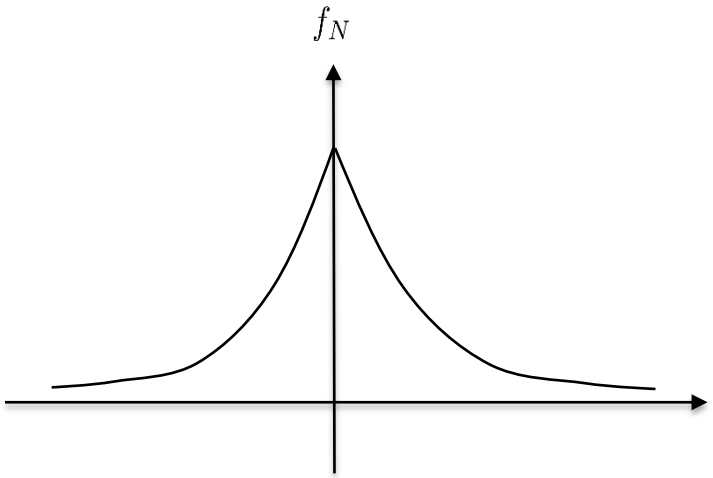
- **data independent mechanisms** are optimal

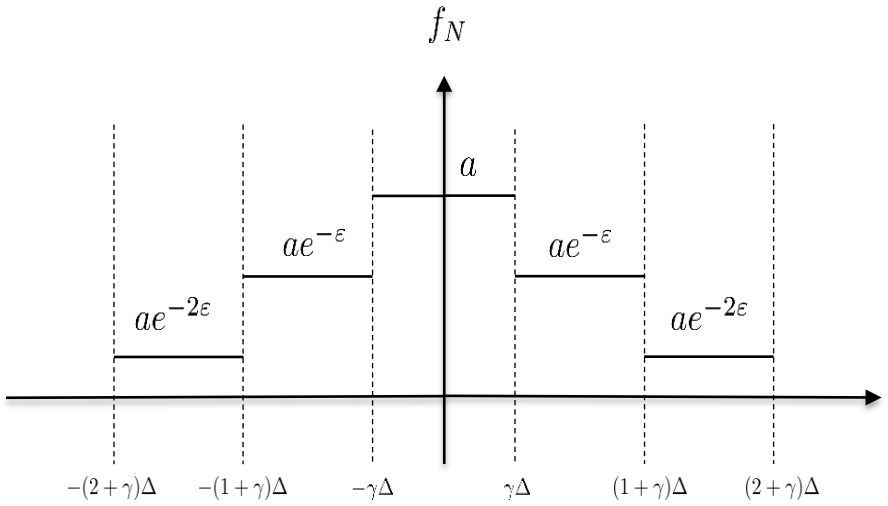


- **staircase mechanisms** are optimal



Staircase mechanisms





$\forall$ neighboring datasets D_1, D_2

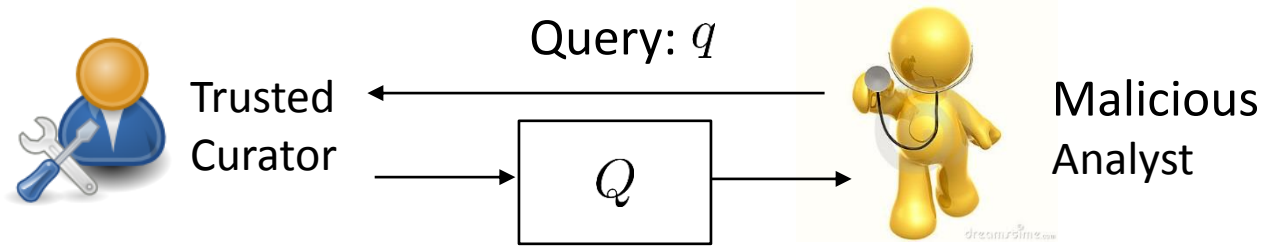
differentially private mechanisms

$$e^{-\varepsilon} \leq \frac{Q(Y|q(D_1))}{Q(Y|q(D_2))} \leq e^{\varepsilon}$$

staircase mechanisms

$$\frac{Q(Y|q(D_1))}{Q(Y|q(D_2))} \in \{e^{-\varepsilon}, 1, e^{+\varepsilon}\}$$

