THE FUNDAMENTAL LIMITS OF DATA & METADATA PRIVACY

Peter Kairouz

ECE Department

University of Illinois at Urbana Champaign

30 years ago…

Pre-internet

Human to human

Then came the Internet

Pre-internet → **Internet of content**

Human to human

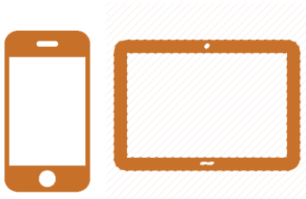
World Wide Web

Google

YAHOO!

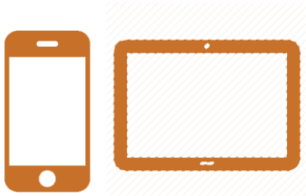
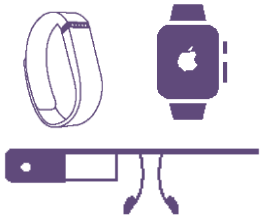
And then the Internet became smarter

Pre-internet → **Internet of content** → **Internet of services** → **Internet of people**

Human to human	*World Wide Web*	*Smart phones and tablets*	*Social Media*
	Google YAHOO!		

Unprecedented level of connectivity

Pre-internet → **Internet of content** → **Internet of services** → **Internet of people** → **Internet of things**

Human to human	*World Wide Web*	*Smart phones and tablets*	*Social Media*	*Machine to machine*
	Google YAHOO!			

WE'RE BEING WATCHED!

Pre-internet | Internet of content | Internet of services | Internet of people | Internet of things

It's okay, our data is encrypted

iCloud

CHASE

National Institutes of Health

NETFLIX

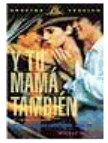
Other Movies You Might Enjoy

Amelie

Add

Not Interested

Y Tu Mama Tambien



Add

Not Interested

Eiken has been added to your Queue at position 2.

This movie is available now.

Move To Top Of My Queue

< Continue Browsing

Visit your Queue >

Guys and Balls

Add

Not Interested

Mostly Martha

Add

Not Interested

Only Human

Add

Not Interested

Russian Dolls

Add

Not Interested

Data privacy

NETFLIX

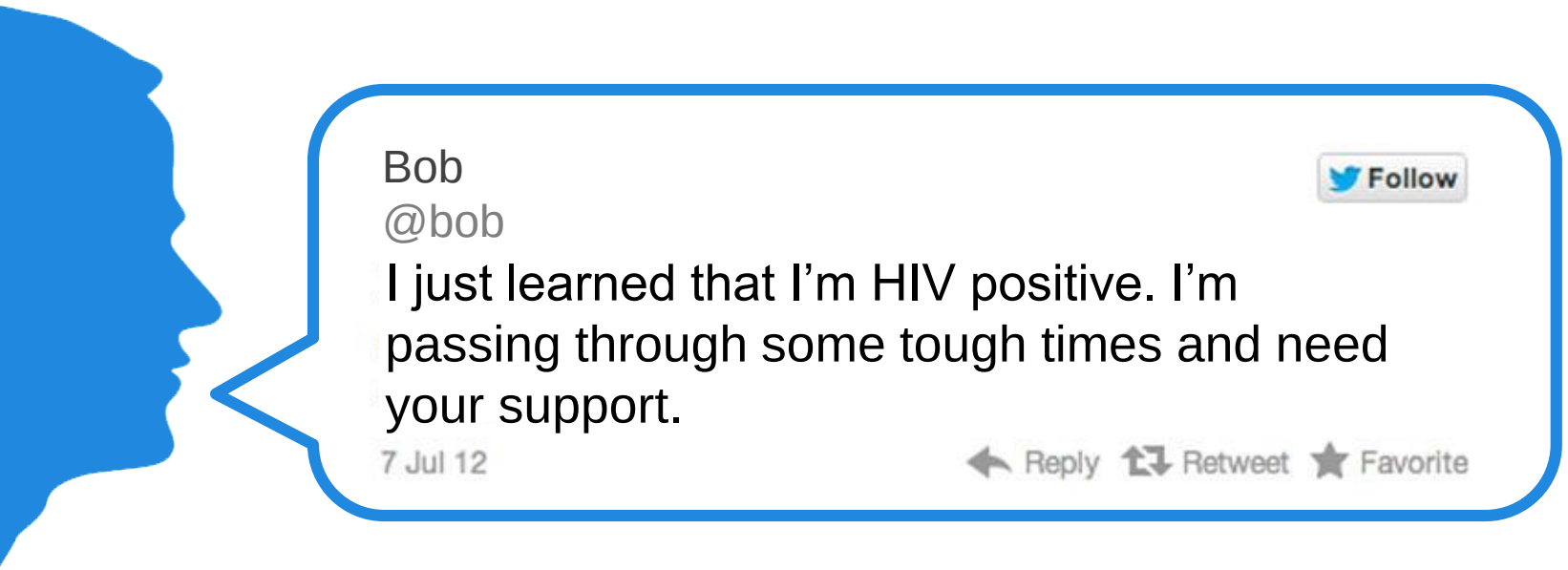
de-anonymizing Netflix watch histories

Science AAAS

identifying surnames and ages from anonymized genomes

from anonymous faces to social security numbers

Metadata privacy



Jason Rezaian's Year of Imprisonment in Iran

Wednesday marks the one-year anniversary of the *Washington Post* reporter's detention in the Islamic Republic.

Russian Activists and Journalists Attacked at Chechen Border

Saudi Man Gets 10 Years, 2,000 Lashes Over Atheist Tweets

By THE ASSOCIATED PRESS ·
RIYADH, Saudi Arabia — Feb 27, 2016, 8:26 AM ET

Politics | Fri Nov 23, 2007 4:54pm EST

Related

Syria blocks Facebook in Internet crackdown

DAMASCUS | BY KHALED YACOUB OWEIS

Metadata and data privacy

a distributed, **anonymous messaging protocol**

[**Best Paper Award** at SIGMETRICS 15, SIGMETRICS 16]

[NIPS 14, NIPS 15, ICML 15, TSTSP 15, CISS 16, JMLR 16, TIT 16]

privacy mechanisms that achieve the optimal privacy-utility tradeoff

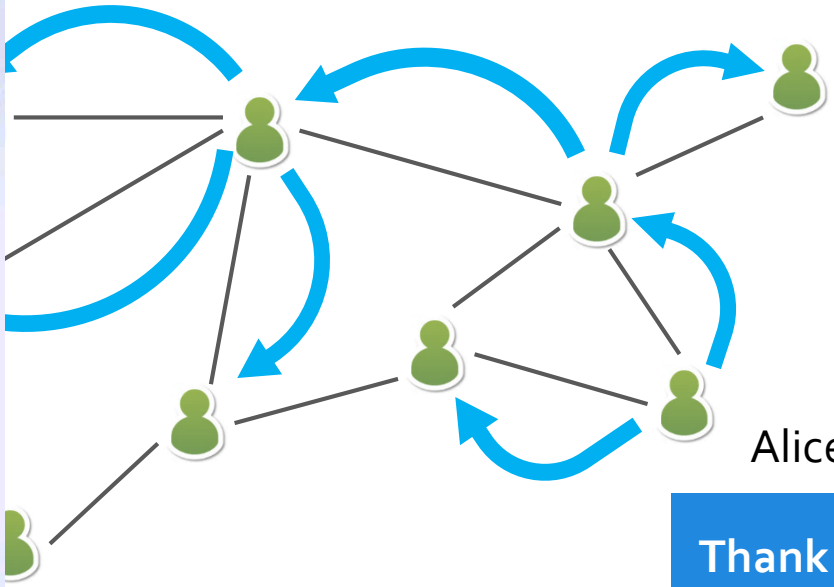
Part 1: Metadata privacy

Anonymous messaging

whisper

secret

Yik Yak



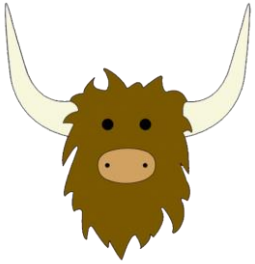
Alice

Thank you to all the blood donors…

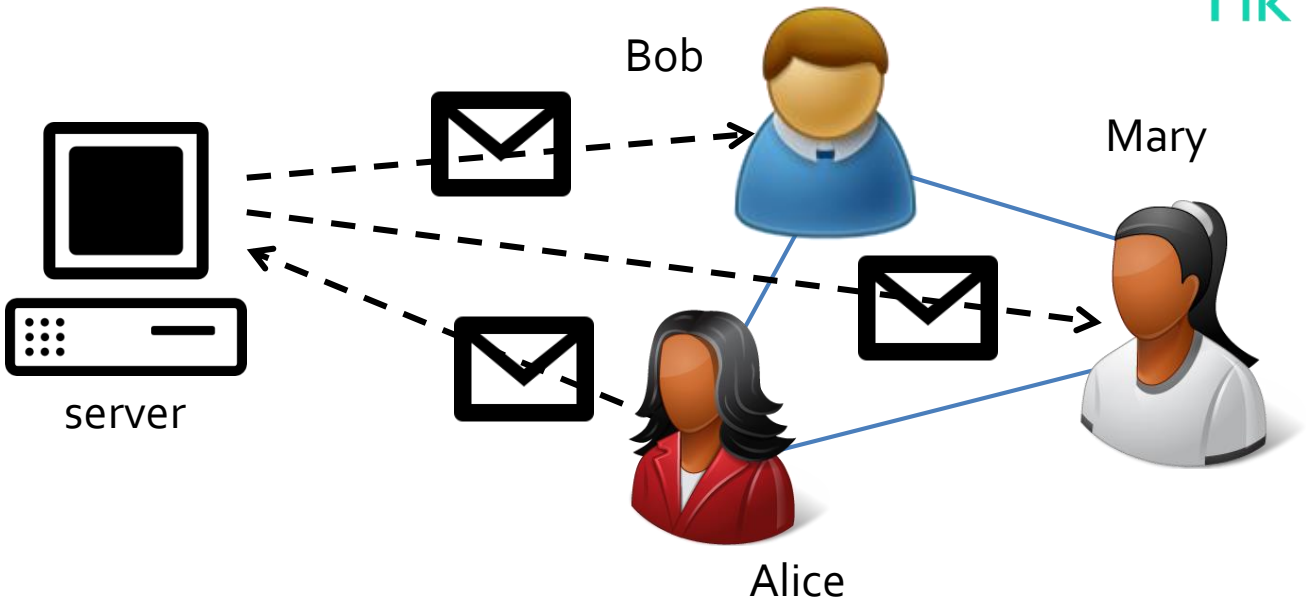
What's wrong with this

whisper

secret



Yik Yak



centralized networks **are not** truly anonymous!

Objective

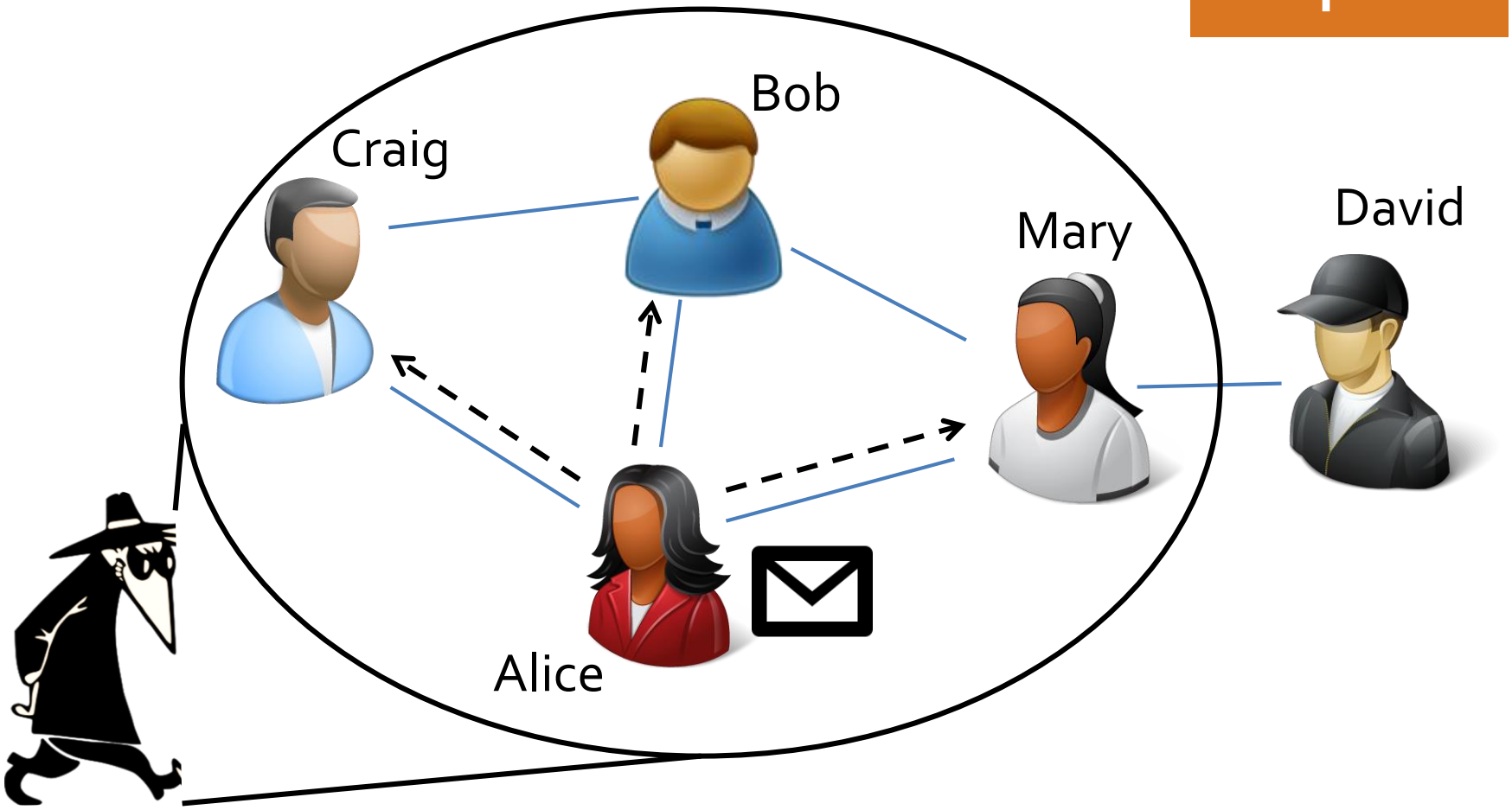
design a decentralized messaging protocol that:

