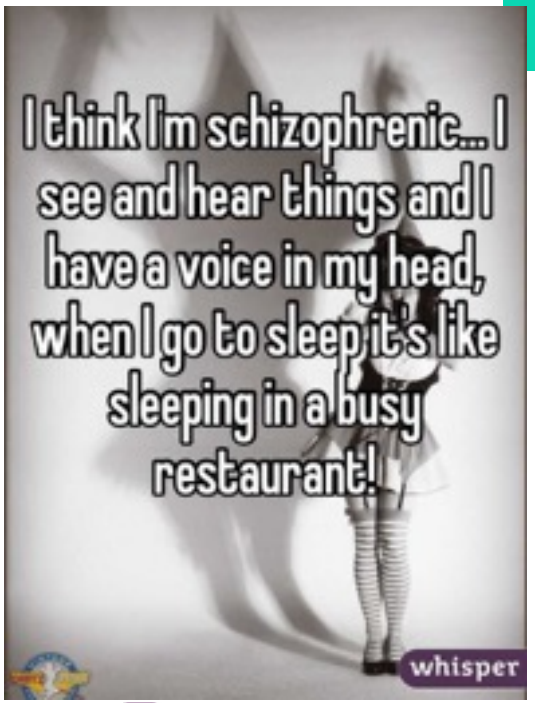
Hiding the Rumor Source in Anonymous Messaging

Sewoong Oh

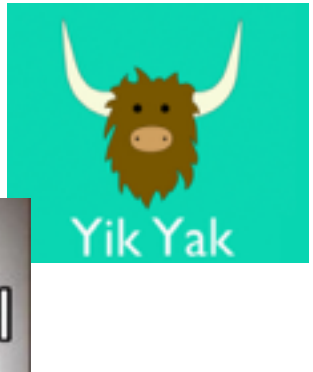
University of Illinois at Urbana-Champaign