Multidimensional queries



$$q(D) = (q_1(D), ..., q_d(D)) \qquad Y = (Y_1, ..., Y_d)$$

Loss functions:

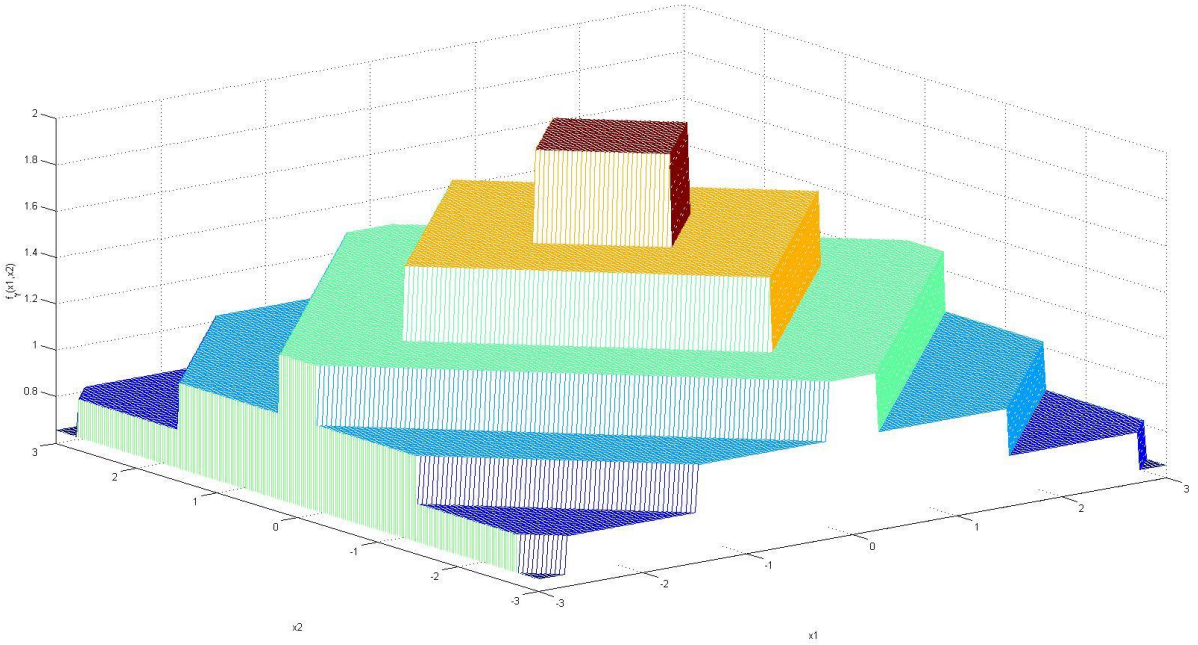
$$L(Y - q(D)) = ||Y - q(D)||_1$$

$$L(Y - q(D)) = ||Y - q(D)||_2$$

average loss when $q(D) = x$**:** $\quad C(Q, x) = \int ... \int L(x - y)Q(y|x)dy$

$$\begin{aligned} \underset{Q}{\text{minimize}} \quad & \sup_x C(Q, x) \\ \text{subject to} \quad & Q \in \mathcal{D}_\varepsilon \end{aligned}$$

Optimality of staircase mechanisms



- the two dimensional **staircase mechanism** is **optimal** for 2 dimensional queries
- we conjecture that the same result holds for any number of dimensions