(a) spreads message **fast**

(b) gives authors **anonymity guarantees**

Adversary without timing

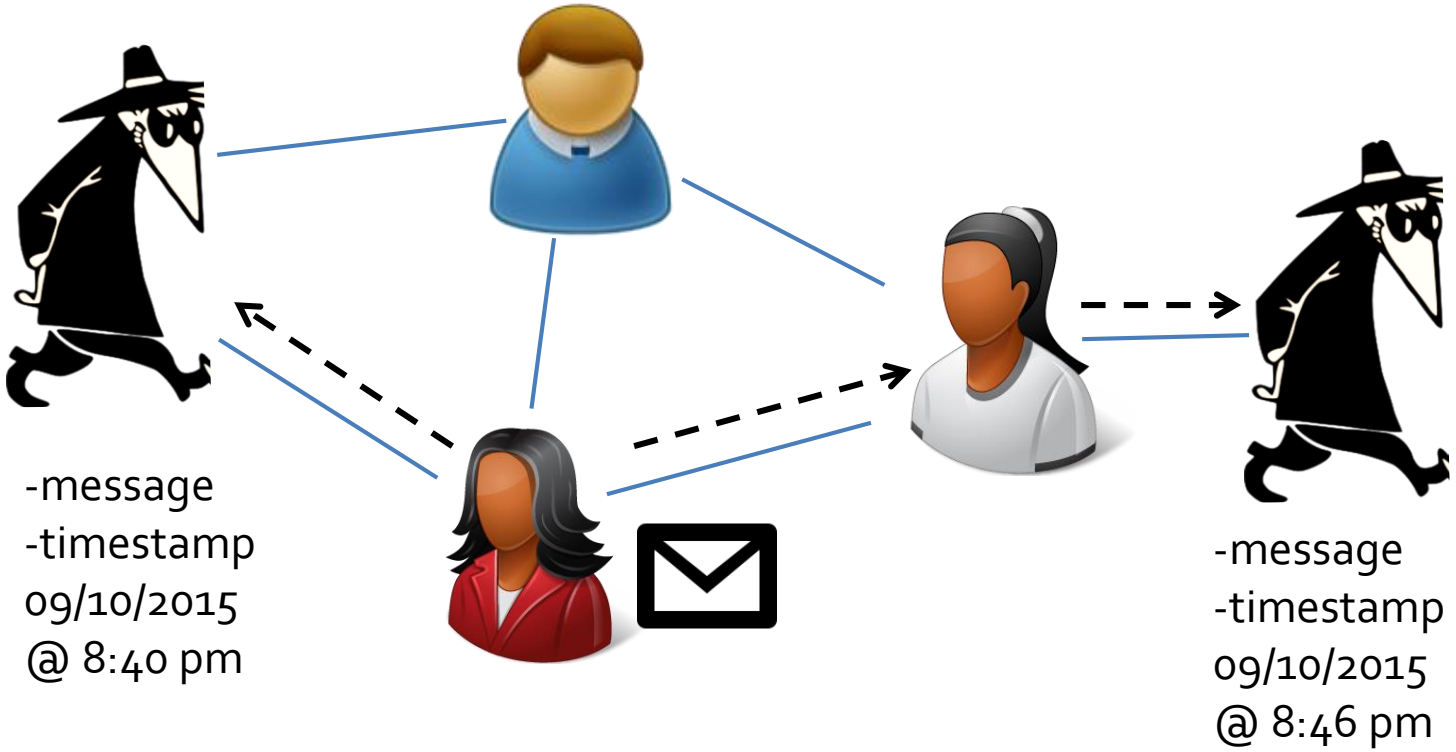
snapshot



adversary can figure out **who got the message**

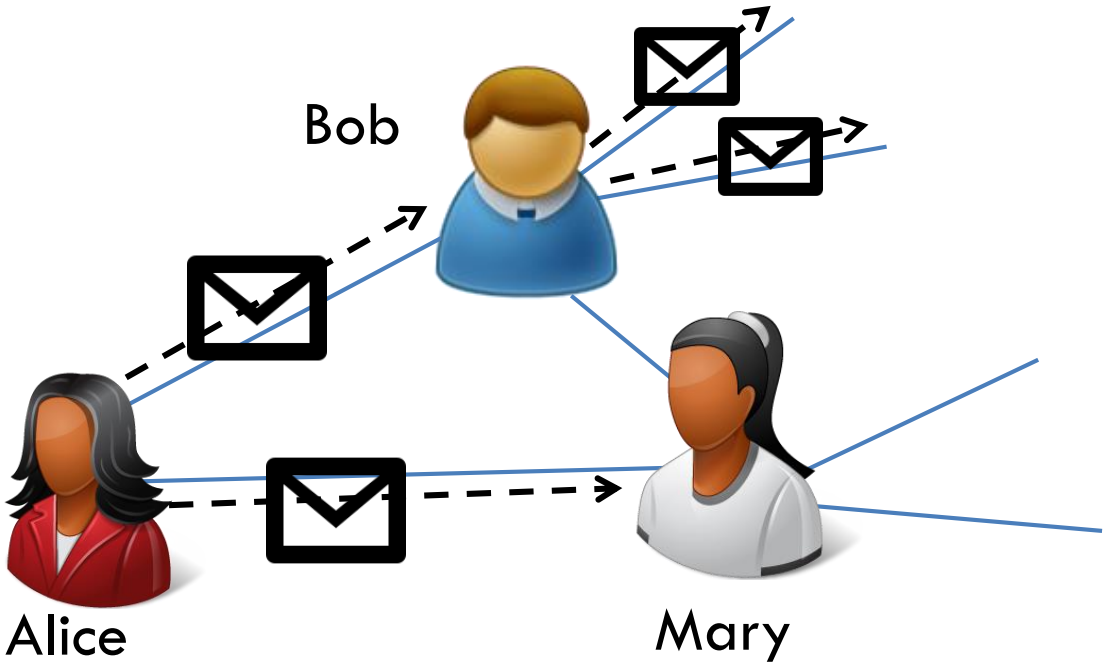
Adversary with timing

spy-based



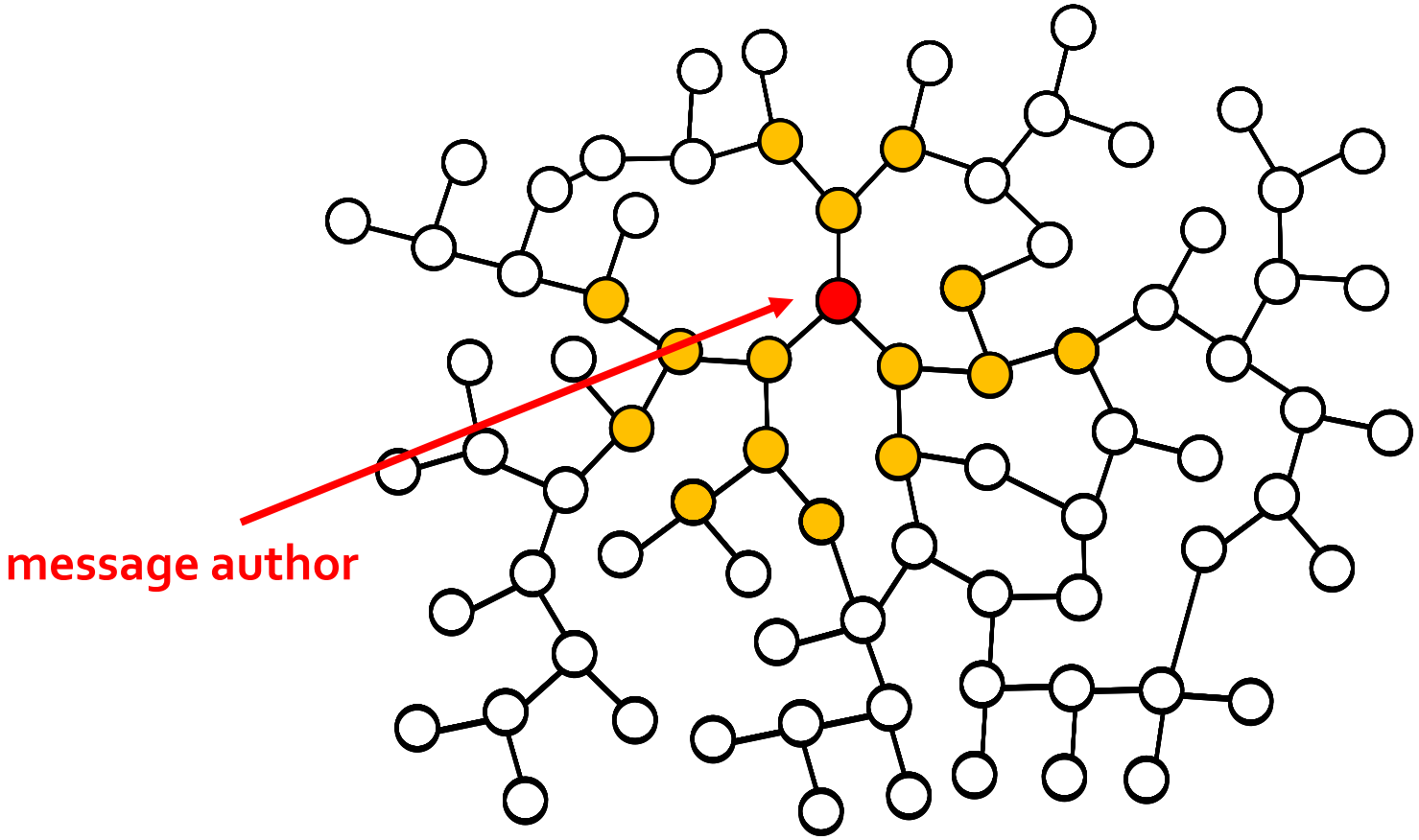
adversary can **collect timing information**

First attempt: distributed networks



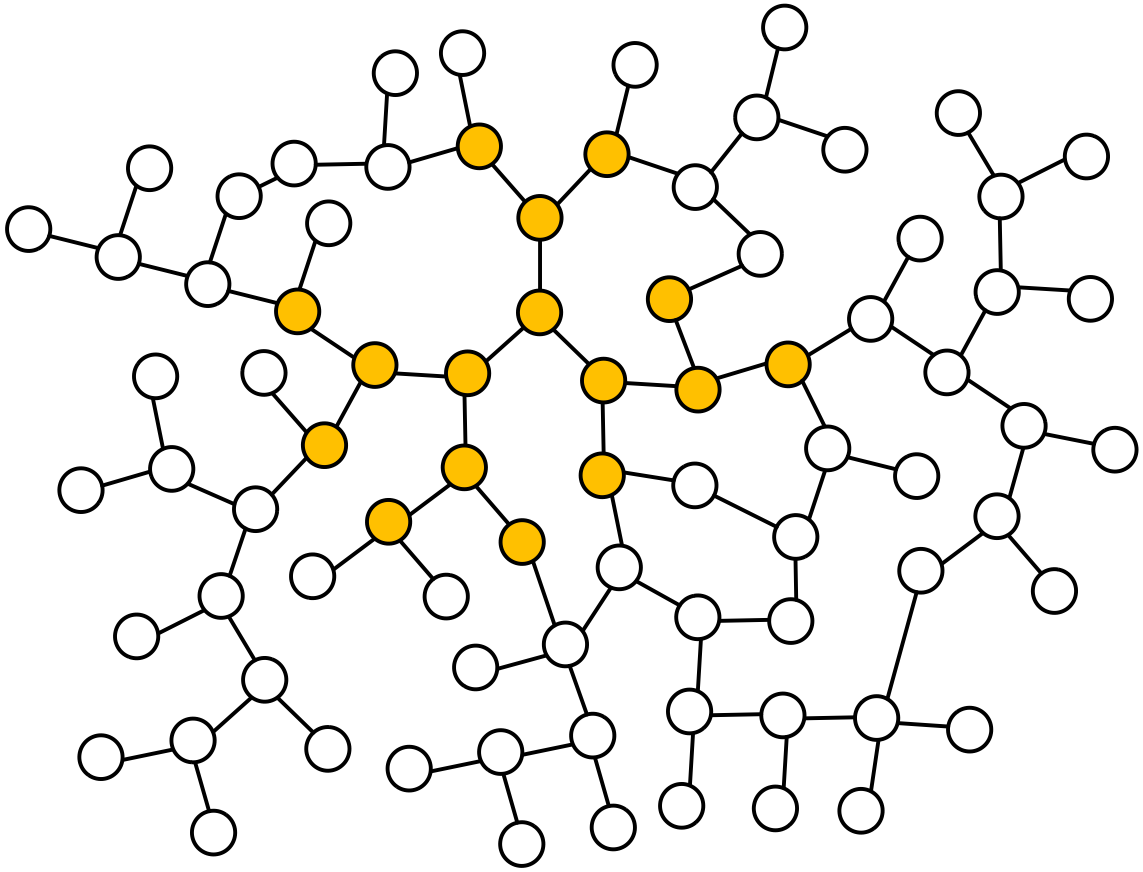
adversary can **still infer the author**

Information flow in social networks



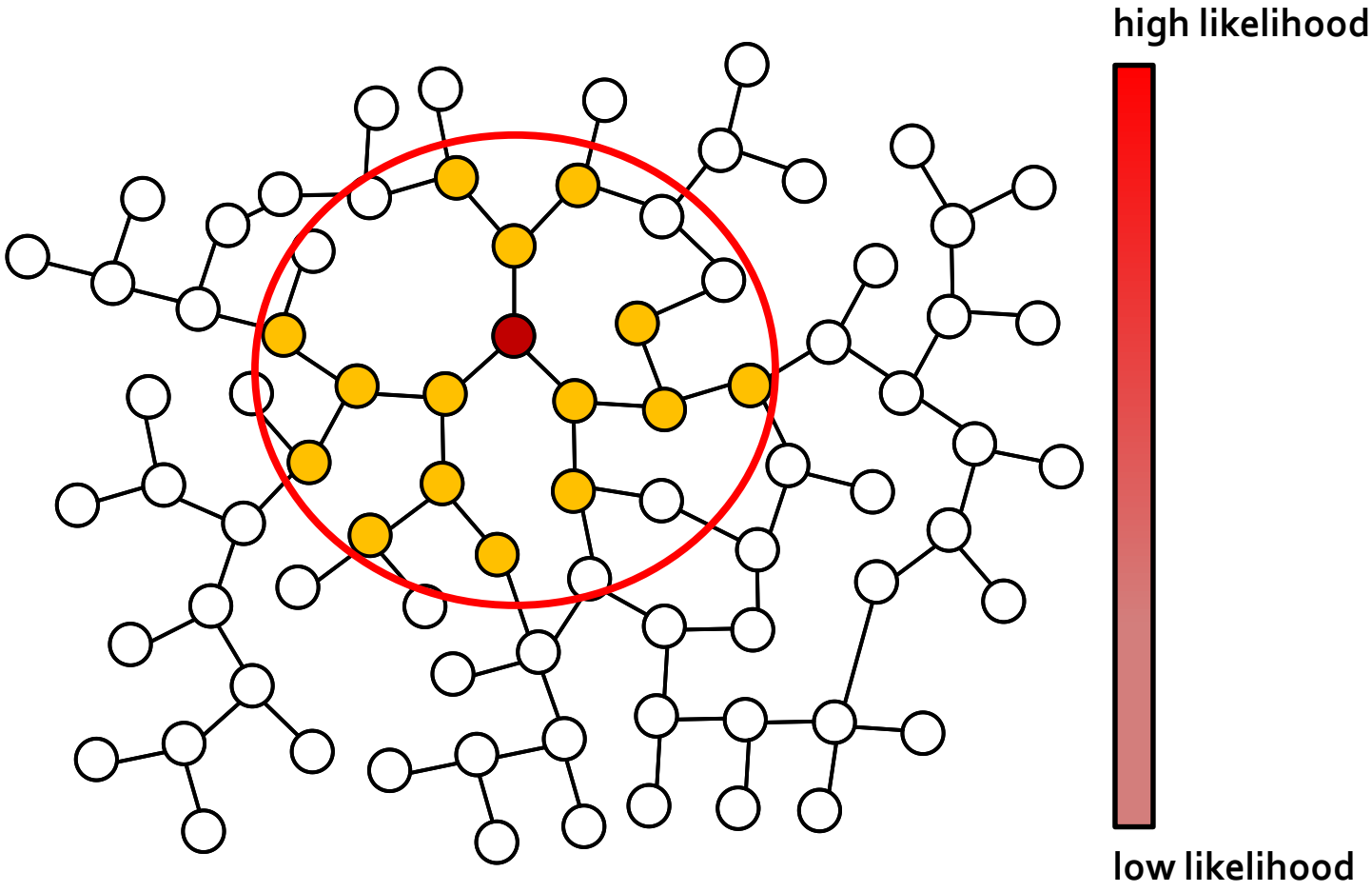
information spreads at the **same rate** in **all direction**

Can you find the source?



[*Shah et al. 2011, Pinto et al. 2012, Zhu et al. 2013*]

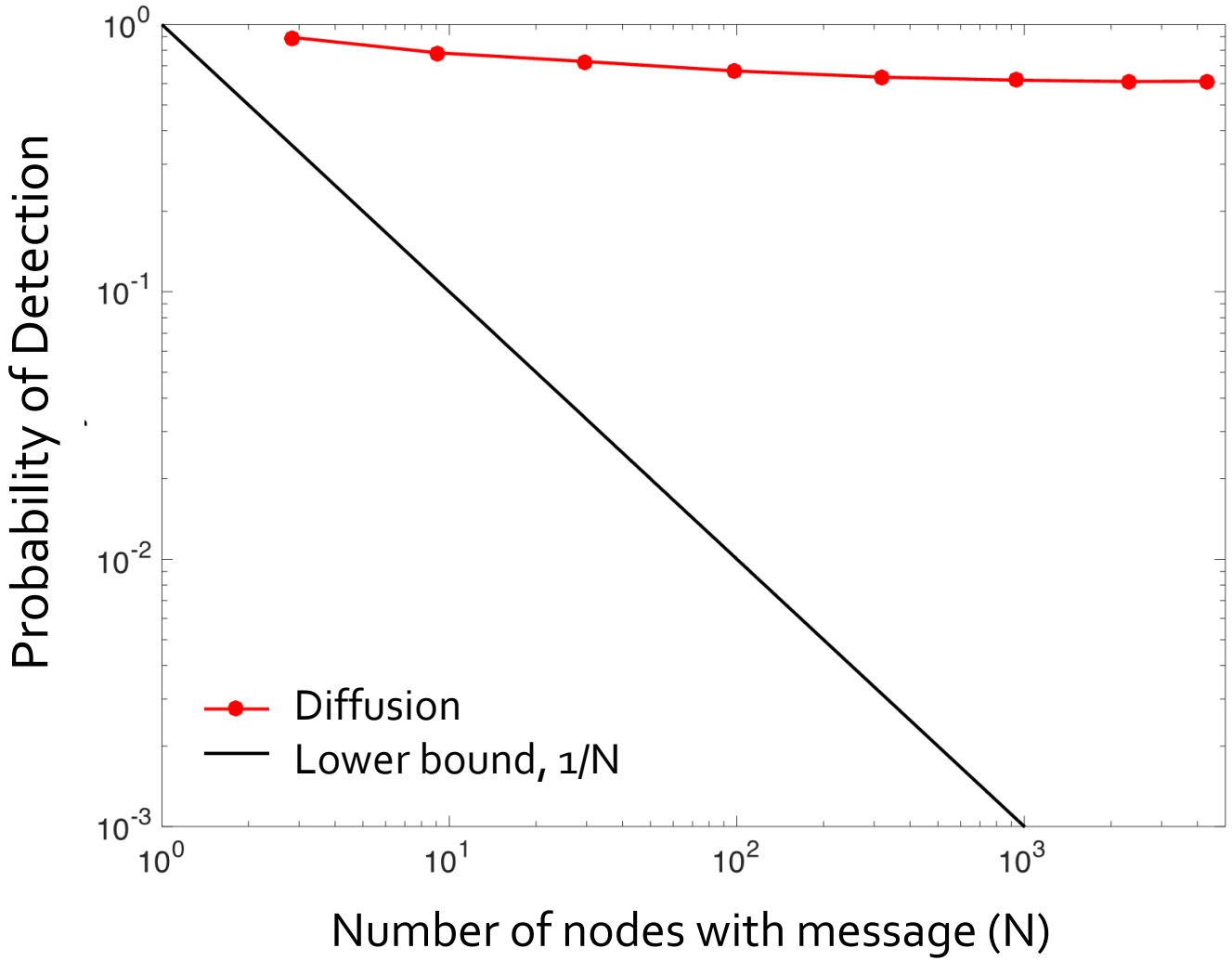
Concentration around the center



diffusion spreading = **de-anonymization**

[*Shah et al. 2011, Pinto et al. 2012, Zhu et al. 2013*]

De-anonymization on social networks



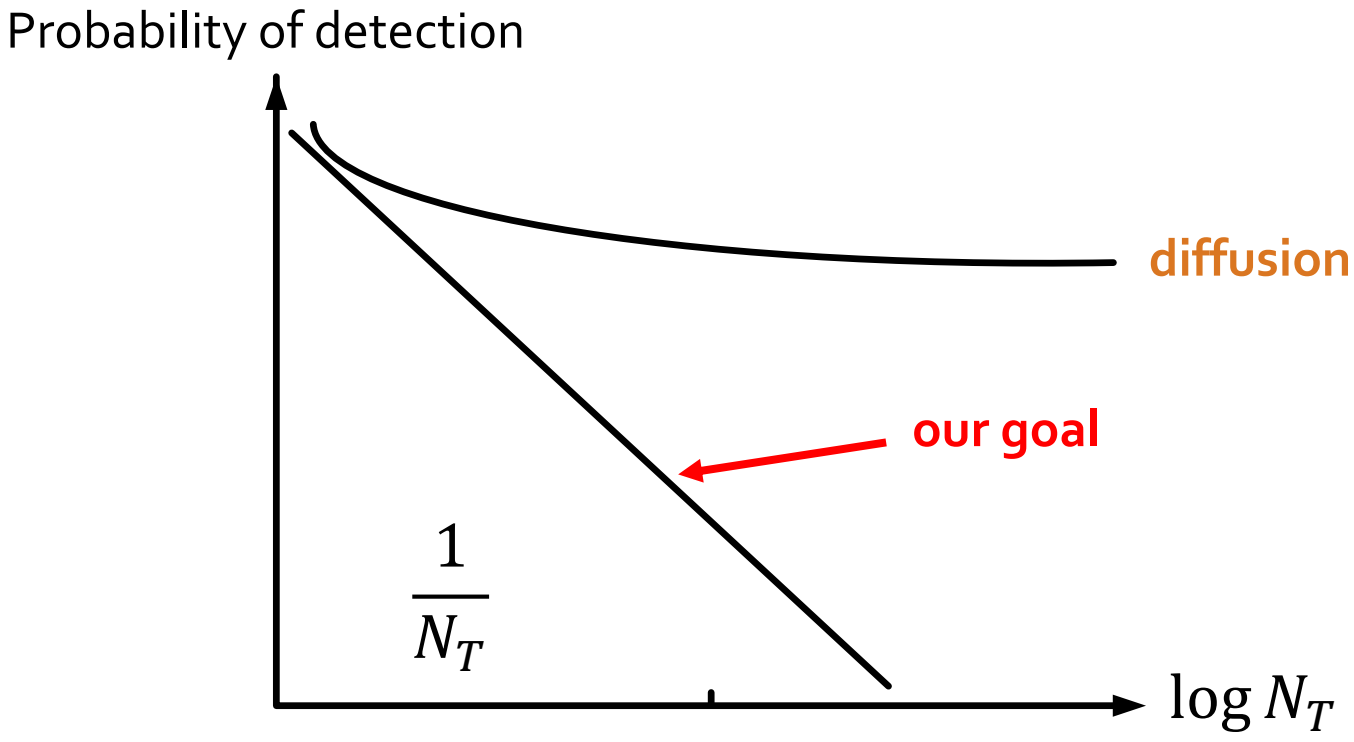
[*Seo et al. 2012*]

our first attempt failed!