Giulia Fanti Peter Kairouz K. Ramchandran P. Viswanath

Anonymous Social Media provide meta-data privacy

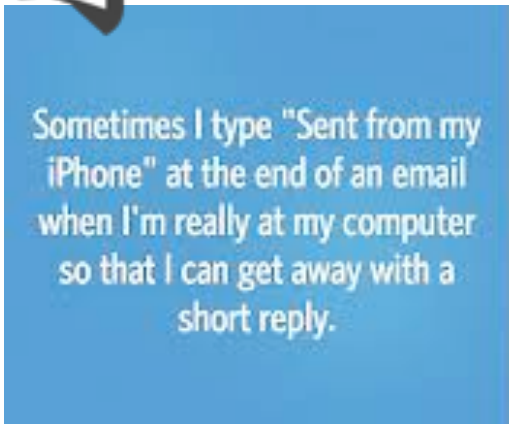


whisper

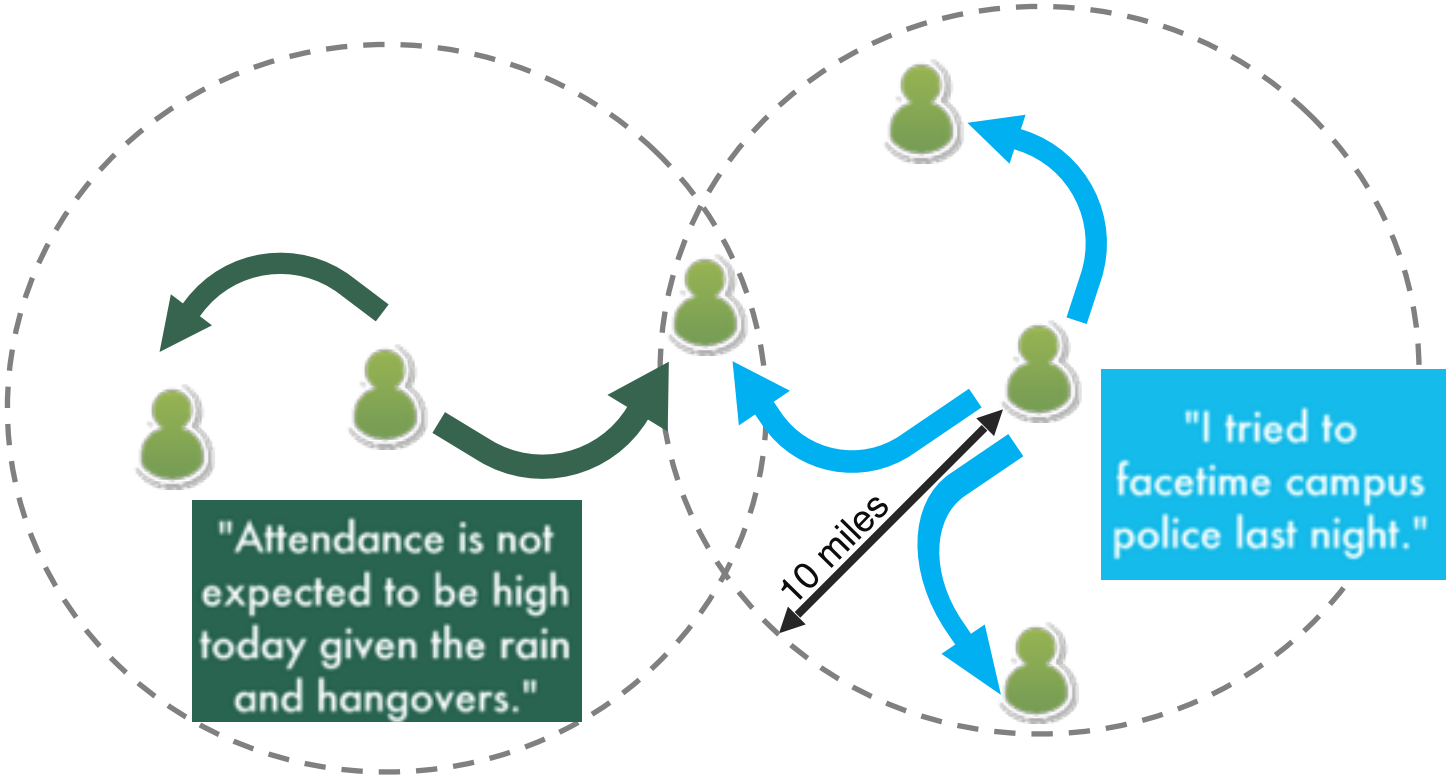


Rooms