Part 2.2:
Local Privacy

Local privacy

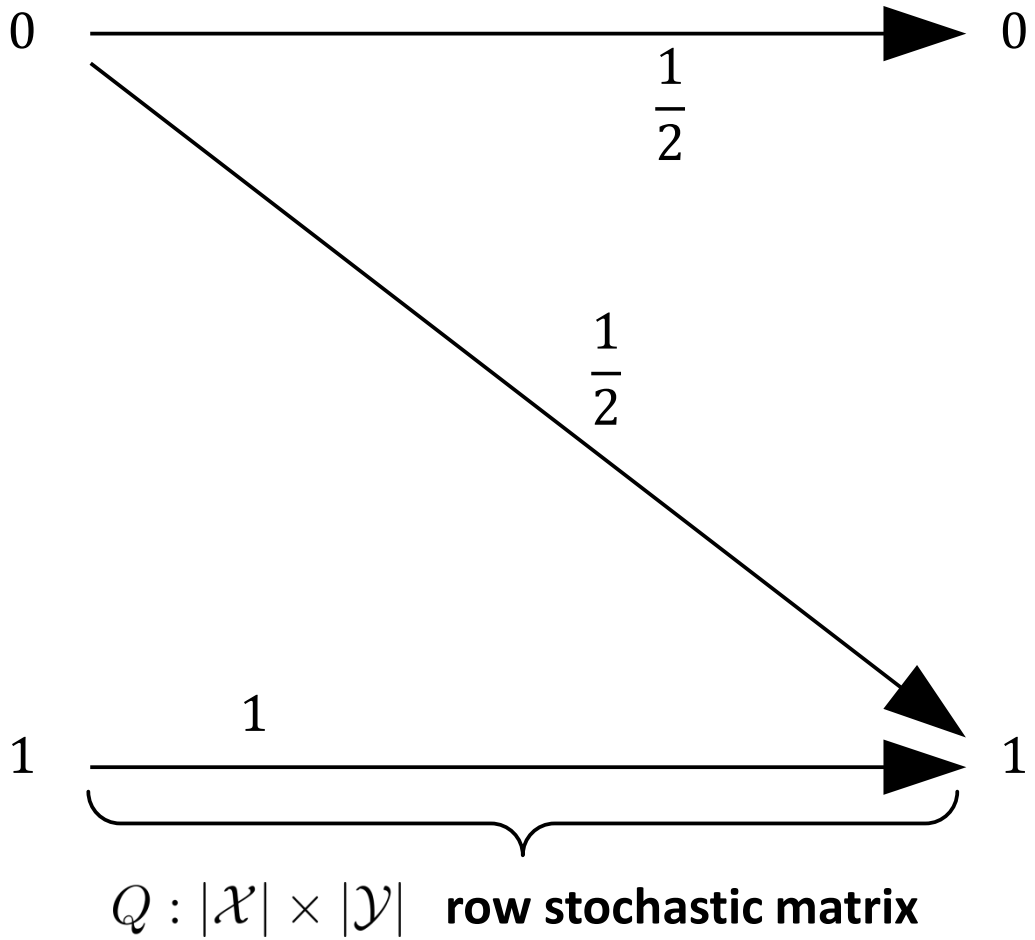
have you ever used illegal drugs?

say yes

answer truthfully

[*Warner* 1965]

Privacy via plausible deniability



- instead of $X = x$, share $Y = y$ w.p. $Q(y|x)$
- we focus on discrete alphabets

Local differential model

Q is ε differentially private if $\forall\ X, X'$, and Y

$$e^{-\varepsilon} \leq \frac{Q(Y|X)}{Q(Y|X')} \leq e^{+\varepsilon}$$

ε controls the level of privacy

large ε, low privacy

small ε, high privacy

[*Duchi, et. al.*, 2013]

Privacy-utility tradeoff

- the **more private** you want to be, the **less utility** you get
- there is a **fundamental tradeoff** between **privacy** and **utility**