Spreads fast ☺

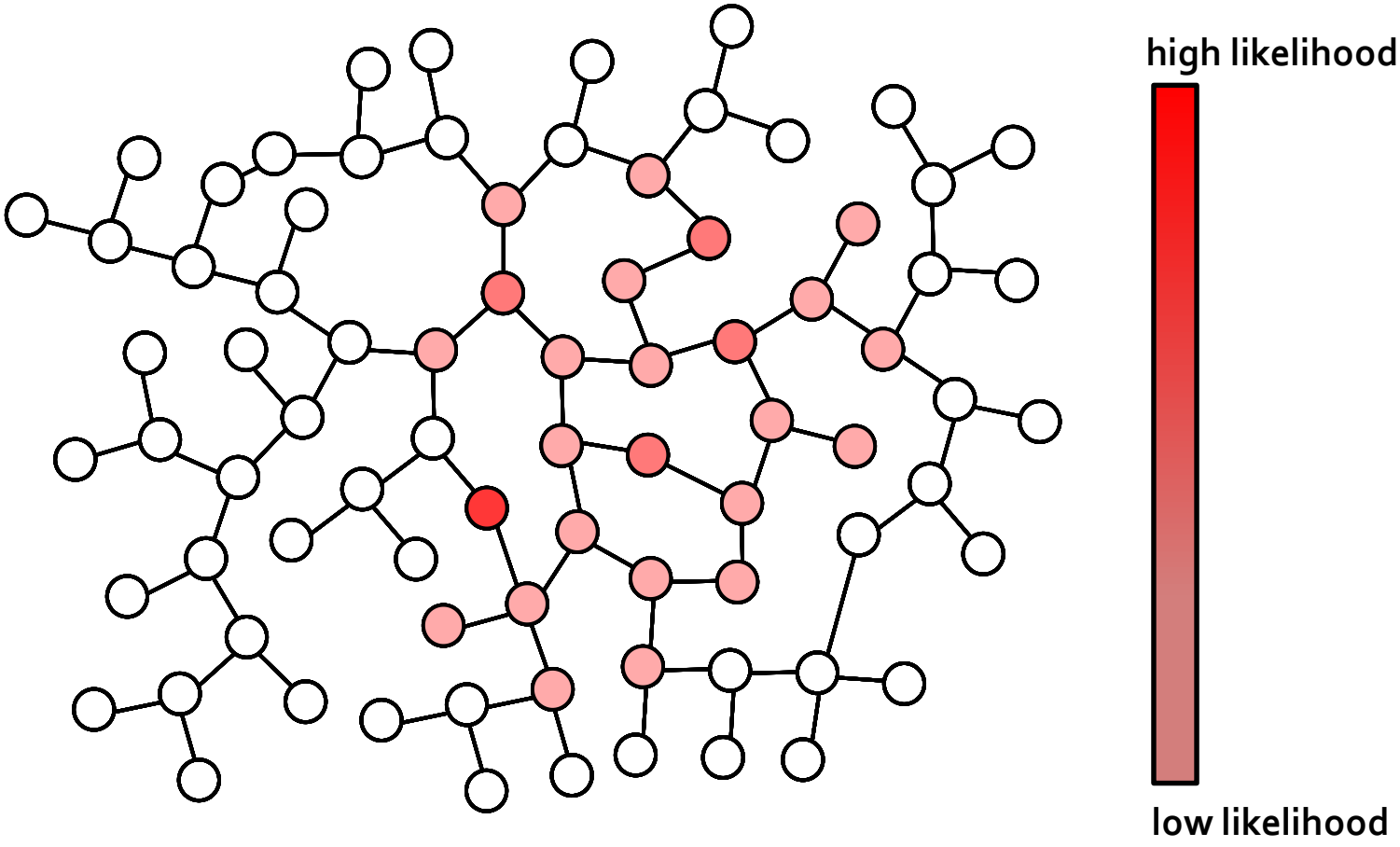
Bad anonymity properties ☹

Objectives



N_T = number of nodes with the message at time T

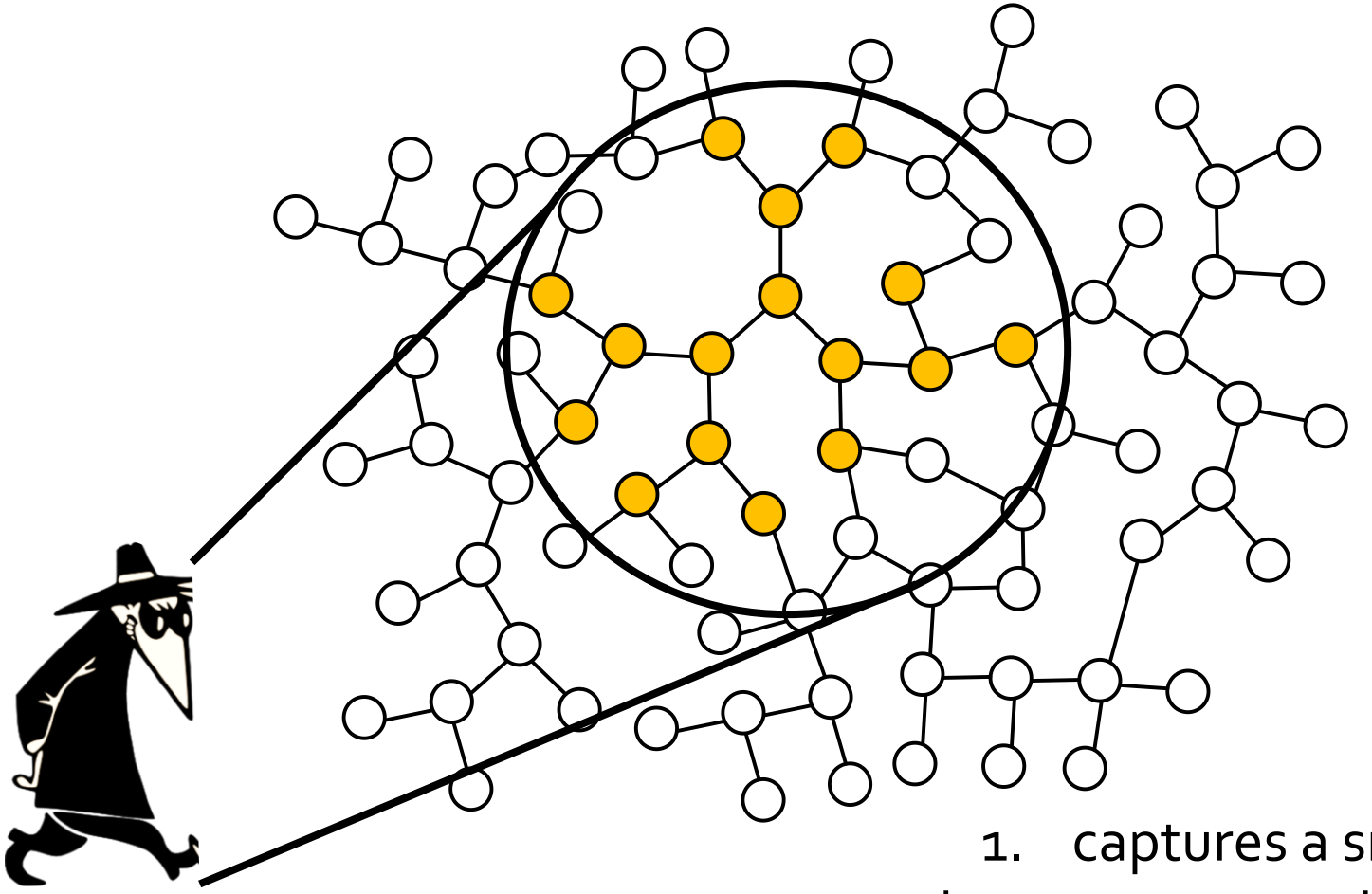
Adaptive diffusion



provides provable anonymity guarantees

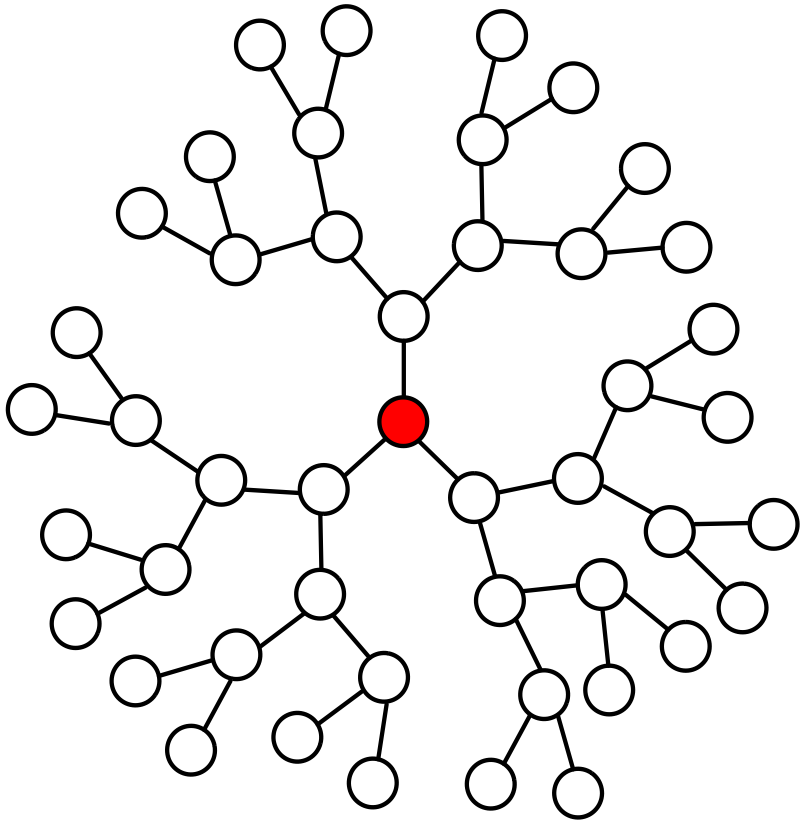
[Best Paper Award at SIGMETRICS 15, SIGMETRICS 16]

Snapshot adversary



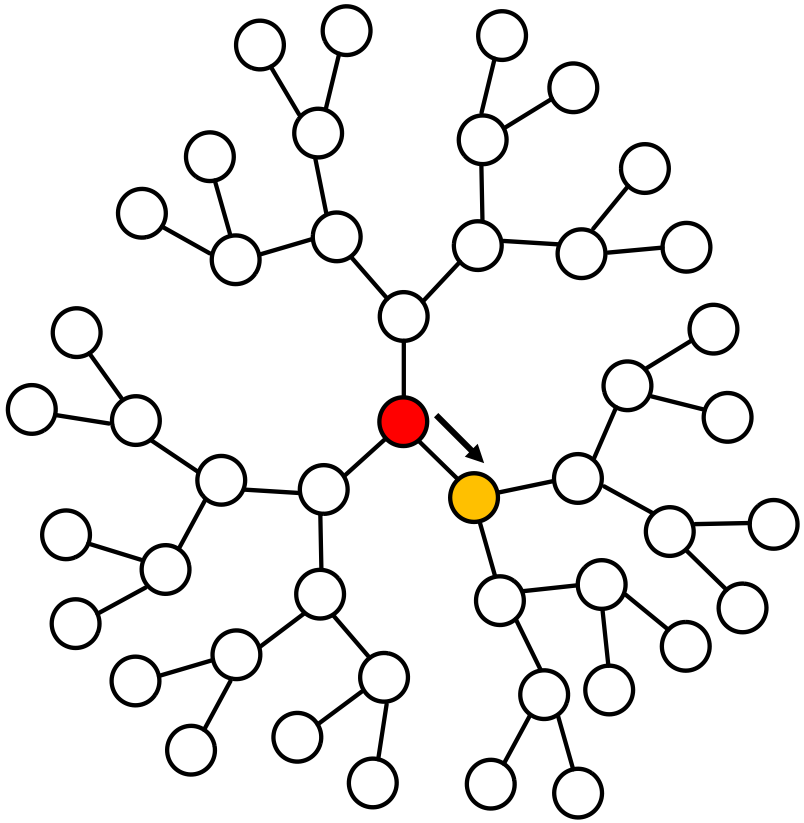
1. captures a snapshot
2. has access to underlying graph
3. knows the spreading protocol
4. is computationally unbounded

d-regular trees: **adaptive diffusion**



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

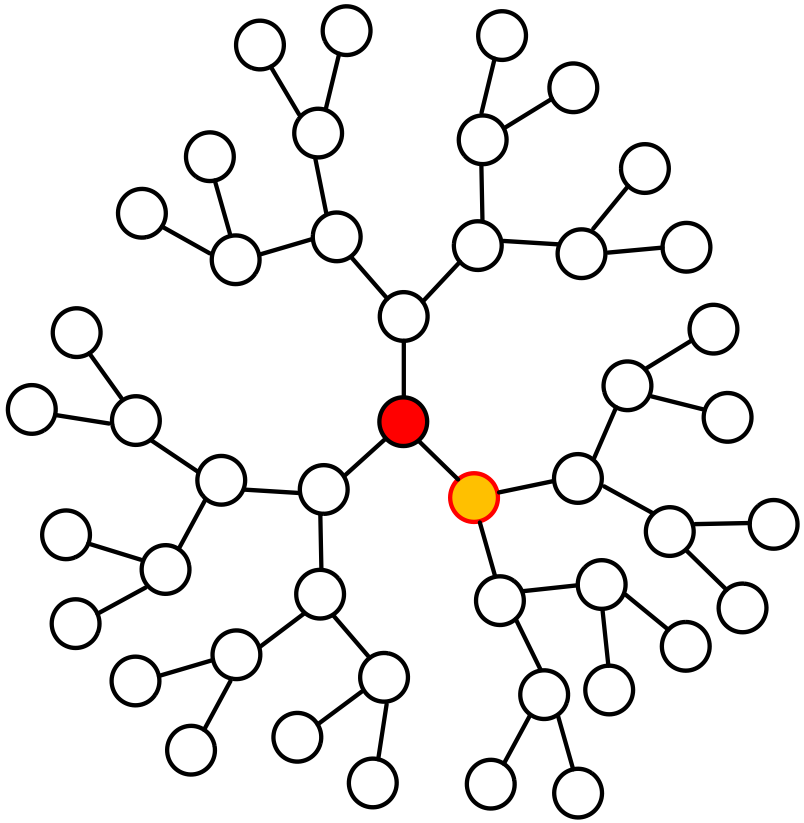
d-regular trees: **adaptive diffusion**



break directional symmetry

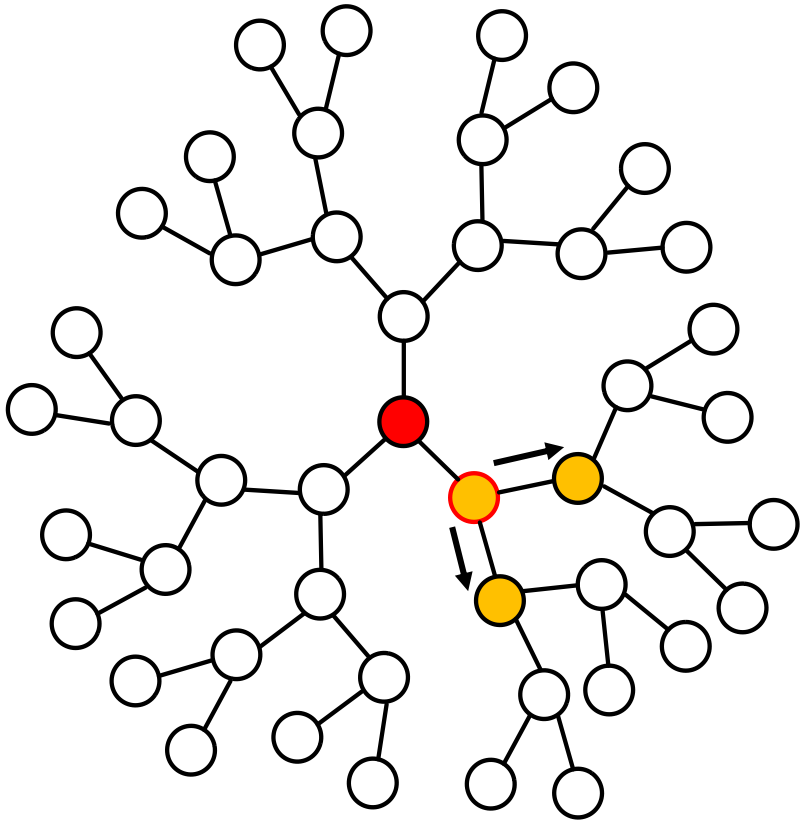
$T = 1$, the author selects one neighbor at random

d-regular trees: **adaptive diffusion**



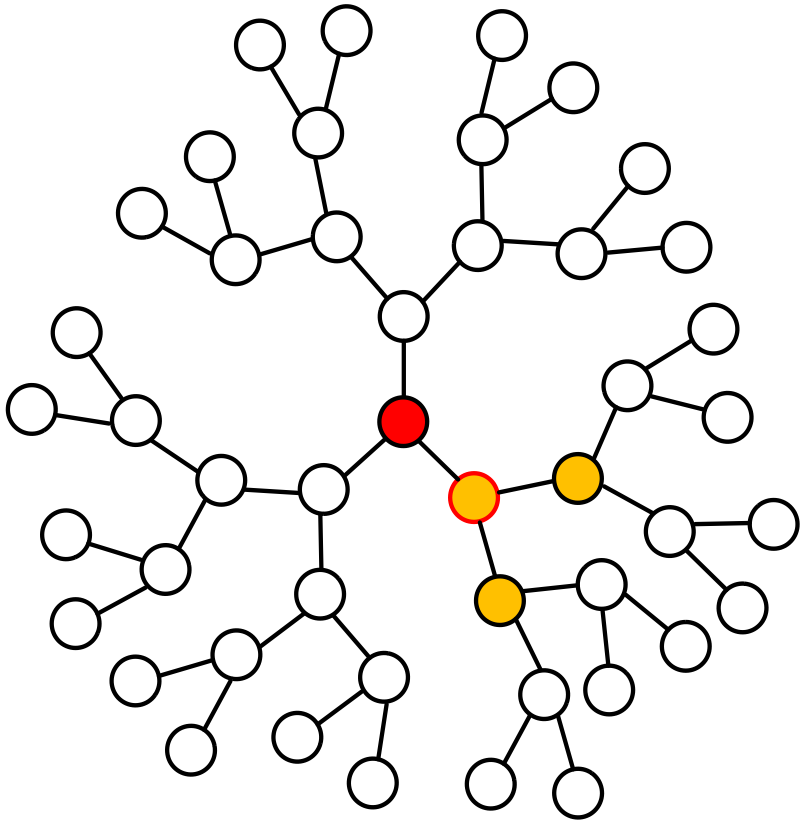
chosen neighbor = new virtual source

d-regular trees: **adaptive diffusion**



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

d-regular trees: **adaptive diffusion**

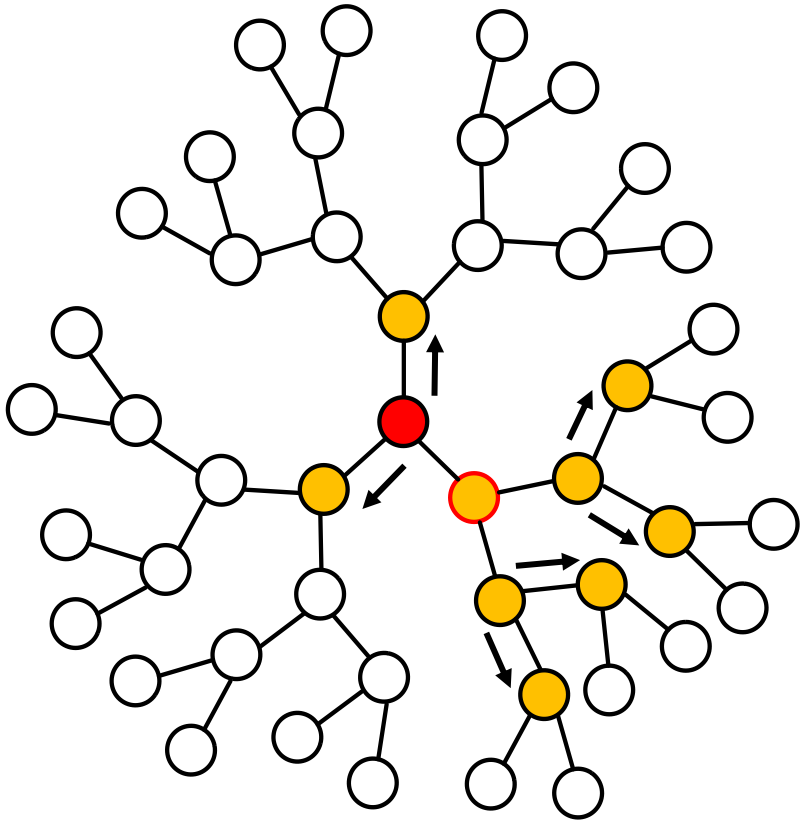


break temporal symmetry

keep the virtual source token

pass the virtual source token

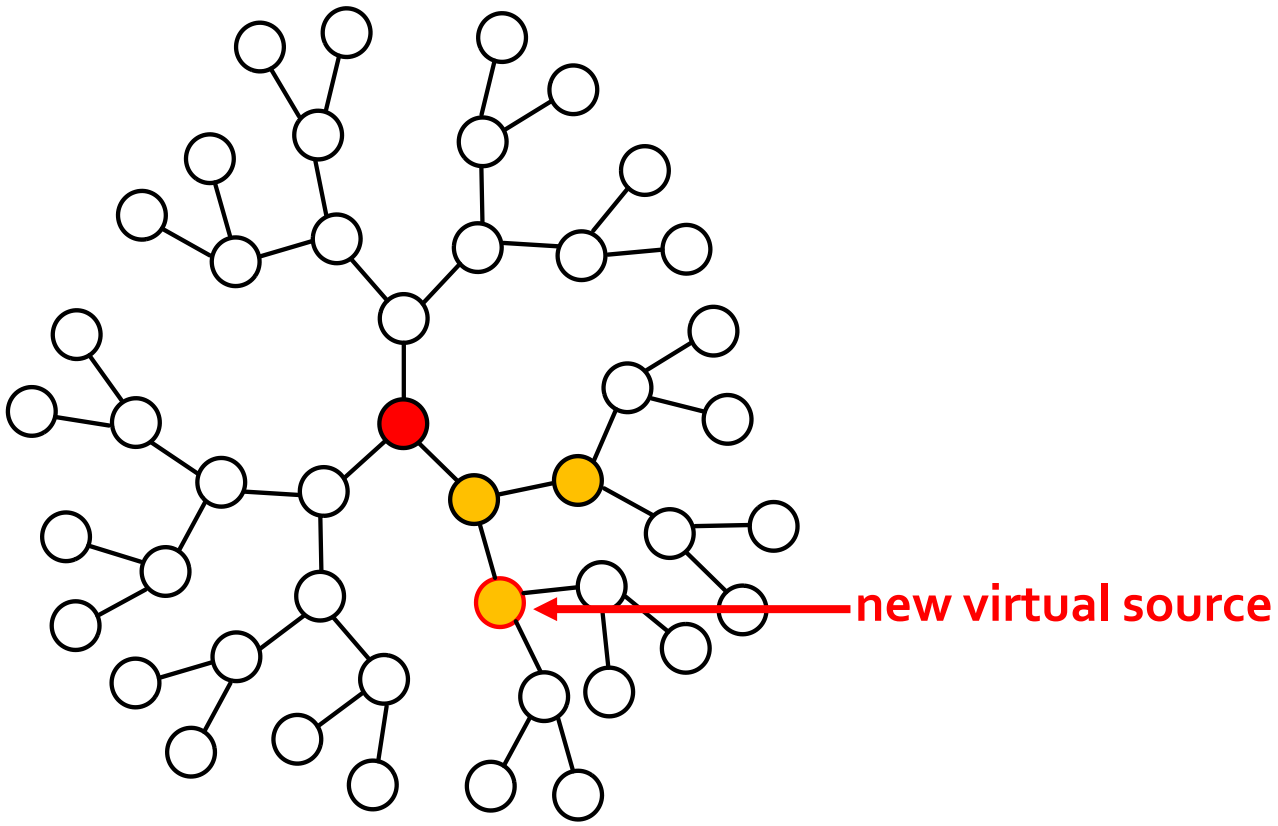
keep the virtual source token



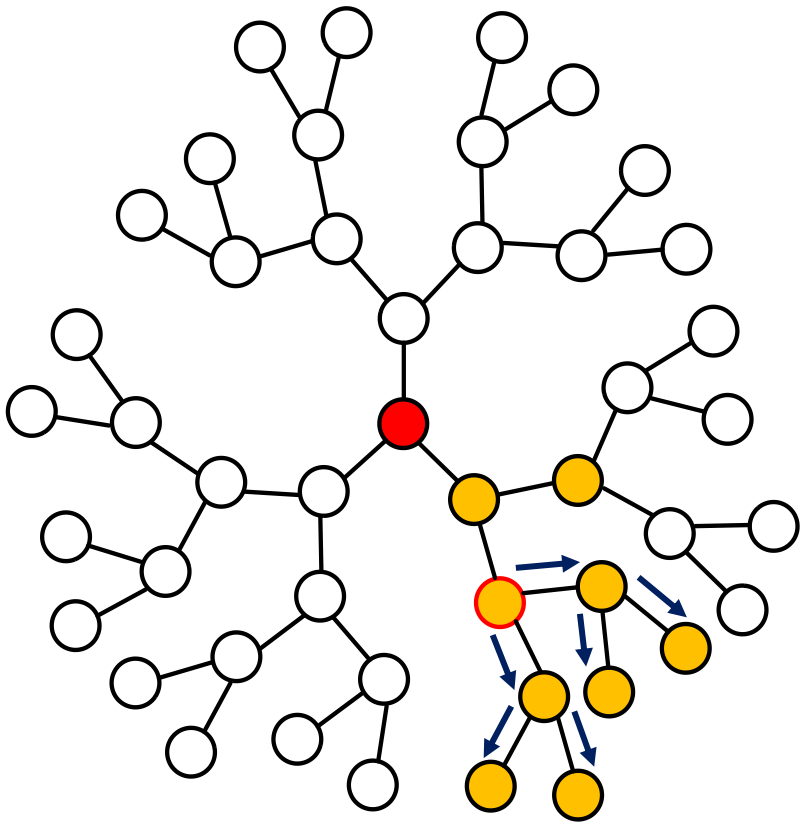
happens in $T = 3$ and $T = 4$

spread the message in all directions

pass the virtual source token



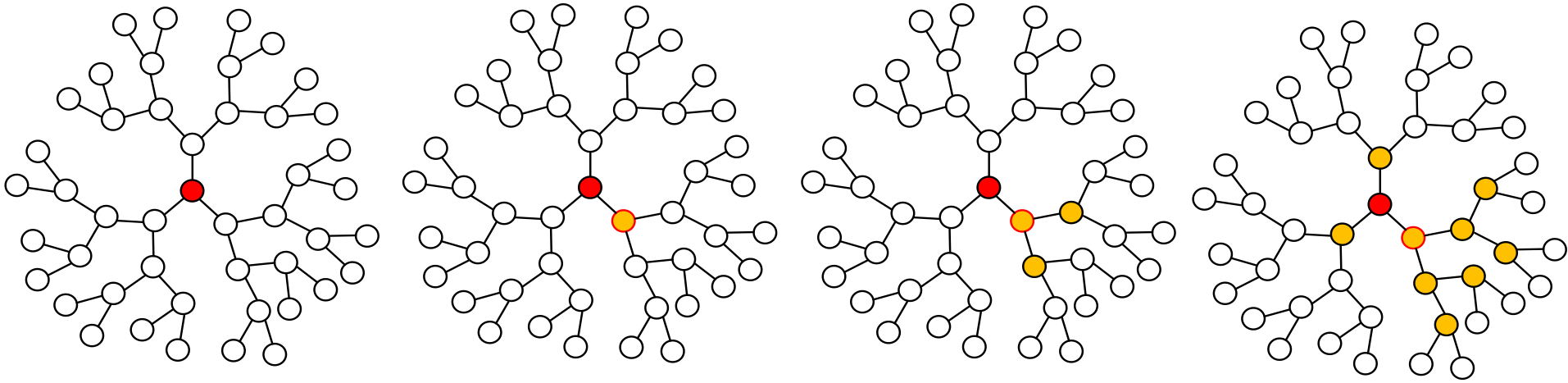
pass the virtual source token



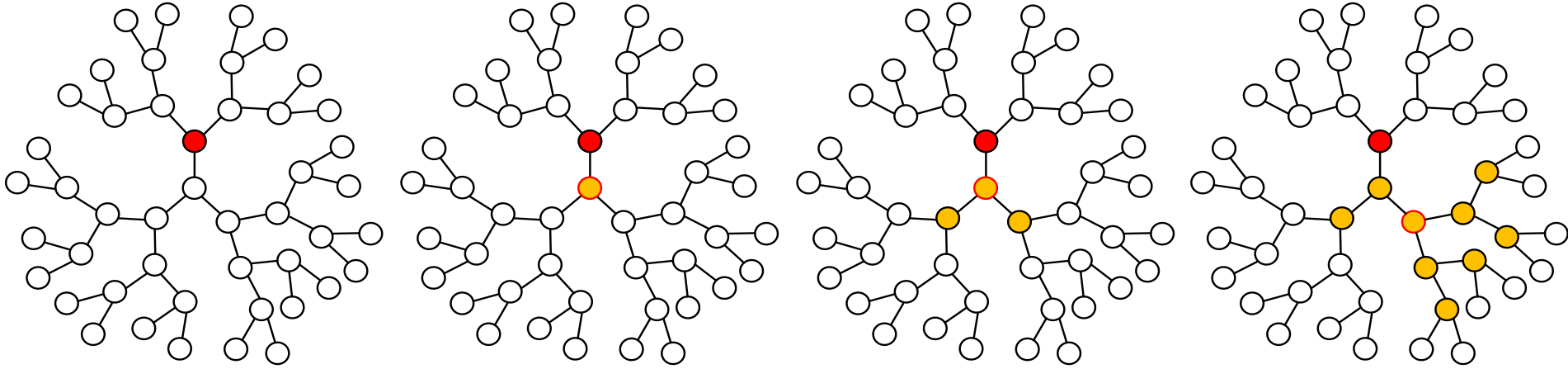
happens in $T = 3$ and $T = 4$

spread the message in the direction of: old → new virtual source

Sample paths

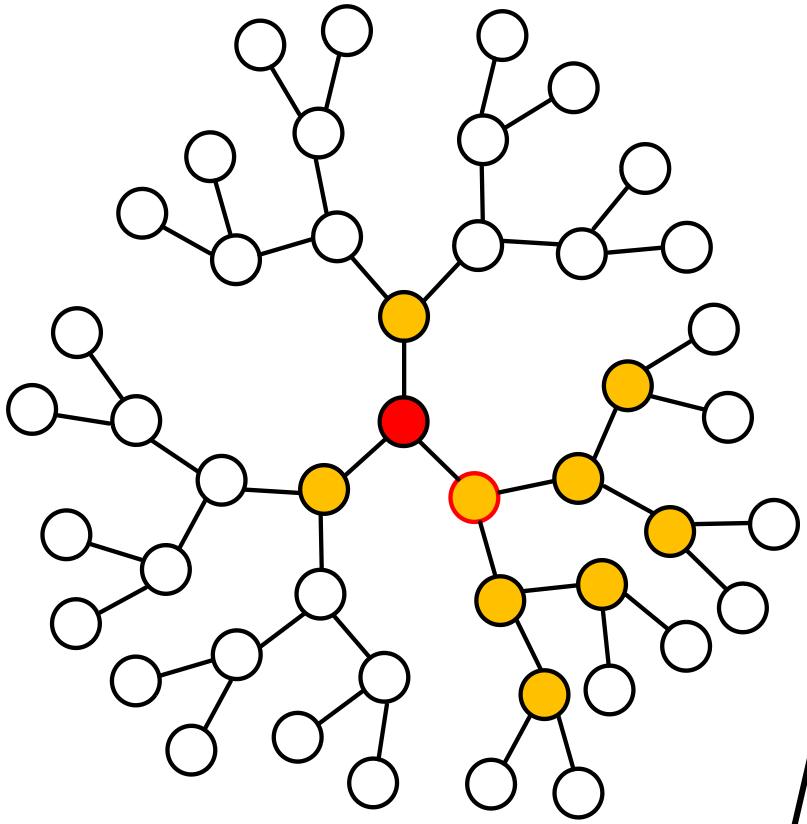


choose a neighbor at random → keep the virtual source token



choose a neighbor at random → pass the virtual source token

When to keep the virtual source



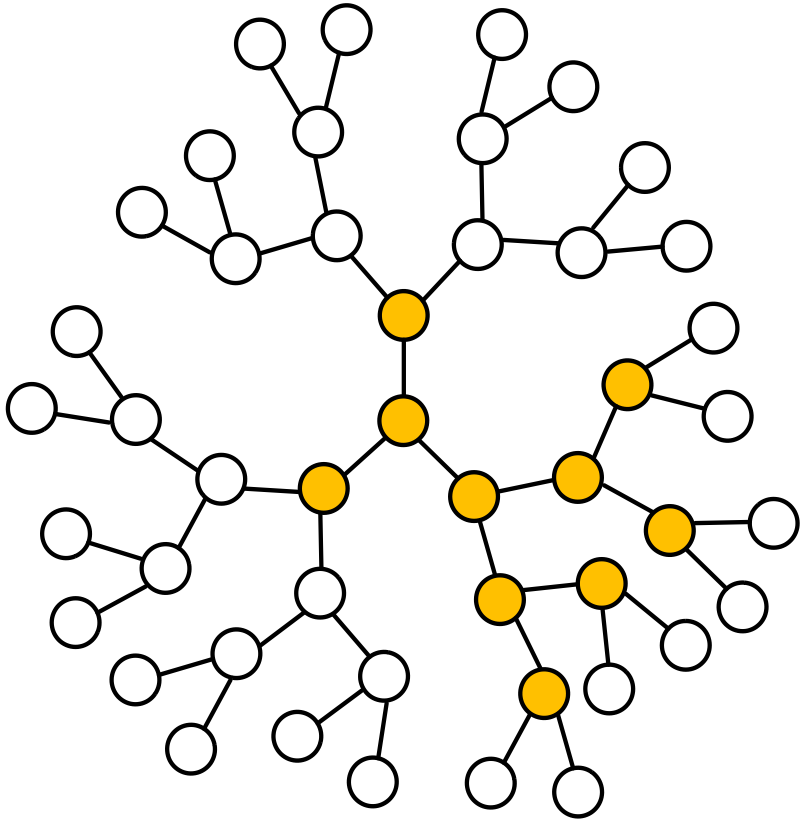
node degree

time index

distance to source

virtual source token is kept w.p. $\alpha = \frac{(d-1)^{\frac{T}{2}-h+1}-1}{(d-1)^{\frac{T}{2}+1}-1}$

Adversary with a snapshot



can the adversary locate the source of the message?

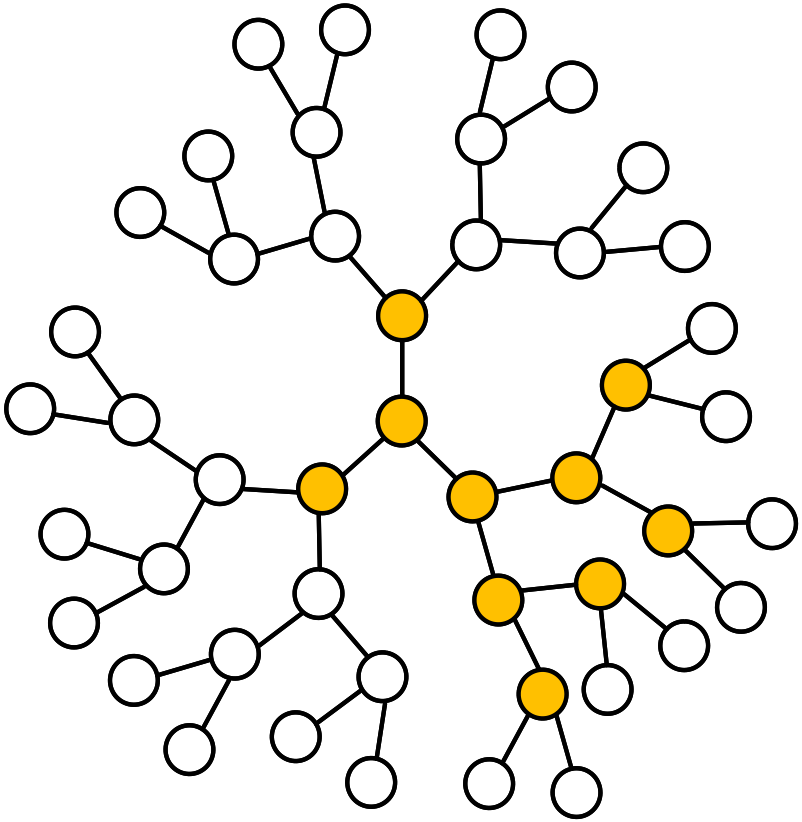
THEOREM

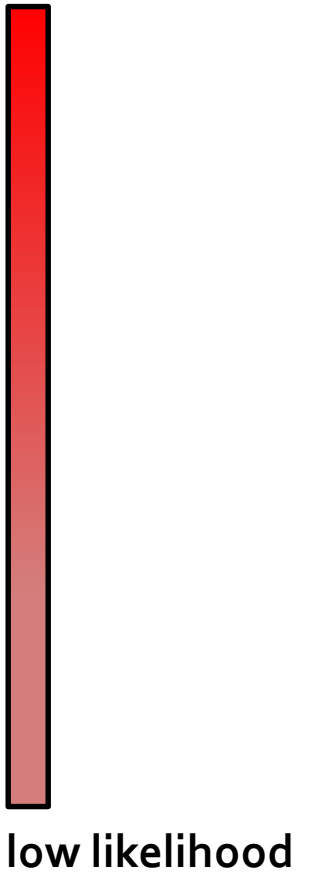
1. We spread fast: $N_T \approx (d-1)^{\frac{T}{2}}$.
2. Probability of source detection under maximum likelihood detection is given by

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

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

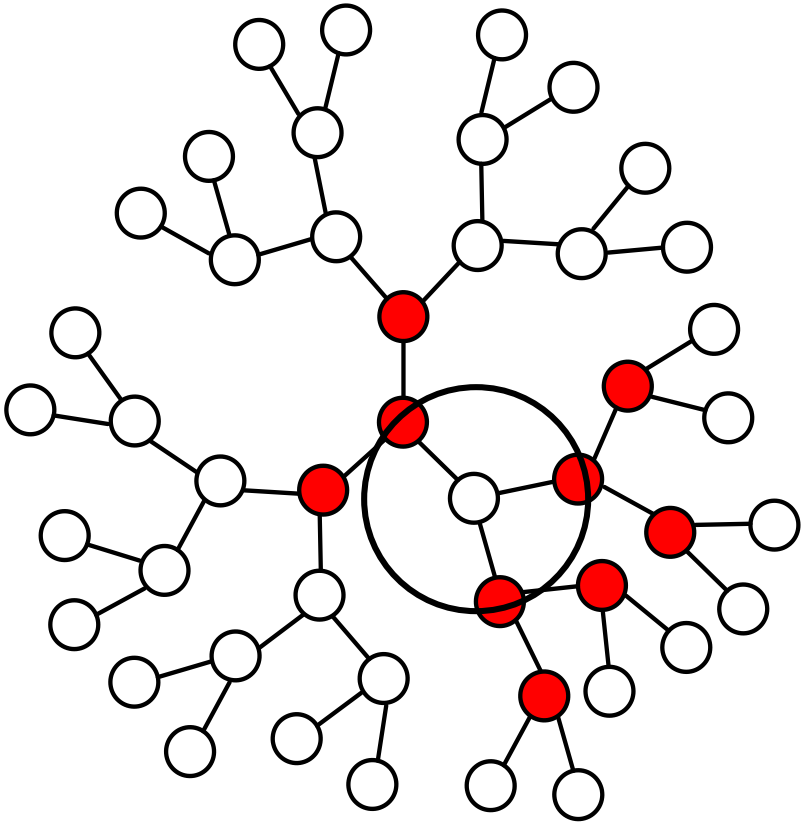
Proof sketch



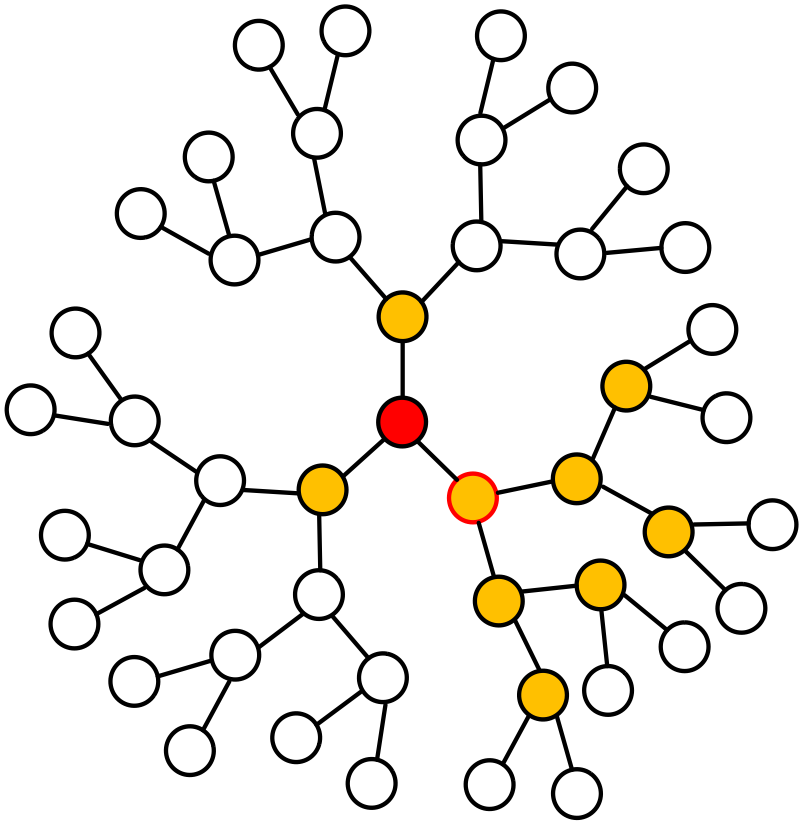


all nodes except for the final virtual source are equally likely

Proof sketch



nodes at the same hop distance from the virtual source are equally likely



Markov chain over the hop distance between the true source and final virtual source

$$h_t = d(v^*, vs_t)$$

After every 2 time steps

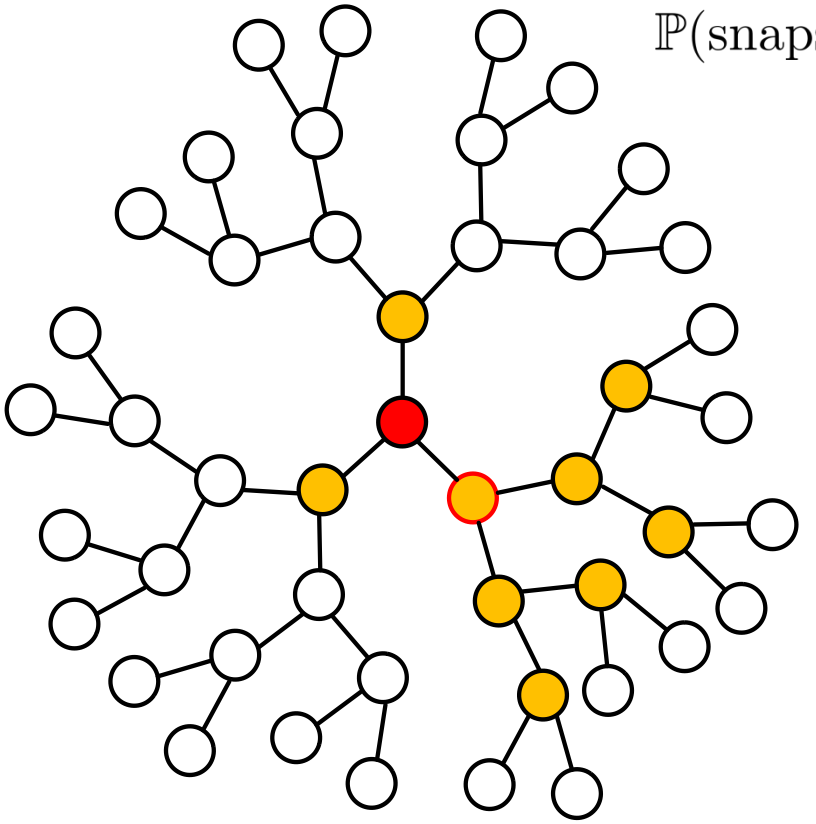
$$h_{t+2} = h_t \text{ or } h_{t+2} = h_t + 1$$