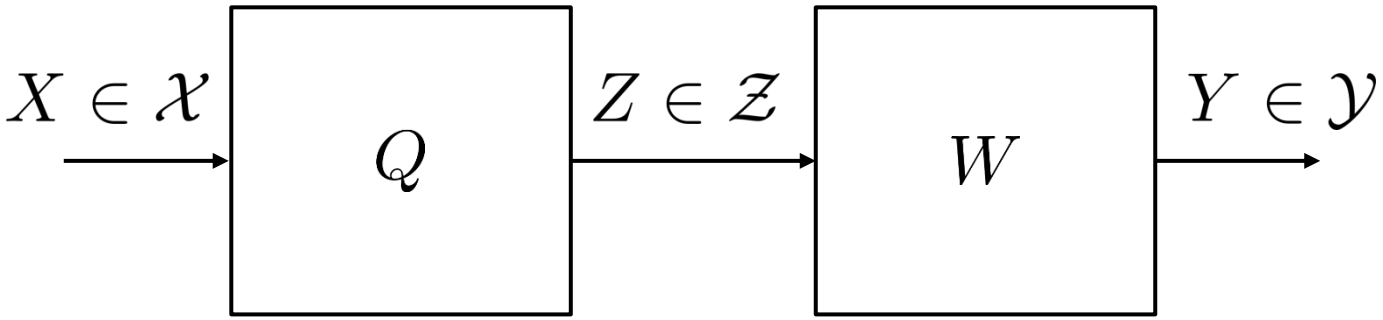
$$\begin{aligned} \underset{Q}{\text{maximize}} \quad & U(Q) \\ \text{subject to} \quad & Q \in \mathcal{D}_\varepsilon \end{aligned}$$

application dependent utility function

set of all differentially private mechanisms

Binary alphabets

Utility functions



utility functions obeying the **data processing inequality**:

$$T = Q \circ W \implies U(T) \leq U(Q)$$

- further randomization **can only reduce utility**

Main result: binary data

- for binary alphabets: $|\mathcal{X}| = 2$

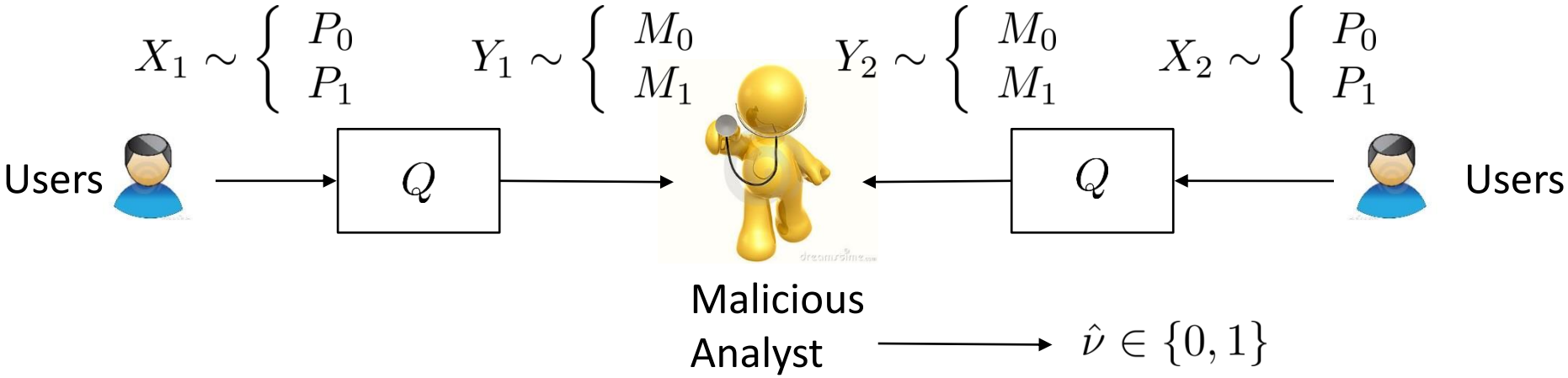
lie w.p. $\dfrac{1}{e^{\varepsilon}+1}$ **say the truth w.p.** $\dfrac{e^{\varepsilon}}{e^{\varepsilon}+1}$

- optimal for **all values of** ε
- optimal for **all** $U(Q)$ **obeying the data processing inequality**

[*JMLR 2015*]

General alphabets

Private hypothesis testing



$$\underset{Q}{\text{maximize}} \quad D_f(\underbrace{QP_0}_{M_0} \,\|\, \underbrace{QP_1}_{M_1})$$

$$\text{subject to} \quad Q \in \mathcal{D}_\varepsilon$$

f-divergences include KL and total variation distances

Staircase mechanisms

Q is ε differentially private if $\forall x, x' \in \mathcal{X}$ and all $y \in \mathcal{Y}$