Probability of keeping token

$$\alpha_d(t,h) = \mathbb{P}\big(h_{t+2} = h_t | h_t = h \big)$$

Probability state vector

$$p^{(t)} = [\mathbb{P}(h_t = h)]_{h \in \{1,\ldots,t/2\}}$$



$$\mathbb{P}(\text{snapshot}|v, d(v, vs_t) = h) = \frac{1}{d(d-1)^{h-1}} p_h^{(t)}$$

$$= \frac{d-2}{d((d-1)^{t/2}-1)}$$

$$p^{(t)} = \frac{d-2}{(d-1)^{t/2}-1} \begin{bmatrix} 1 \\ (d-1) \\ \vdots \\ (d-1)^{t/2-1} \end{bmatrix}$$

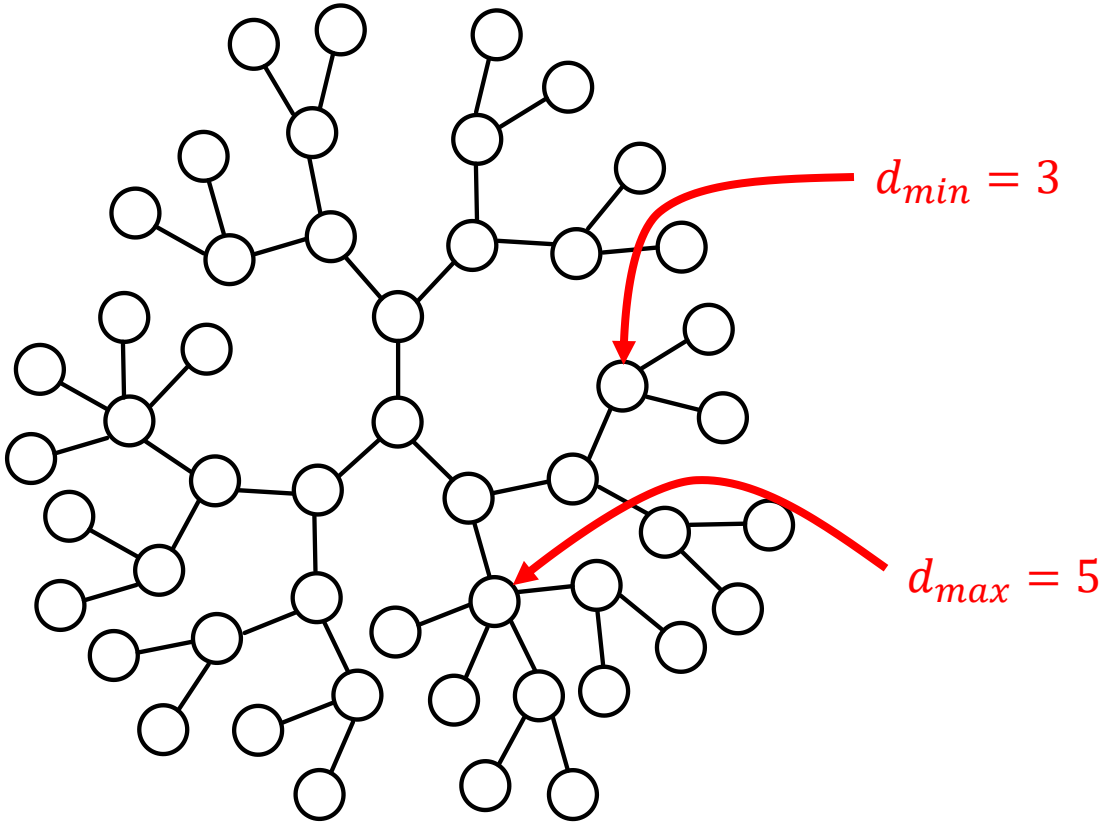
$$\alpha_d(t,h) = \frac{(d-1)^{t/2-h+1}-1}{(d-1)^{t/2+1}-1}$$

State transition matrix

$$p^{(t+2)} = \begin{bmatrix} \alpha_d(t,1) & & & & \\ 1-\alpha_d(t,1) & \alpha_d(t,2) & & & \\ & 1-\alpha_d(t,2) & \ddots & & \\ & & \ddots & \alpha_d(t,t/2) & \\ & & & 1-\alpha_d(t,t/2) & \end{bmatrix} p^{(t)}$$

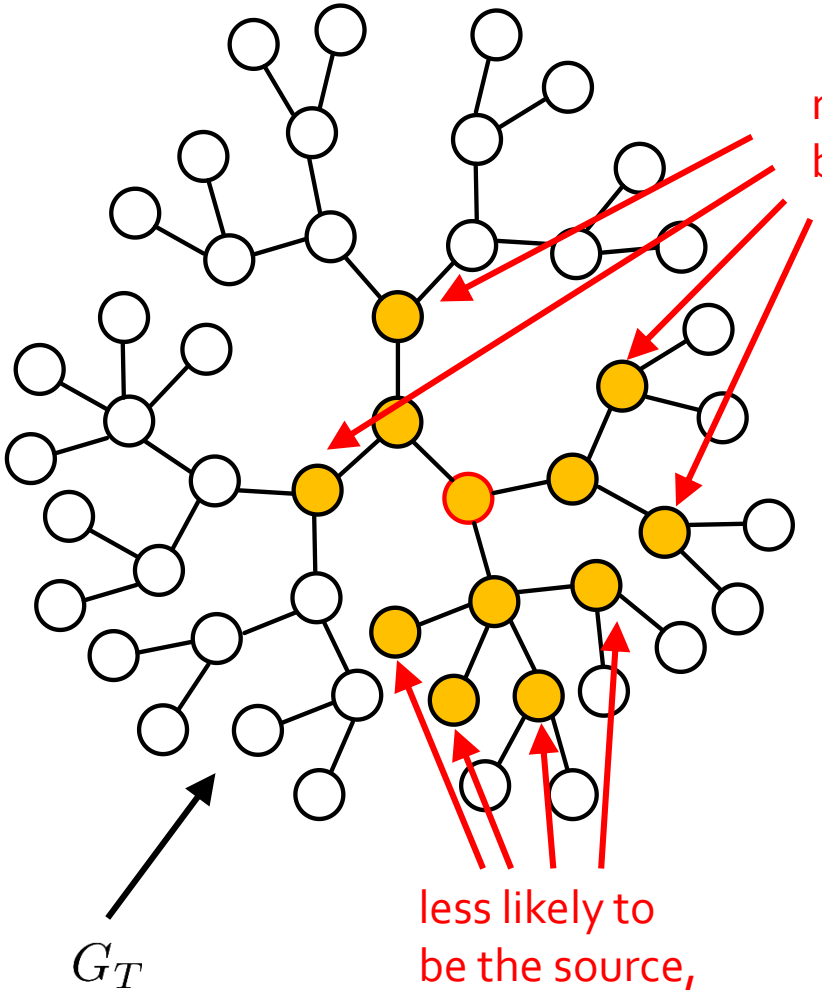
What about irregular trees?

$$d_v = \begin{cases} 3 & w.p. \quad 0.8 \\ 5 & w.p. \quad 0.2 \end{cases}$$



virtual source is moved with probability 1 ($\alpha = 0$)

Maximum likelihood detection

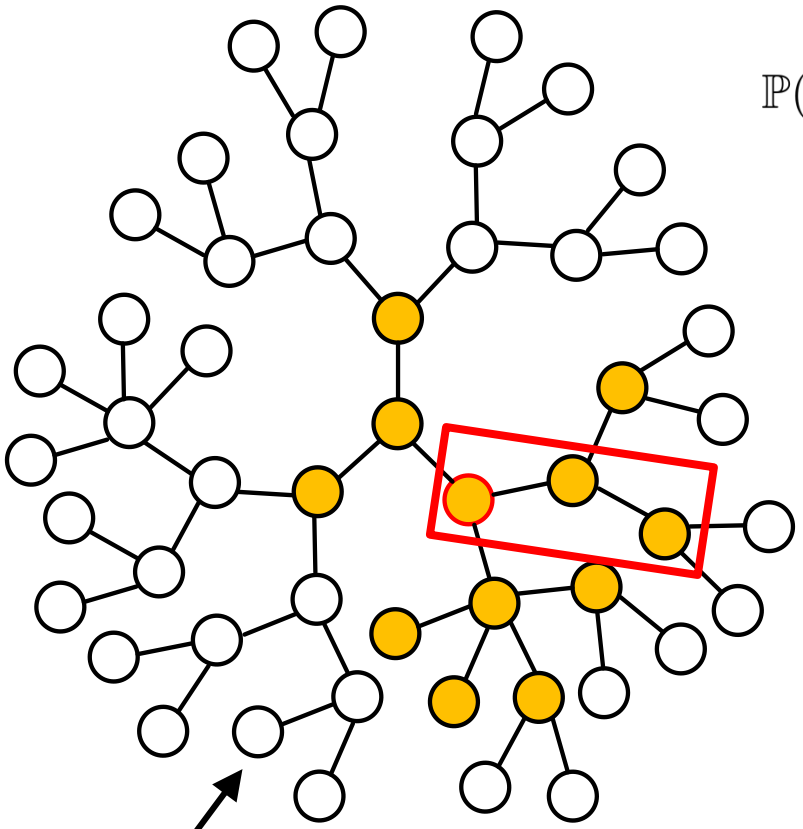


$$\hat{v}_{\mathrm{ML}} = \arg \min_{v \in \partial G_T} \prod_{\substack{w \in \phi(v,vs) \\ \setminus \{vs,v\}}} (d_w - 1)$$

path from leaf
node v to
virtual source

Performance of ML detection

$$d_v = \begin{cases} d_{min} & w.p. \quad p_{min} \\ d_{max} & w.p. \quad p_{max} \end{cases}$$



$$\mathbb{P}(\hat{v}_{\mathrm{ML}} = v^* | G_T) = \frac{1}{d_{vs} \min\limits_{v \in \partial G_T} \prod\limits_{\substack{w \in \phi(v,vs) \\ \setminus \{vs,v\}}} (d_w - 1)}$$

path from leaf node v to virtual source

degree of node w

$$\mathbb{P}(\hat{v}_{\mathrm{ML}} = v^*) = \mathbb{E}_{G_T}[\mathbb{P}(\hat{v}_{\mathrm{ML}} = v^* | G_T)]$$

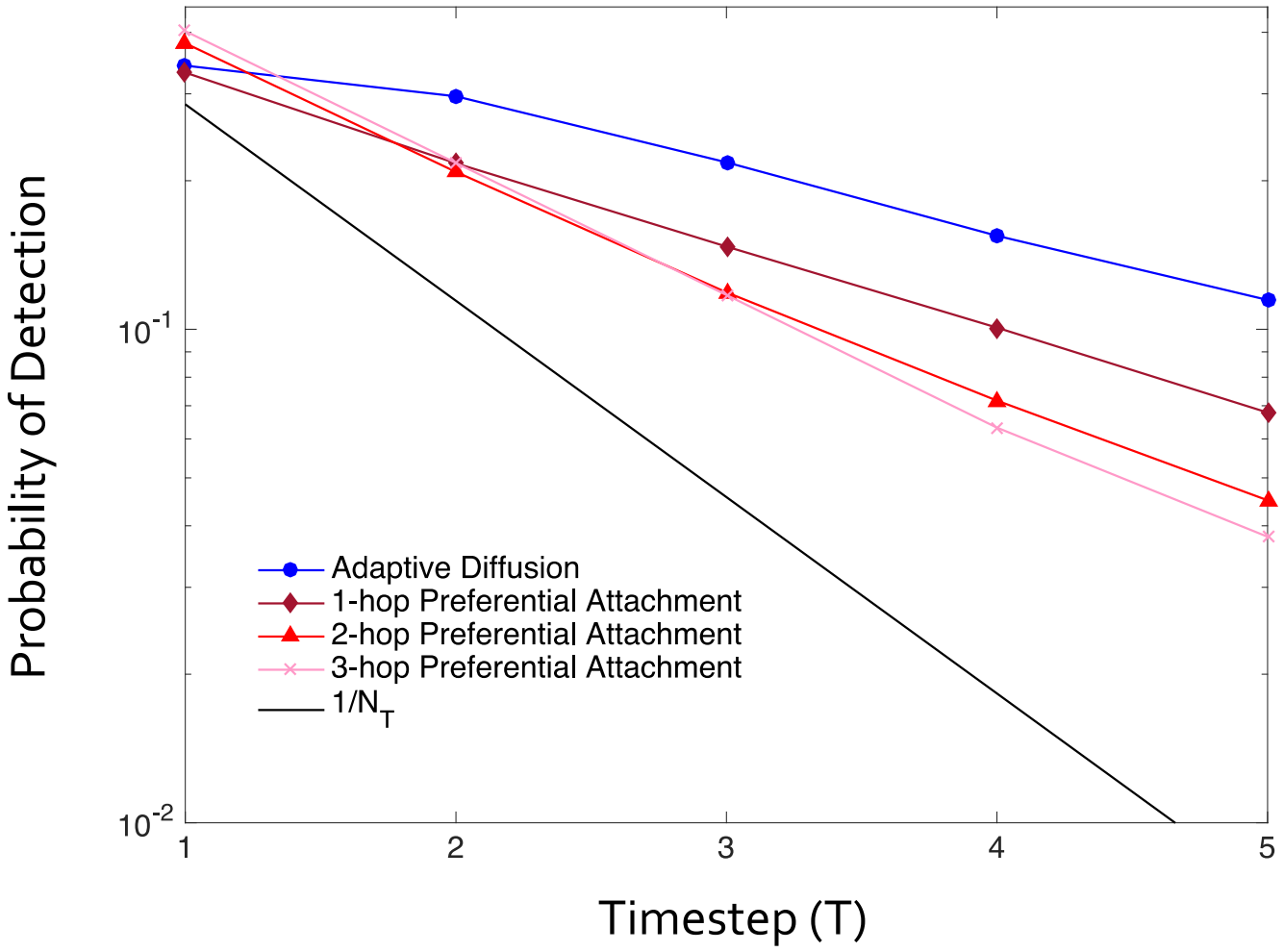
$\frac{1}{9}$ If $p_{min}(d_{min} - 1) > 1$

$$\mathbb{P}(\hat{v}_{\mathrm{ML}} = v^* | G_T) \approx (d_{\min} - 1)^{-T/2}$$

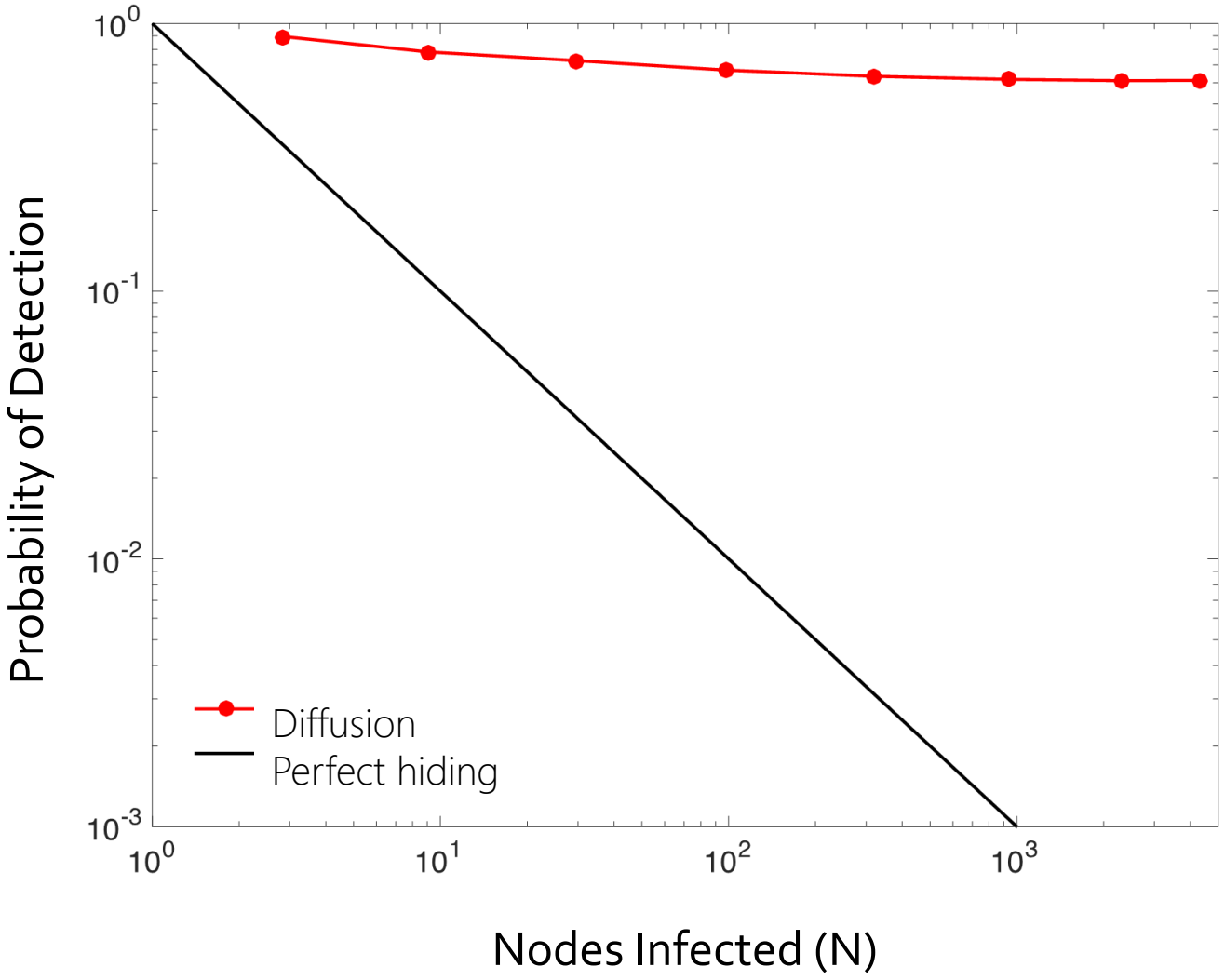
G_T

THEOREM: probability of detection $\approx \frac{1}{(d_{min}-1)^{T/2}}$

Preferential attachment protocol



Facebook graph



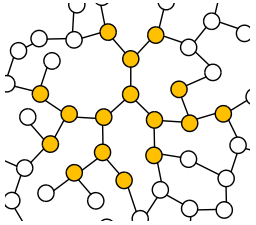
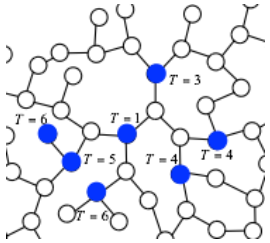
Lessons for snapshot adversary

1) Diffusion = de-anonymization

2) For anonymity, break symmetry.

3) For *more* anonymity, hide in a crowd.

Our results

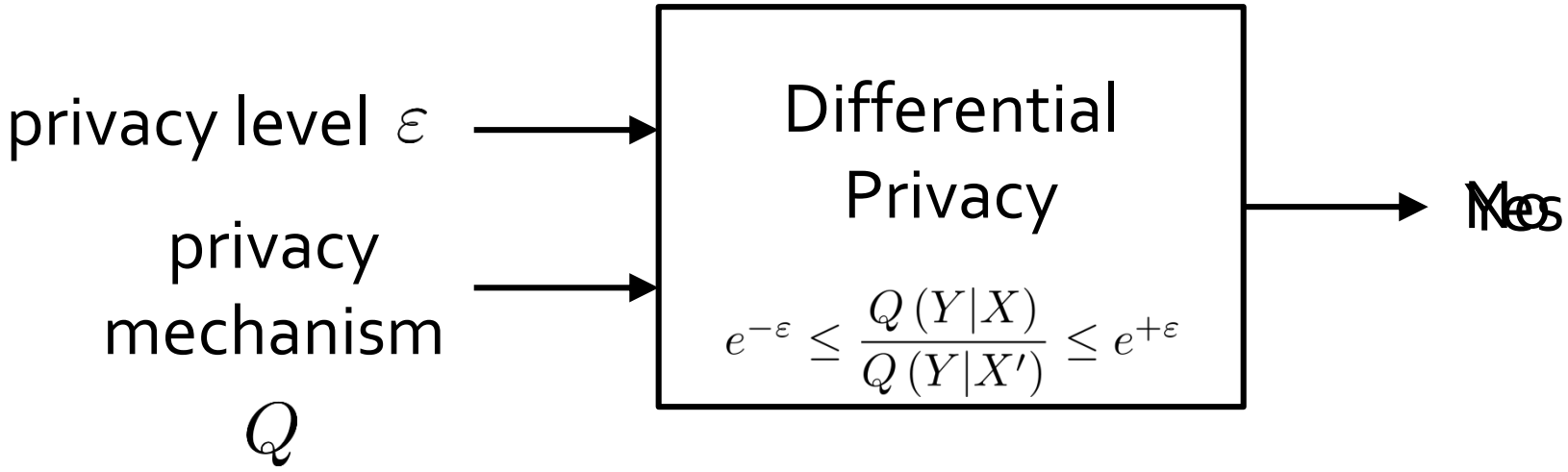
	d-Regular trees	Irregular trees	Facebook graph
snapshot 	[1]	[1, 2]	[1]
spy-based 	[3]	[3]	[3]

[1] *Spy vs. Spy: Rumor Source Obfuscation*, SIGMETRICS 2015
[2] *Rumor Source Obfuscation on Irregular Trees, to appear in SIGMETRICS 2016*
[3] *Under review*

Part 2: Data Privacy

Differential privacy

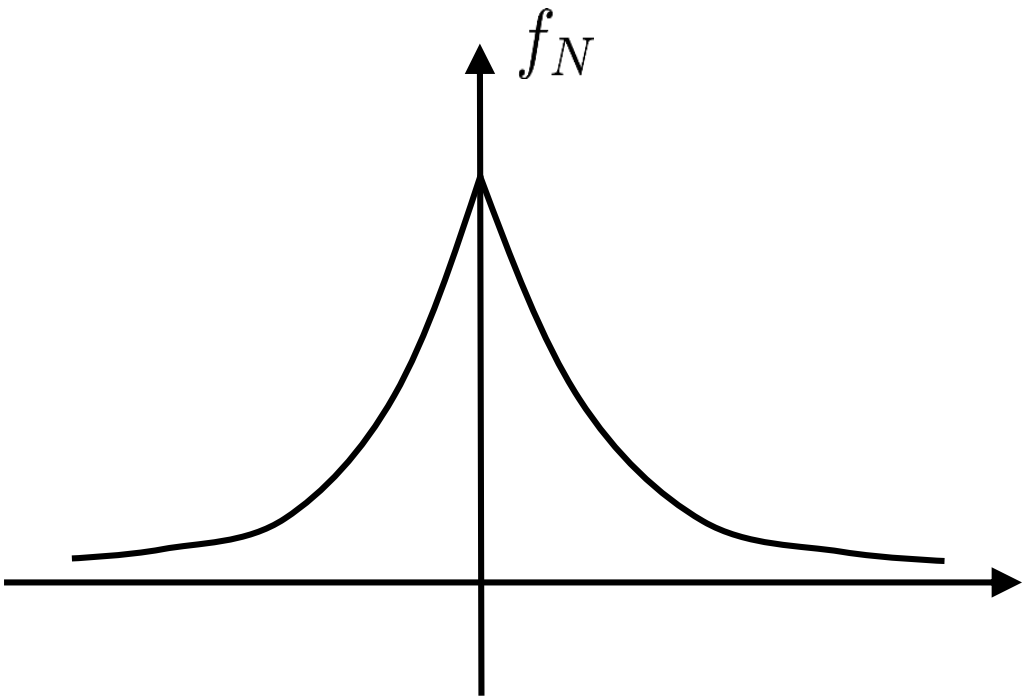
we need **context free** privacy guarantees



[*Dinur et al. 2003, Dwork et al. 2006*]

Differential privacy

Laplace Mechanism



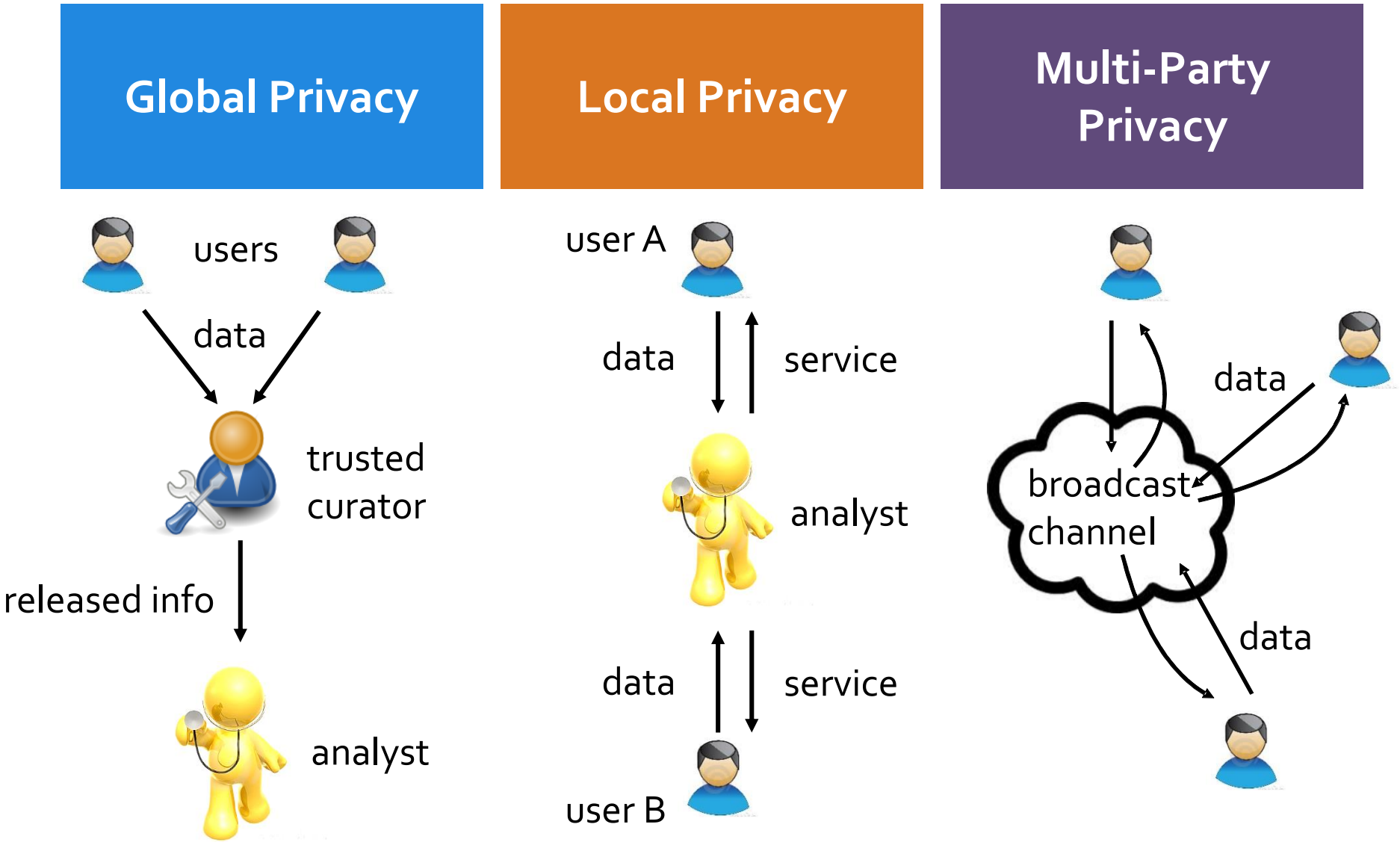
standard deviation proportional to privacy level

Privacy vs. utility

given a **privacy level**

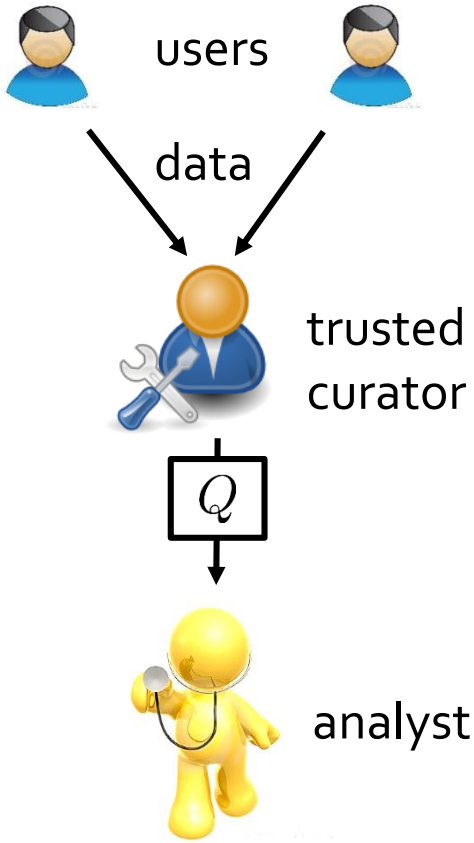
find the **privacy mechanism** that maximizes utility

Three main settings

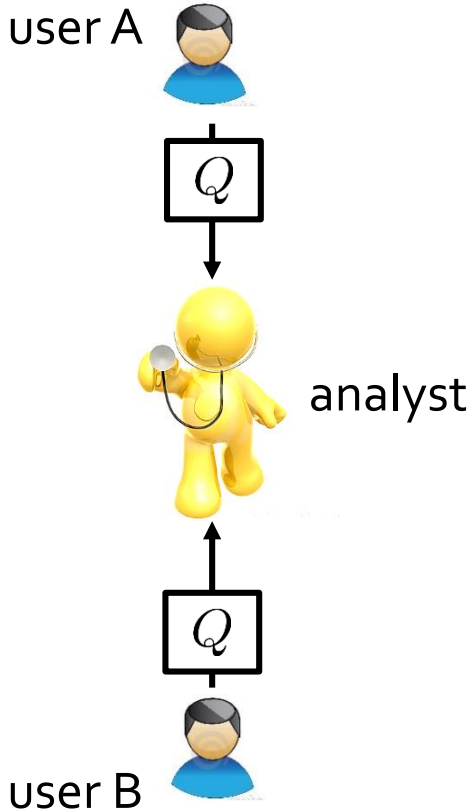


Three main settings

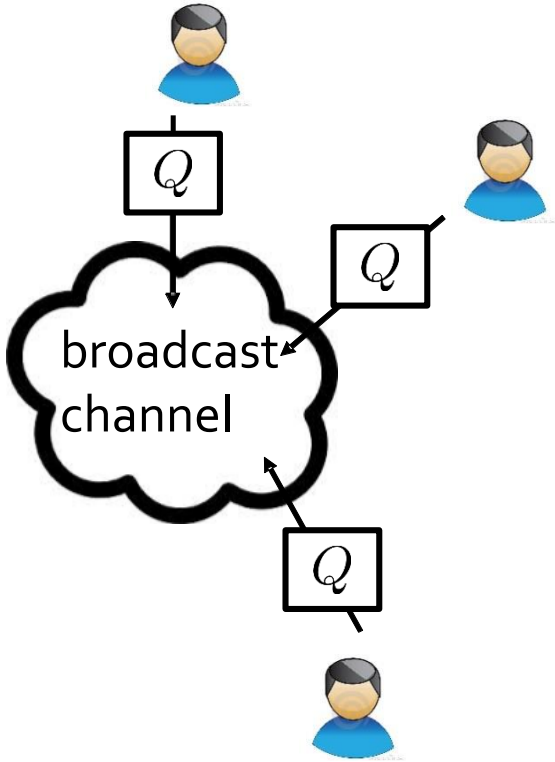
Global Privacy



Local Privacy



Multi-Party Privacy



Our results

Global Privacy	Local Privacy	Multi-Party Privacy

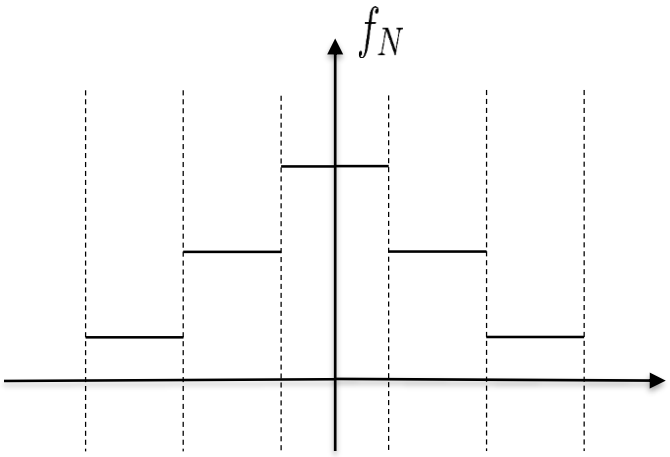
privacy mechanisms that achieve the best privacy-utility tradeoff

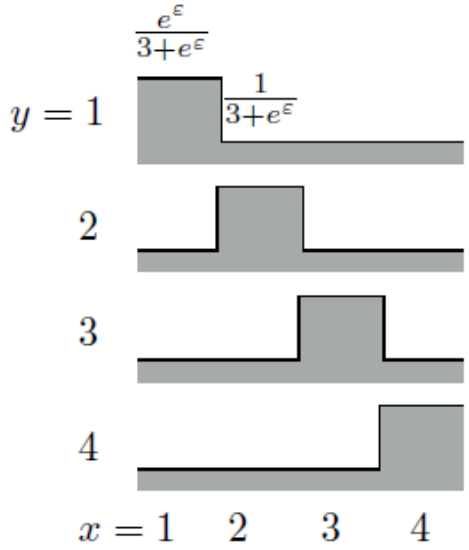
[NIPS 14, NIPS 15, ICML 15, TSTSP 15, CISS 16, JMLR 16, TIT 16]

Our results

Global Privacy | **Local Privacy** | **Multi-Party Privacy**

the optimal mechanisms in all three settings have a **staircase shape**





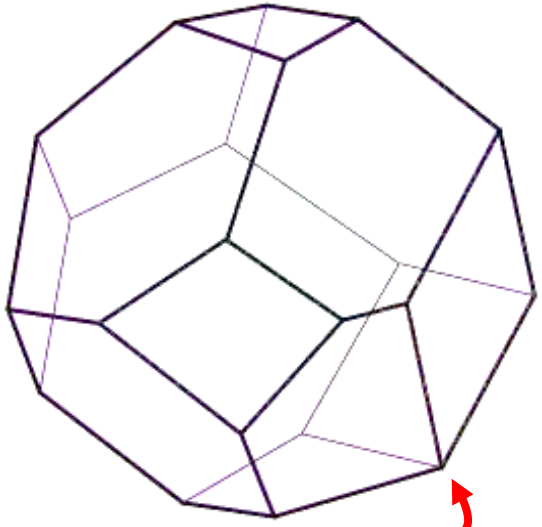
[NIPS 14, NIPS 15, ICML 15, TSTSP 15, CISS 16, JMLR 16, TIT 16]

Main tools used

Geometry of differential privacy

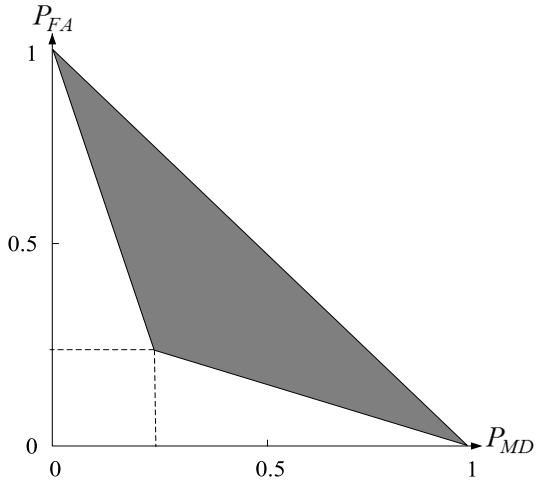
maximize utility

subject to differential privacy



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

Privacy as a region

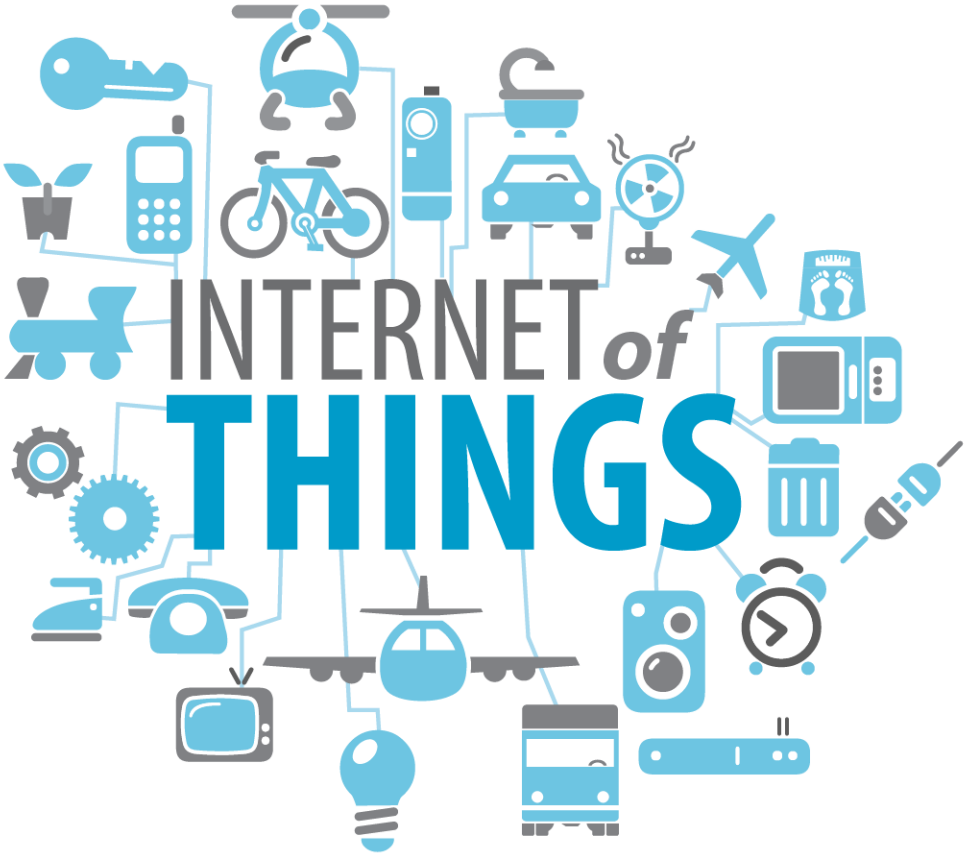


leads to **optimal** results in

[1] Composition
[2] Local Differential Privacy
[3] Multi-party differential privacy

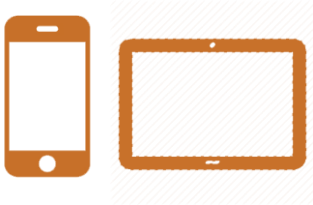
Research Vision

Unprecedented level of connectivity

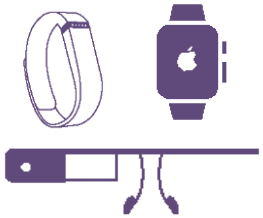


Internet of services → **Internet of things**

Smart phones and tablets



Machine to machine



Privacy preserving technologies

On-device intelligence

Ultra-low power communication

Collaborations

Pramod Viswanath

Sewoong Oh

Kannan Ramchandran

Pavan Hanumolu

Songbin Gong

Giulia Fanti

Keith Bonawitz

Brendan McMahan

Daniel Ramage

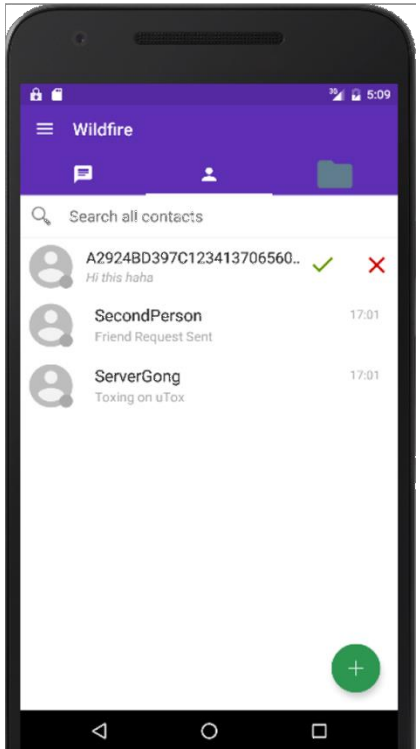
Haitham Hassanieh

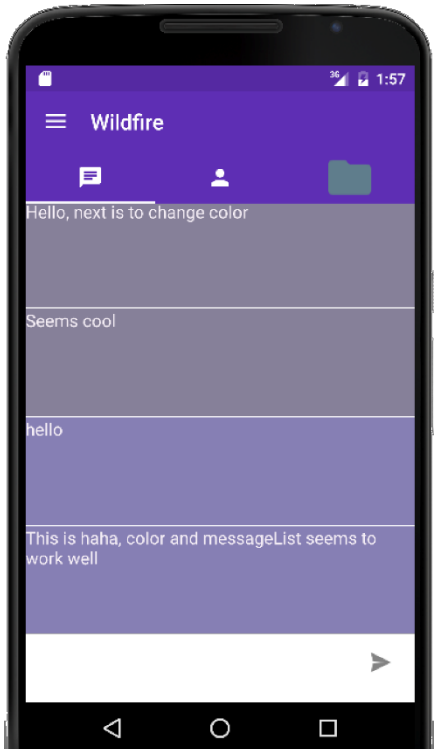
THANK YOU!