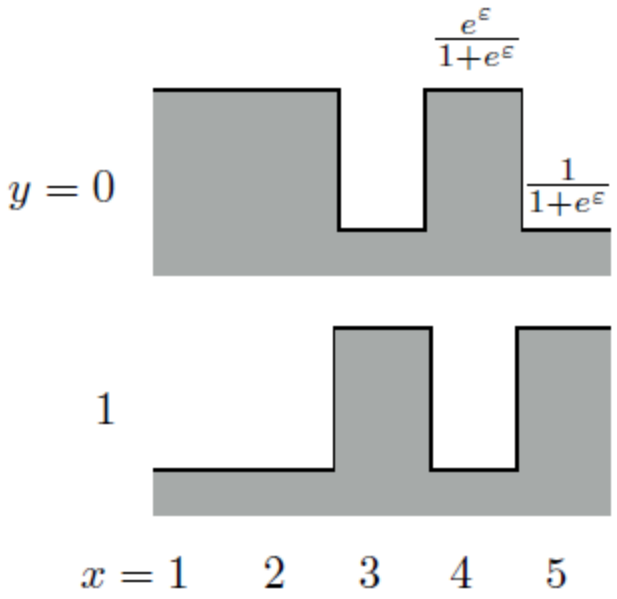
$$e^{-\varepsilon} \leq \frac{Q(y|x)}{Q(y|x')} \leq e^{\varepsilon}$$

Q is a staircase mechanism if $\forall x, x' \in \mathcal{X}$ and all $y \in \mathcal{Y}$

$$\frac{Q(y|x)}{Q(y|x')} \in \{e^{-\varepsilon}, 1, e^{+\varepsilon}\}$$

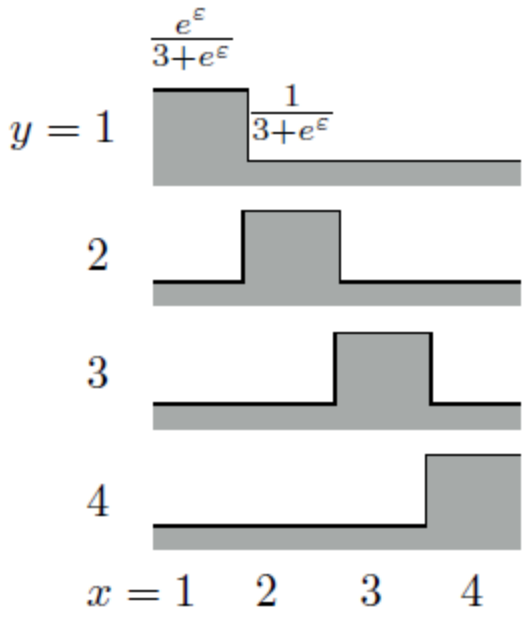
Example of staircase mechanisms

$$Q^T = \frac{1}{1+e^\varepsilon}\begin{bmatrix} e^\varepsilon & e^\varepsilon & 1 & e^\varepsilon & 1 \\ 1 & 1 & e^\varepsilon & 1 & e^\varepsilon \end{bmatrix}$$



Binary Mechanism

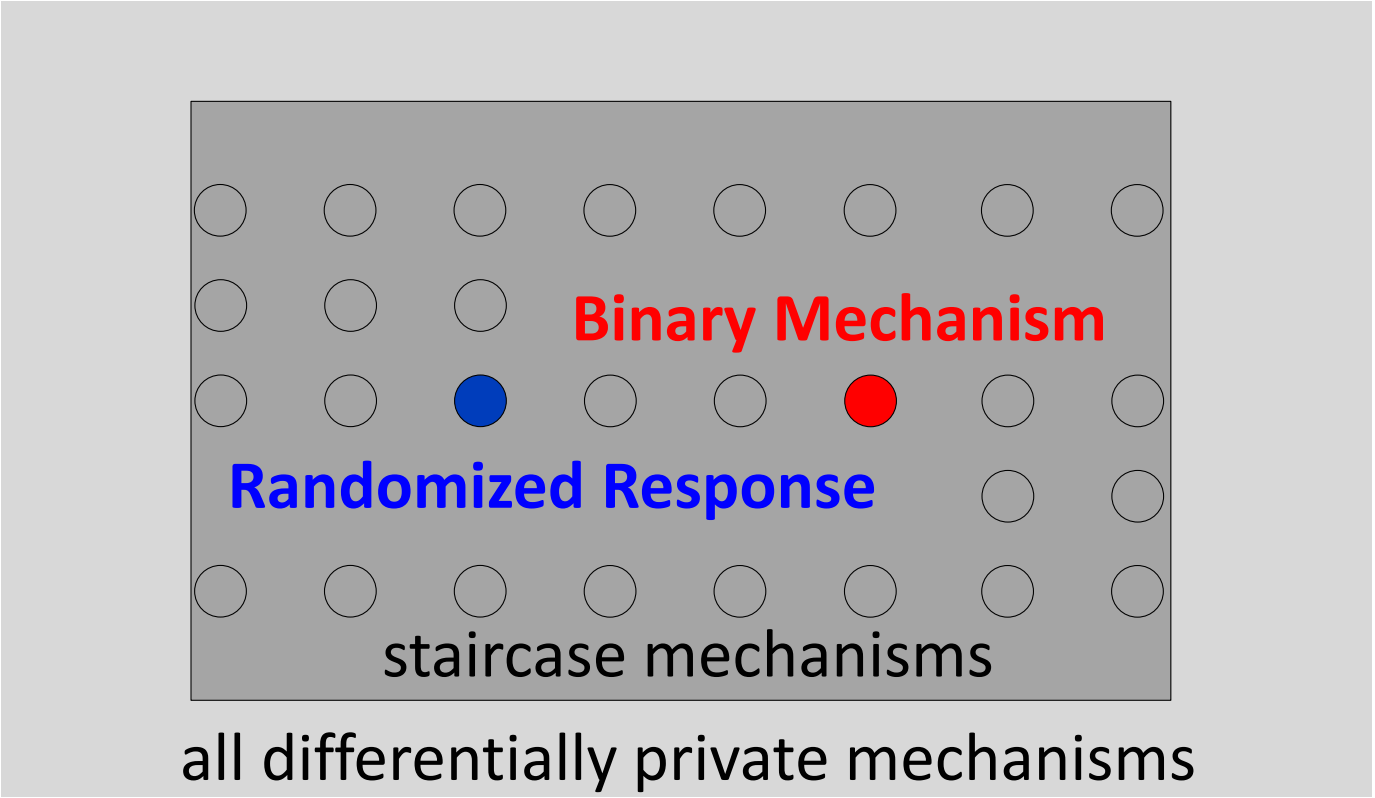
$$Q^T = \frac{1}{3+e^\varepsilon}\begin{bmatrix} e^\varepsilon & 1 & 1 & 1 \\ 1 & e^\varepsilon & 1 & 1 \\ 1 & 1 & e^\varepsilon & 1 \\ 1 & 1 & 1 & e^\varepsilon \end{bmatrix}$$



Randomized Response

Main result: general case

- for $|\mathcal{X}| = k > 2$



- staircase mechanisms are optimal for **all ε**
- **BM optimal for small ε**
- **RR optimal for large ε**

[NIPS 14, *JMLR 15*]