BACK UP SLIDES

Wildfire: P2P anonymous messaging

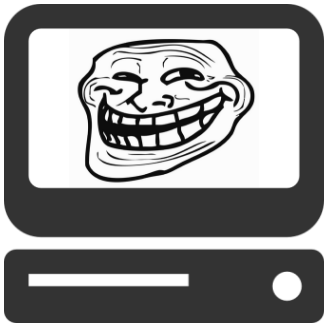
TOX



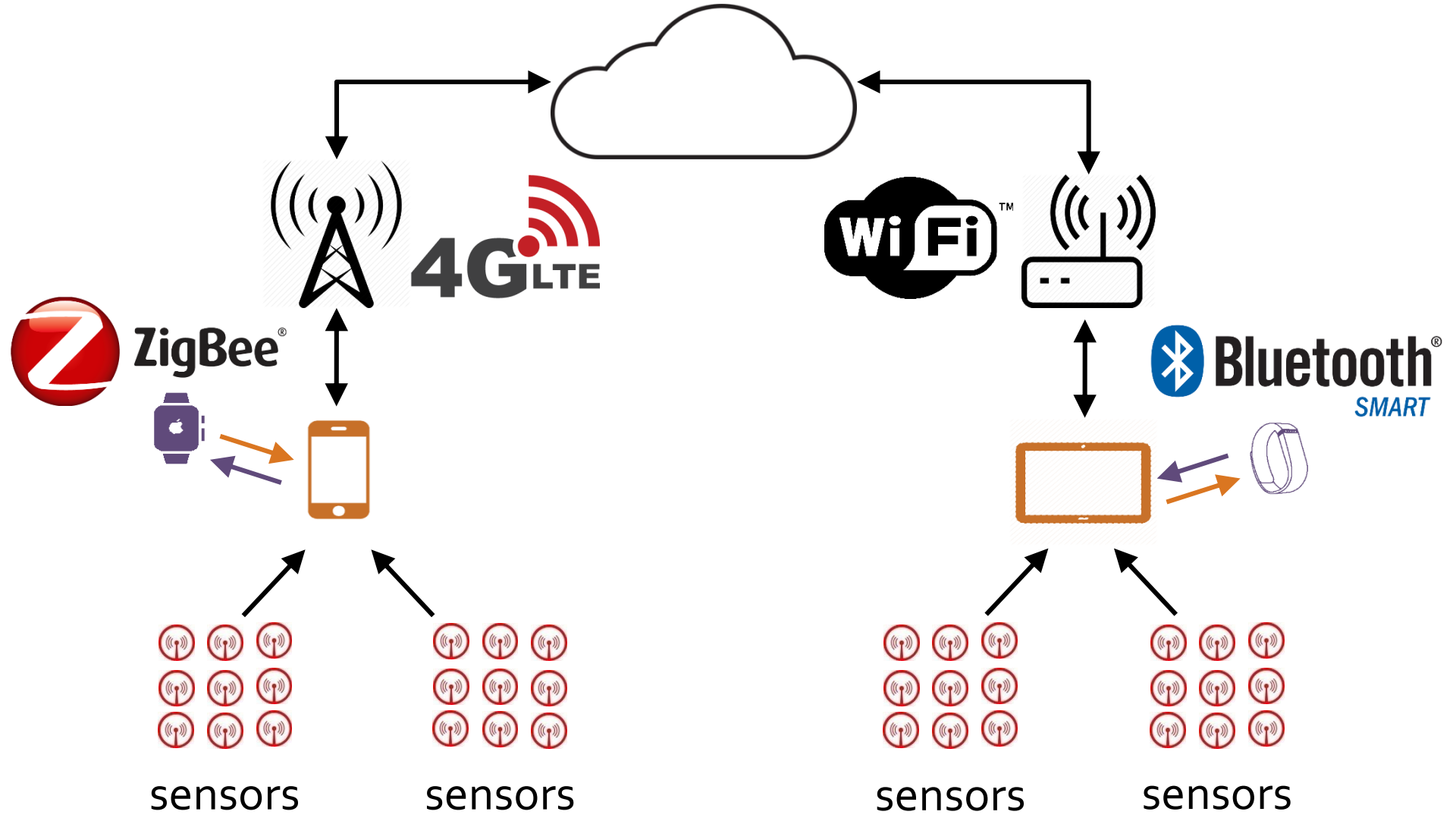


Namespace resolution

Cyberbullying



Ultra-low power communication

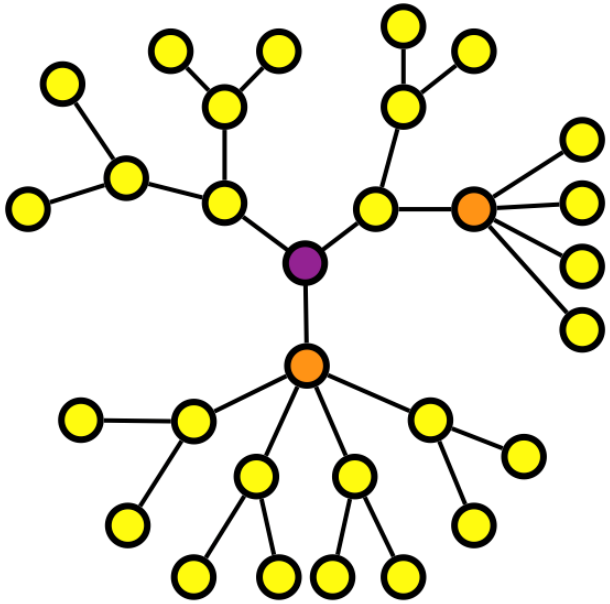


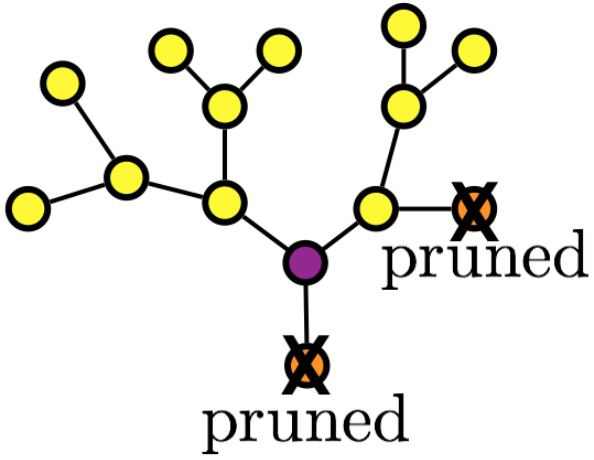
one-way, event driven, asynchronous communication

$$\min_{v \in \text{leaves}} \prod_{v \in P(v, v_T)} d_v \approx (d_{min} - 1)^{T/2}$$

$$d_v = \begin{cases} 3 & w.p. \quad 0.7 \\ 5 & w.p. \quad 0.3 \end{cases}$$

$$d_v = \begin{cases} 3 & w.p. \quad 0.7 \\ 1 & w.p. \quad 0.3 \end{cases}$$

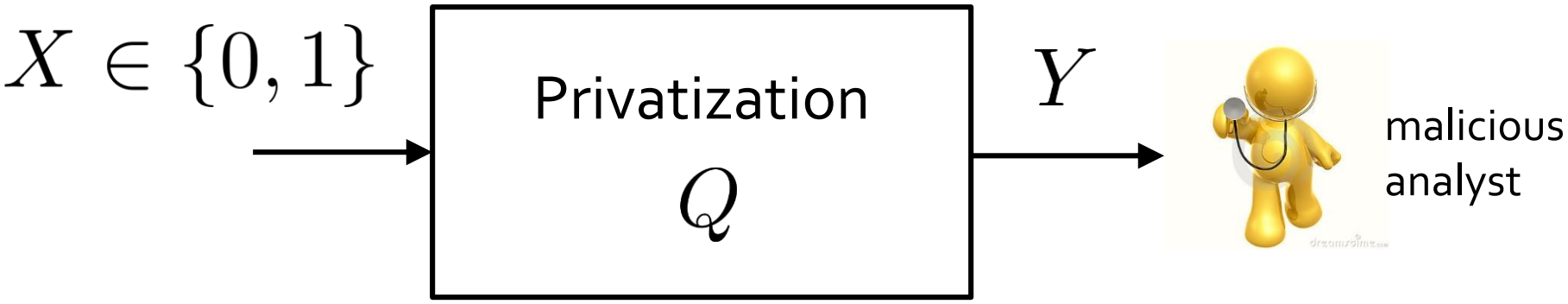


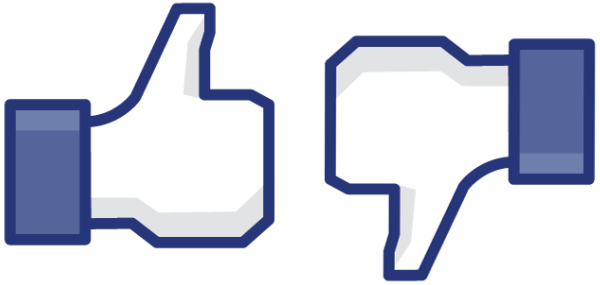


0.7 3

If $p_{min}(d_{min} - 1) > 1$ then the pruned process survives.

Binary data

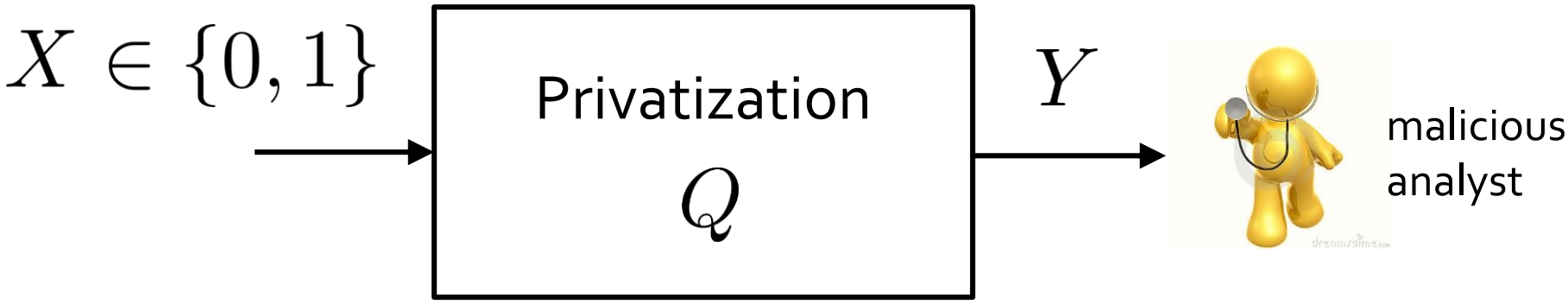




thumbs up
thumbs down

Yes/No
questions

Differential privacy



$$e^{-\varepsilon} \leq \frac{\mathbb{P}(Y|X=1)}{\mathbb{P}(Y|X=0)} \leq e^{+\varepsilon}$$

ε controls the level of privacy
large ε, low privacy
small ε, high privacy

Randomized response

answer truthfully

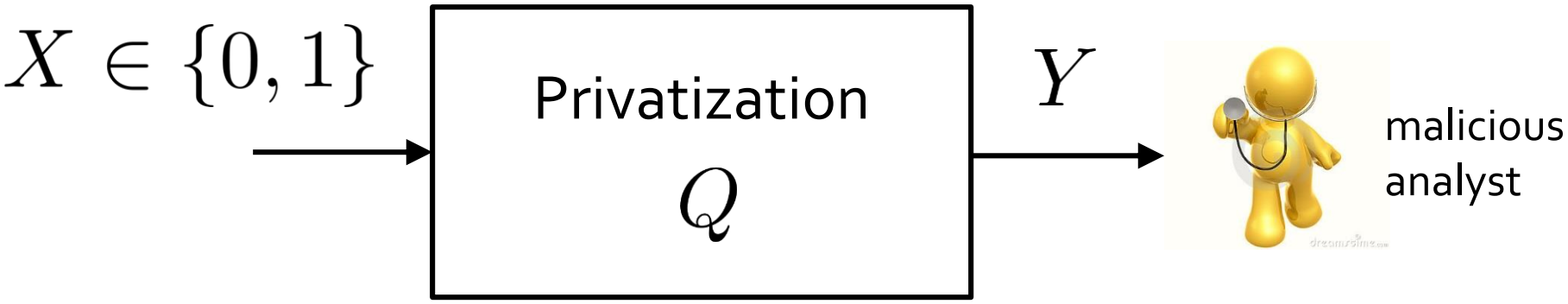
answer wrongly

$$\frac{e^{\varepsilon}}{e^{\varepsilon}+1}$$

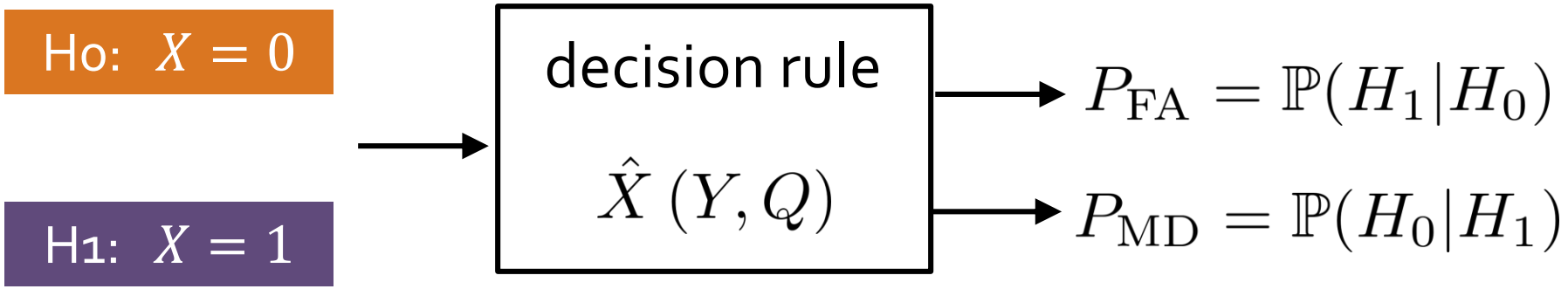
$$\frac{1}{e^{\varepsilon}+1}$$

[*Warner* 1965]

False alarm, missed detection tradeoff

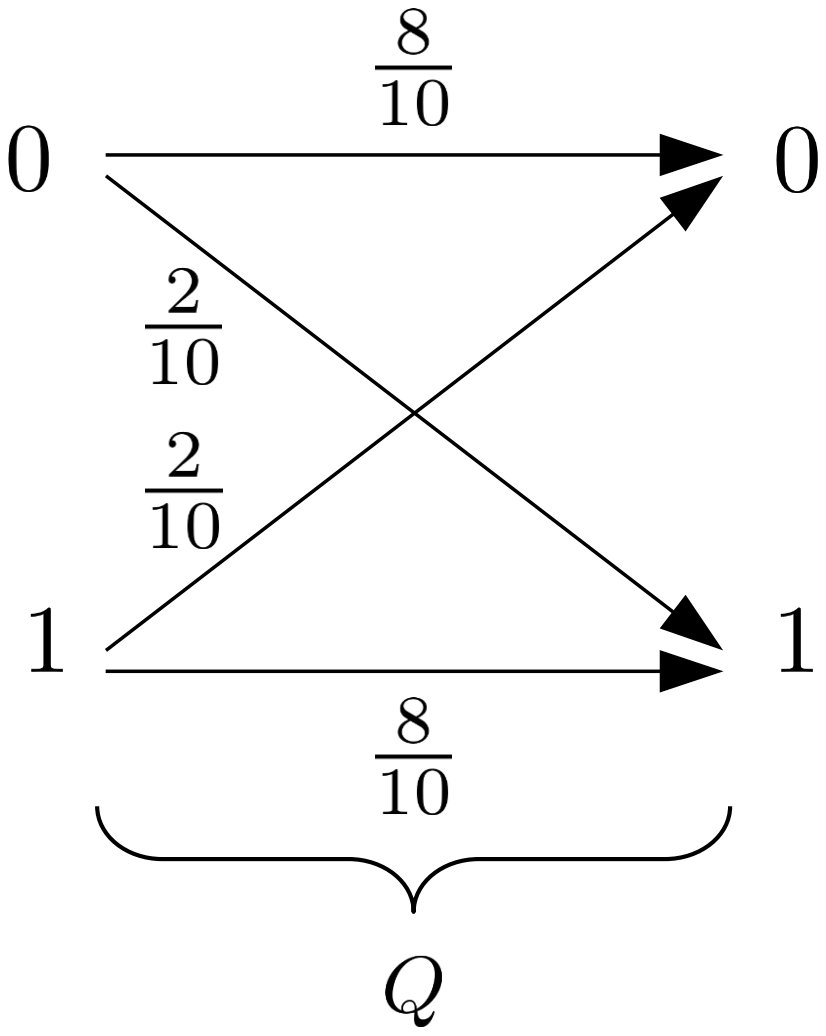


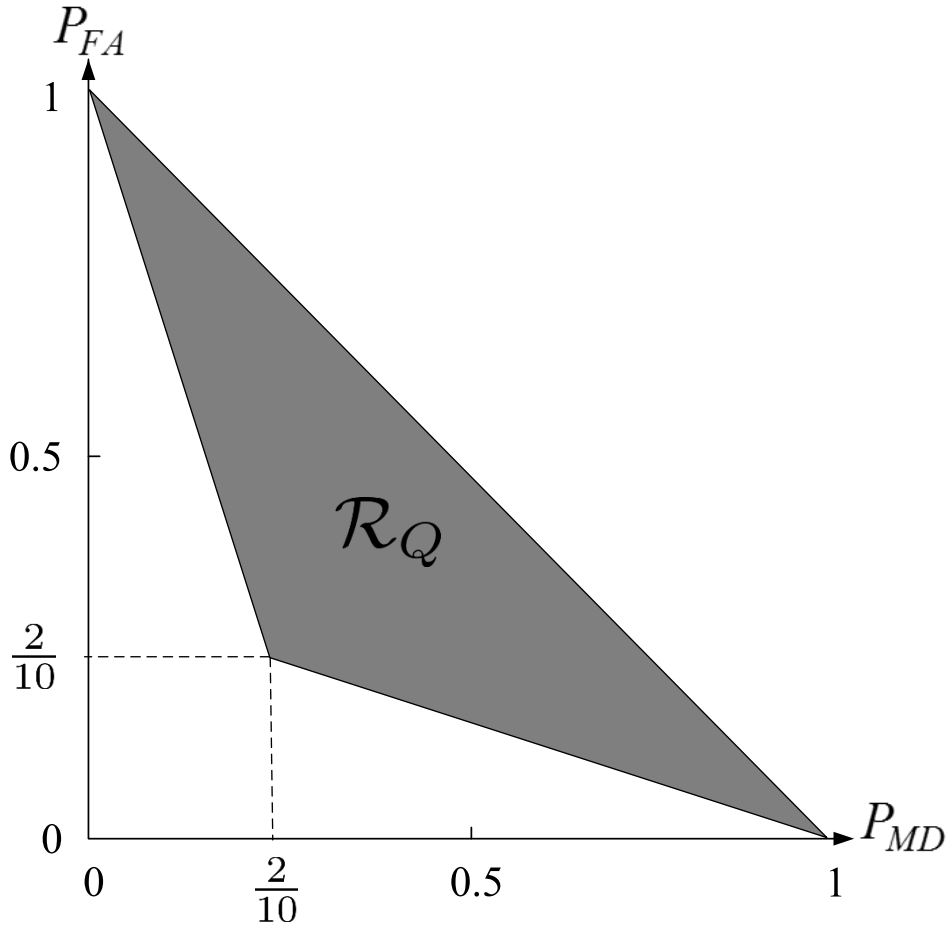
hypothesis test



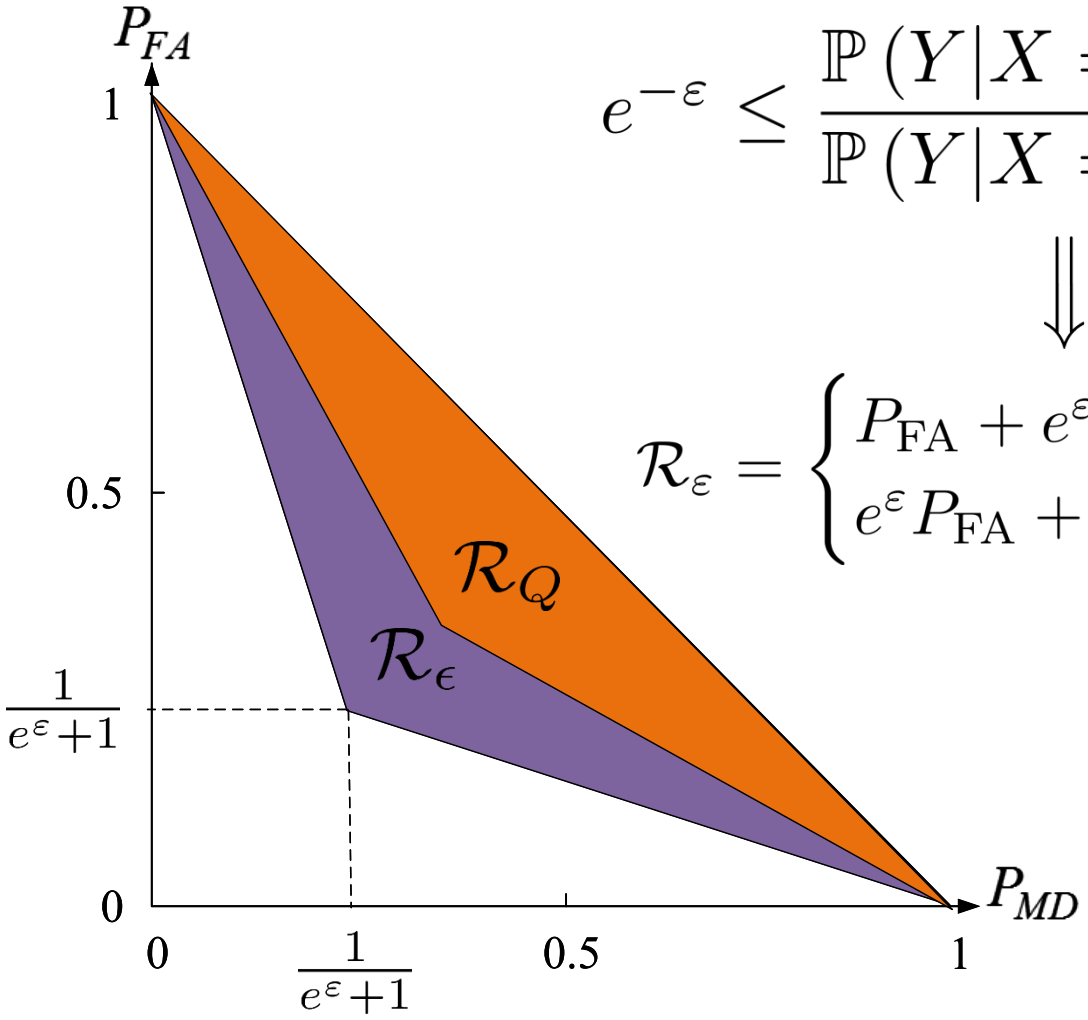
$\mathcal{R}_Q =$ all achievable pairs of $(P_{\mathrm{FA}}, P_{\mathrm{MD}})$

False alarm, missed detection region





Differential privacy as a privacy region



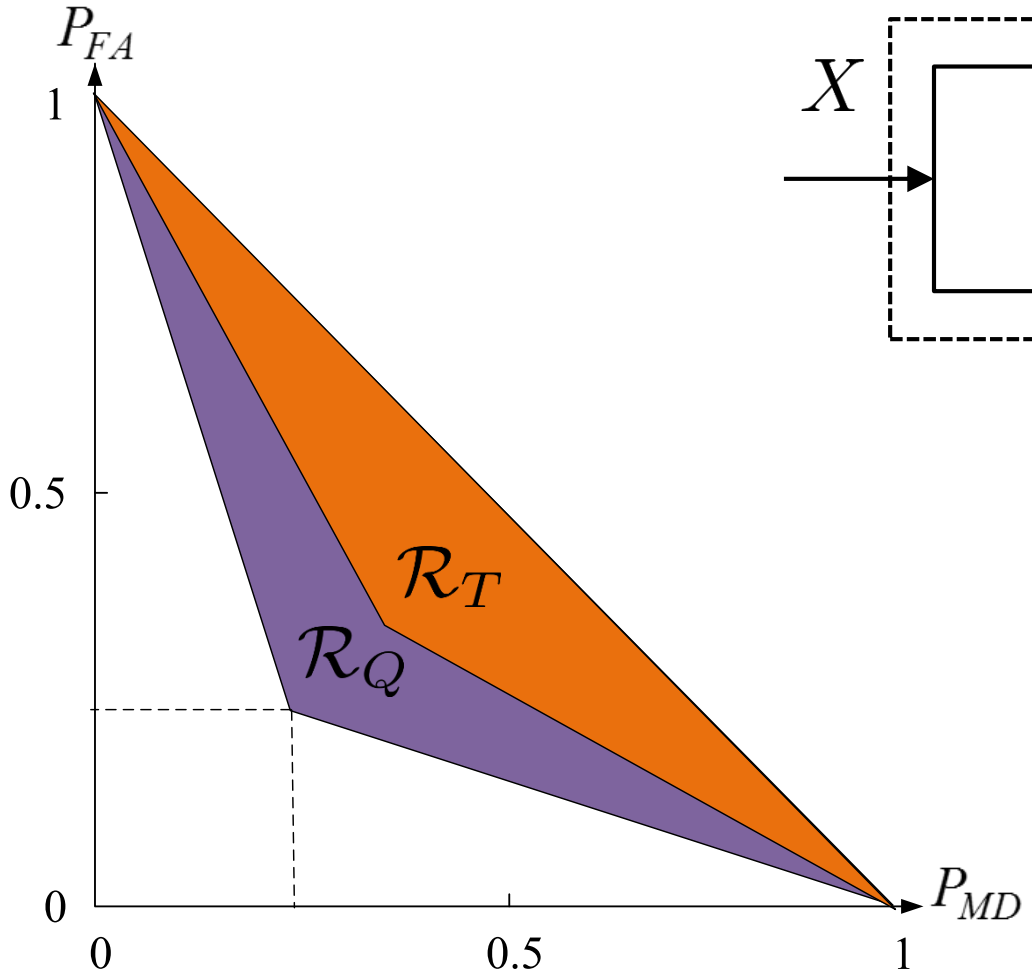
$$e^{-\varepsilon} \leq \frac{\mathbb{P}(Y|X=1)}{\mathbb{P}(Y|X=0)} \leq e^{+\varepsilon}$$

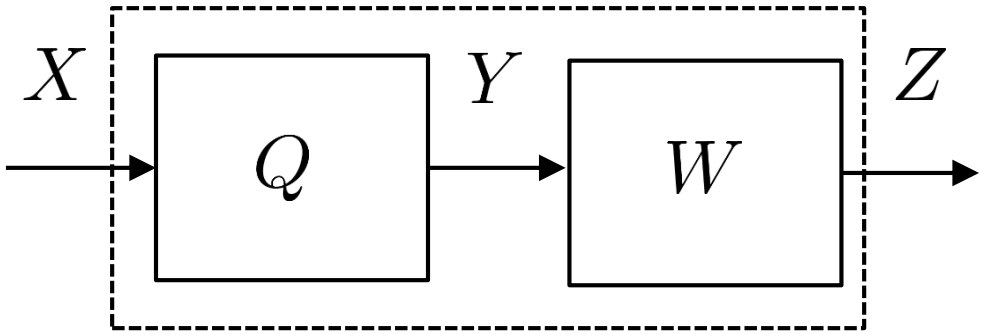
$$\Downarrow$$

$$\mathcal{R}_\varepsilon = \begin{cases} P_{\mathrm{FA}} + e^{\varepsilon} P_{\mathrm{MD}} \geq 1 \\ e^{\varepsilon} P_{\mathrm{FA}} + P_{\mathrm{MD}} \geq 1 \end{cases}$$

Q is differentially private if and only if $\mathcal{R}_Q \subseteq \mathcal{R}_\varepsilon$

Data processing inequality (DPI)





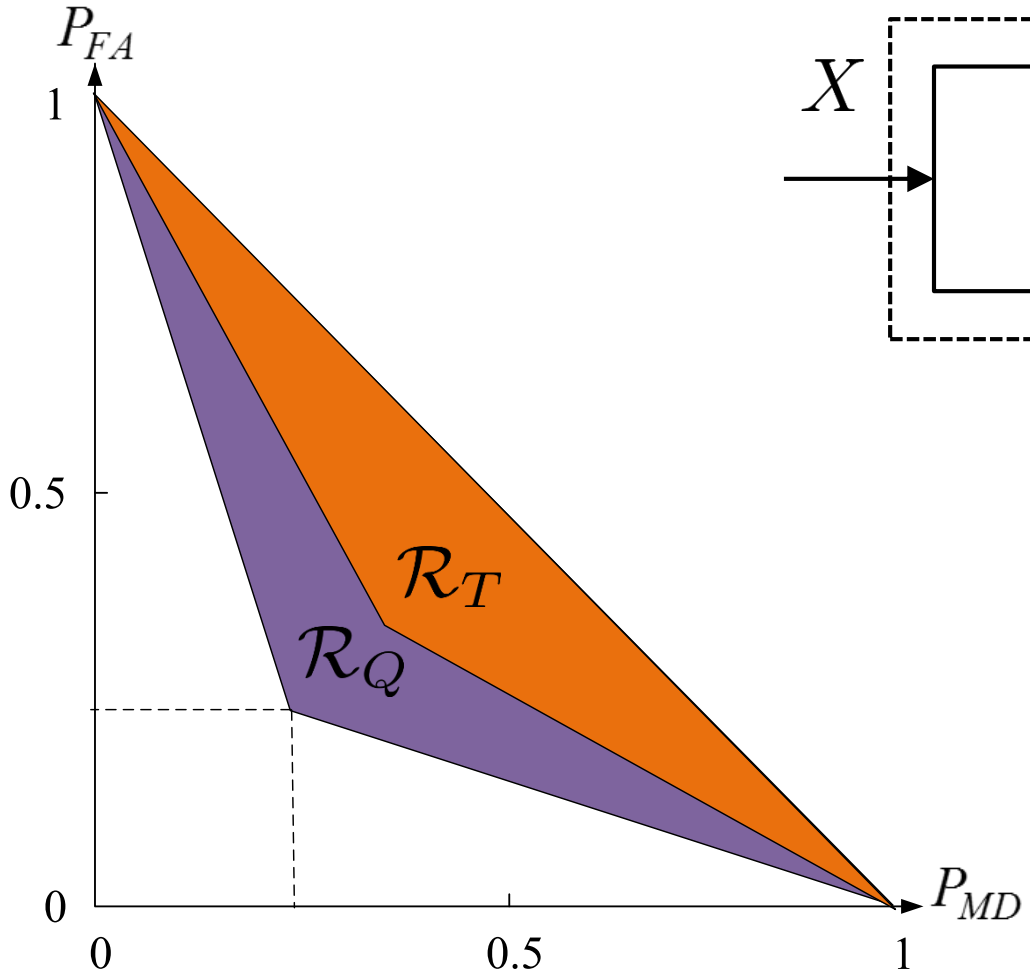
$$T$$

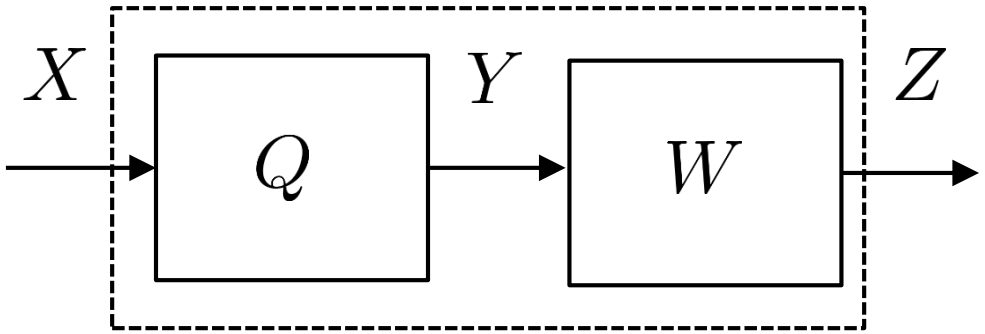
$$\Downarrow$$

$$\mathcal{R}_T \subseteq \mathcal{R}_Q$$

if Q is differentially private then T is differentially prate

Converse to DPI [Blackwell 1953]





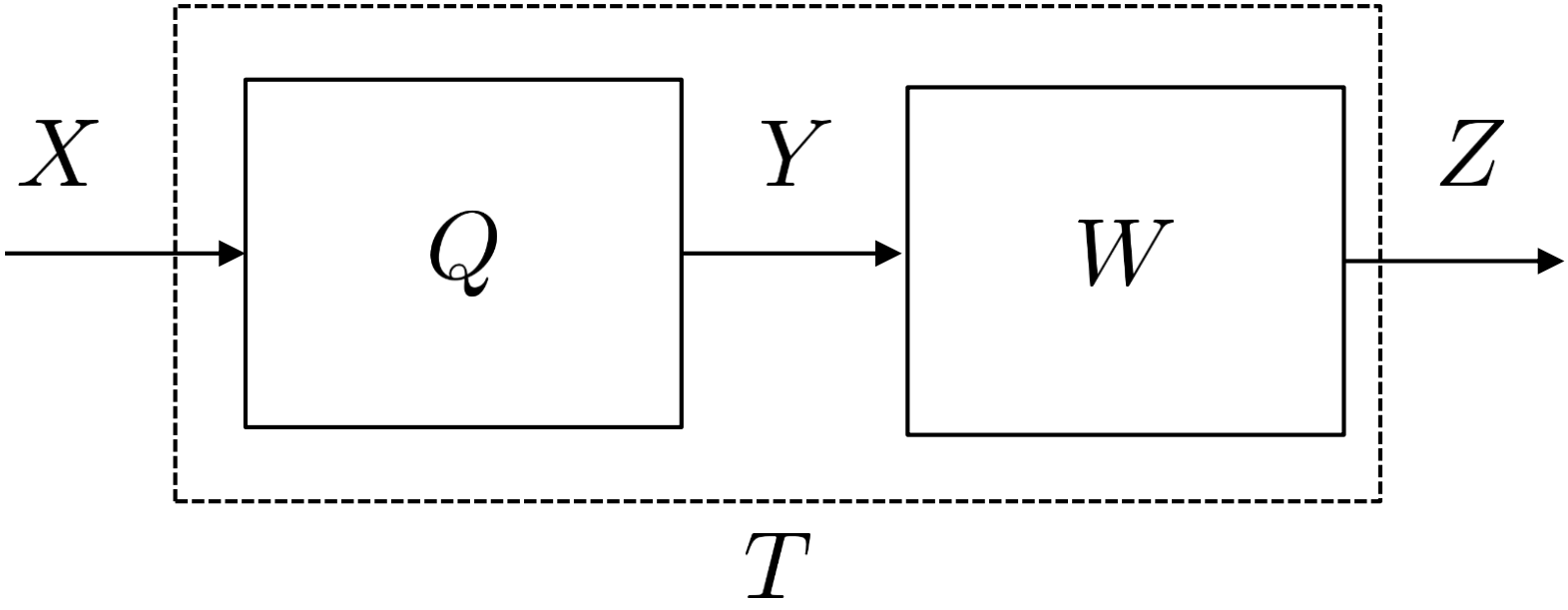
T

$\Uparrow$

$\mathcal{R}_T \subseteq \mathcal{R}_Q$

if $\mathcal{R}_T \subseteq \mathcal{R}_Q$ then there exits a W such that $T = Q \circ W$

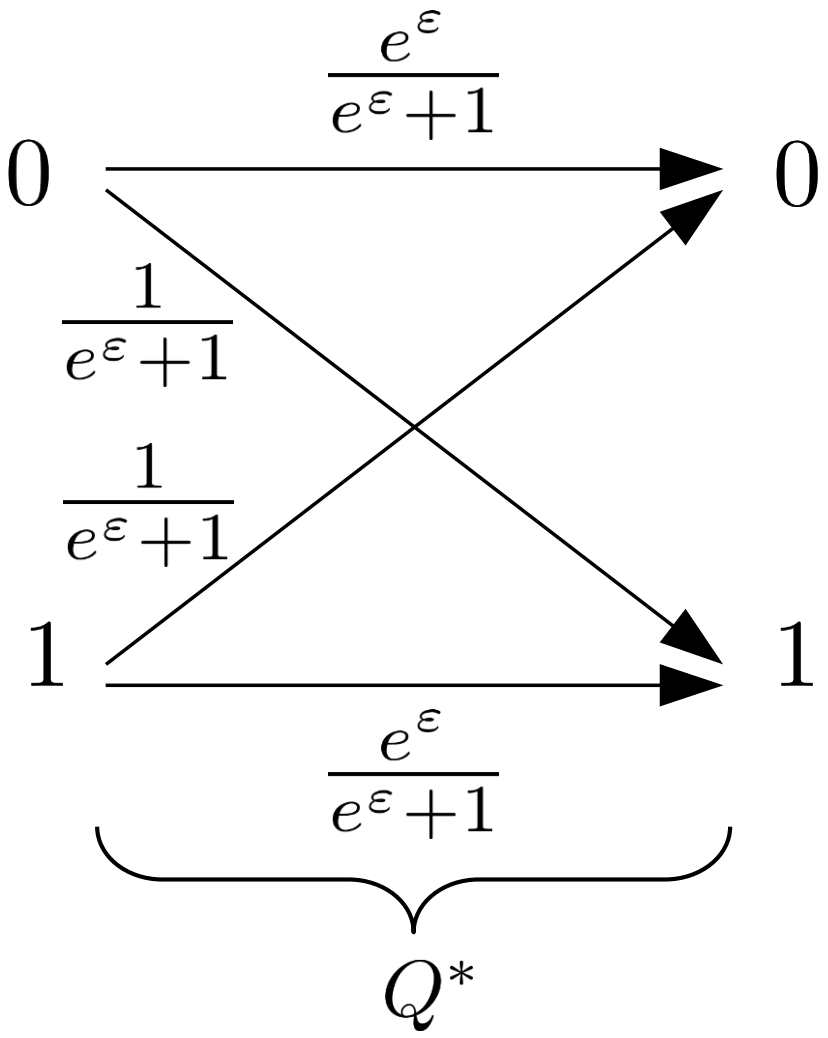
Utility functions obeying DPI

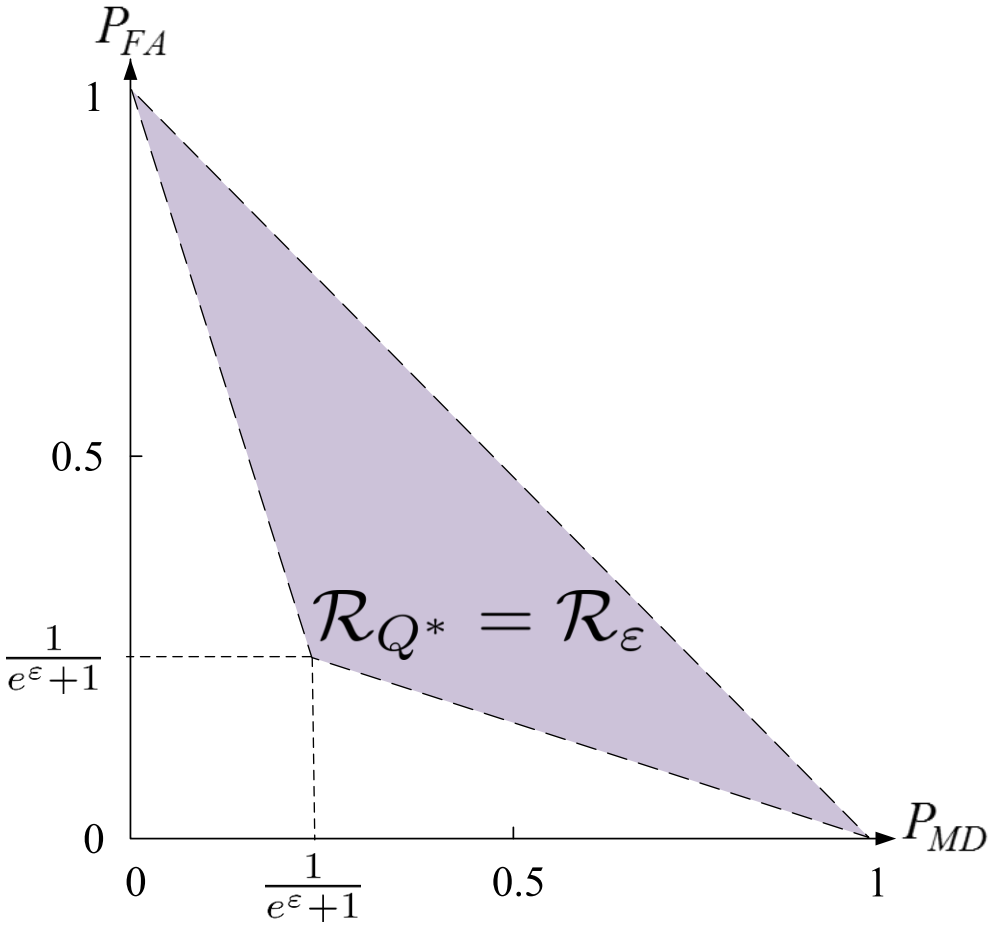


$$U(Q) \geq U(T)$$

further randomization can only reduce utility

Dominant mechanism for DP





optimal for all privacy levels & all utilities obeying DPI