More general utility functions

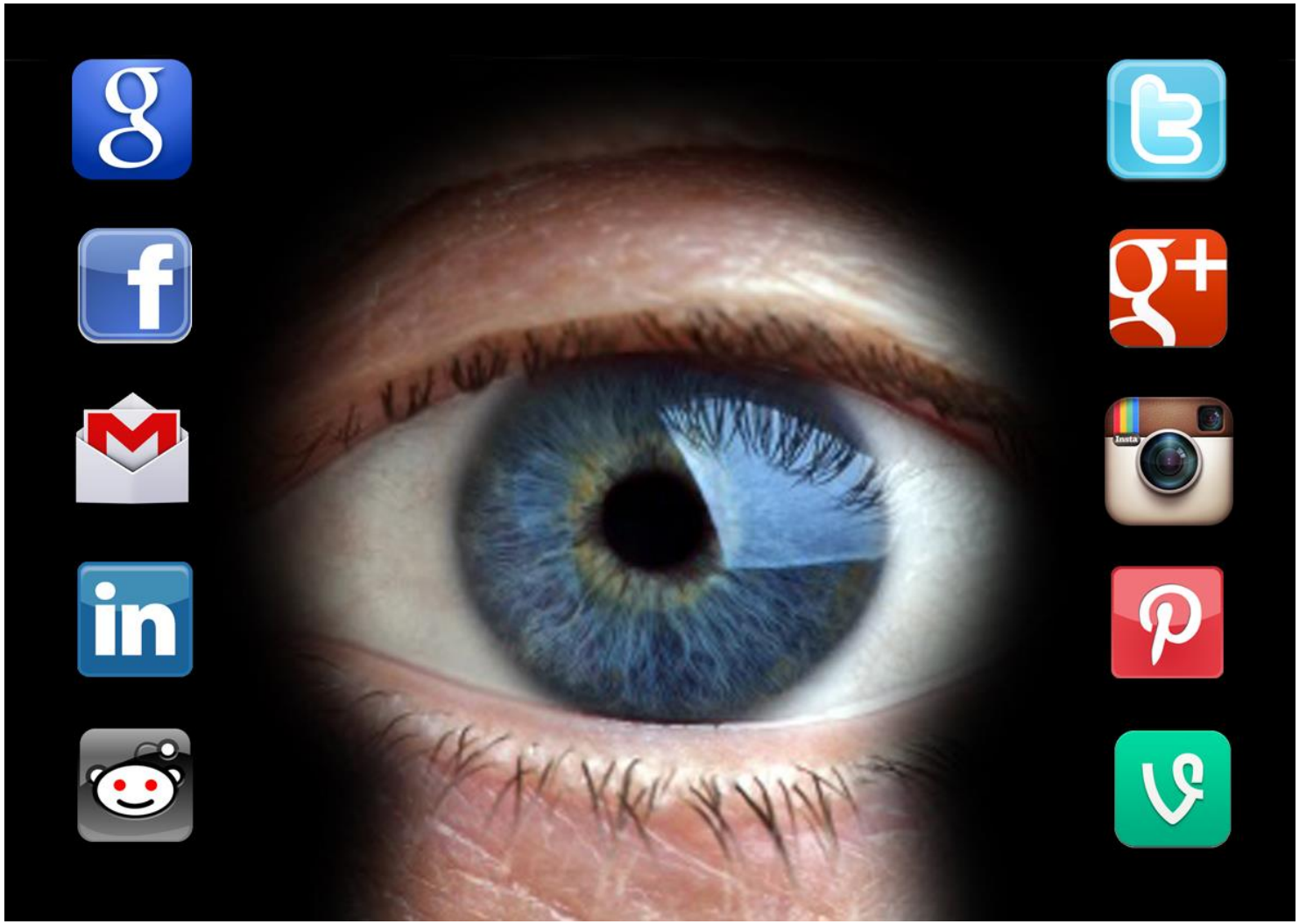
- same results hold for a rich class of **convex utility functions:**

$$\begin{aligned} \underset{Q}{\text{maximize}} \quad & U(Q) = \sum_{y \in \mathcal{Y}} \mu\left(Q_y\right) \\ \text{subject to} \quad & Q \in \mathcal{D}_\varepsilon \end{aligned}$$

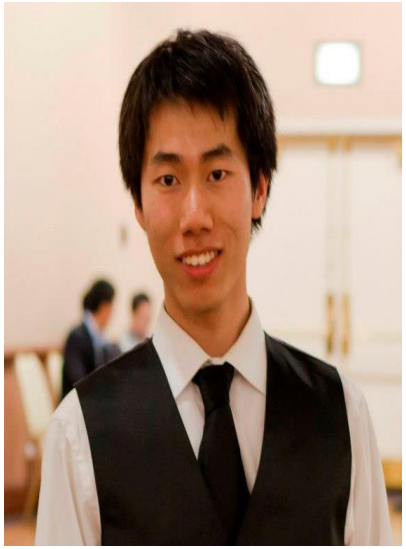
Q_y : the column of Q corresponding to $Q(y|.)$
μ : any sub-linear function

includes all f-divergences and mutual information

Rise of the planet of the apps!



Acknowledgments



Quan Geng

Giulia Fanti

Pramod Viswanath

Sewoong Oh