secret



Existing anonymous messaging apps

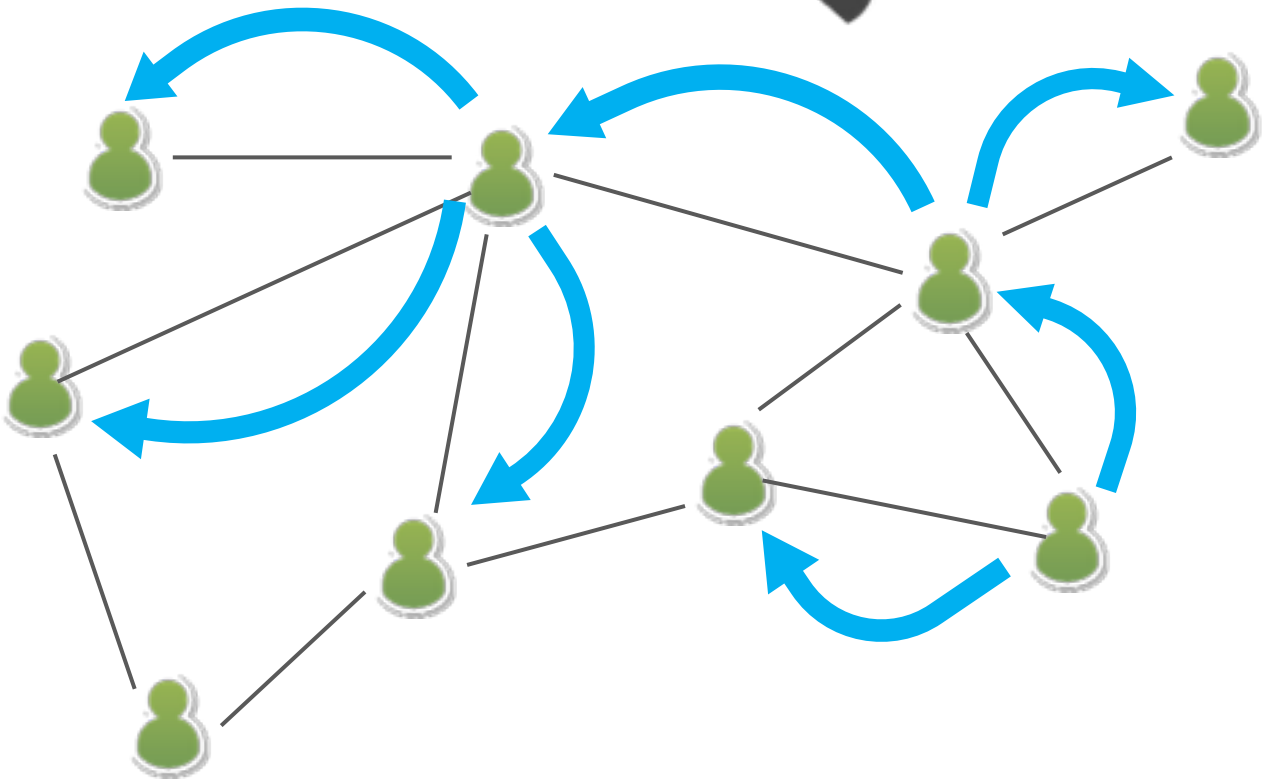


Threat is real

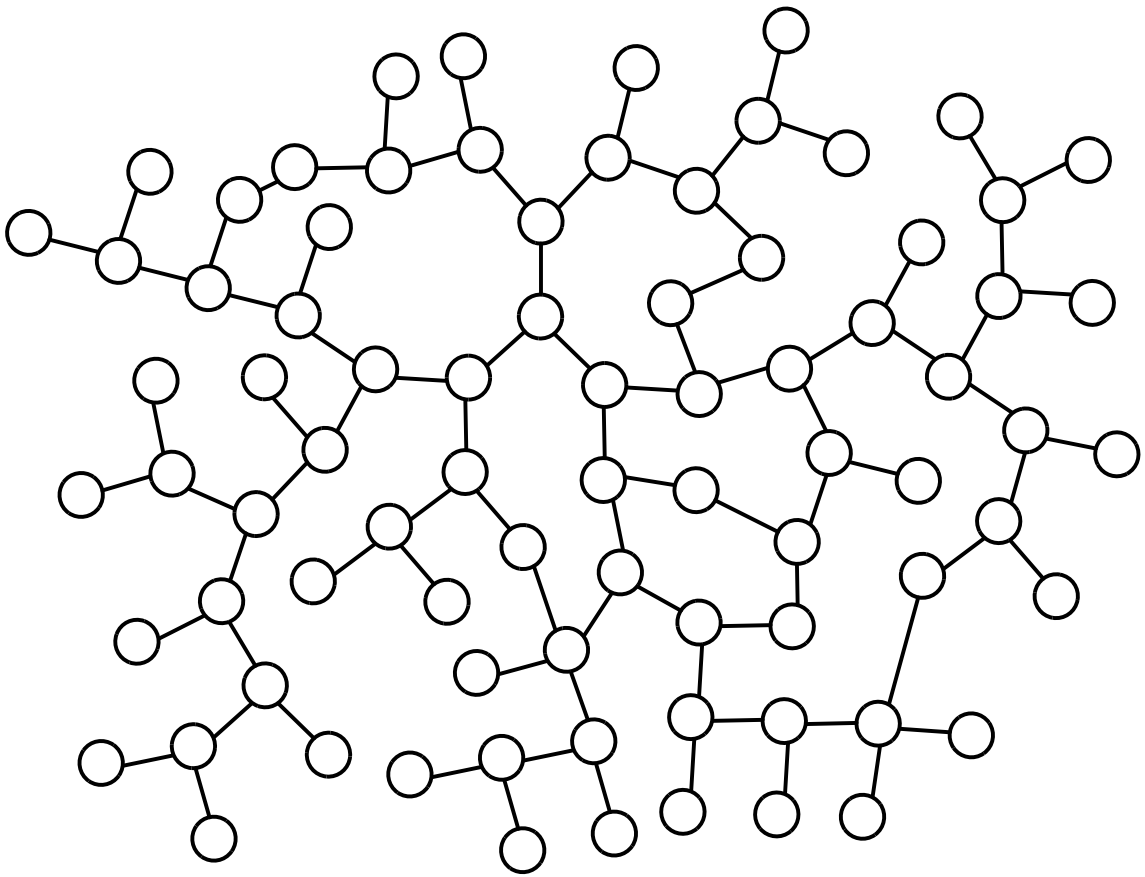
theguardian

Anonymous messaging meets social filtering

secret

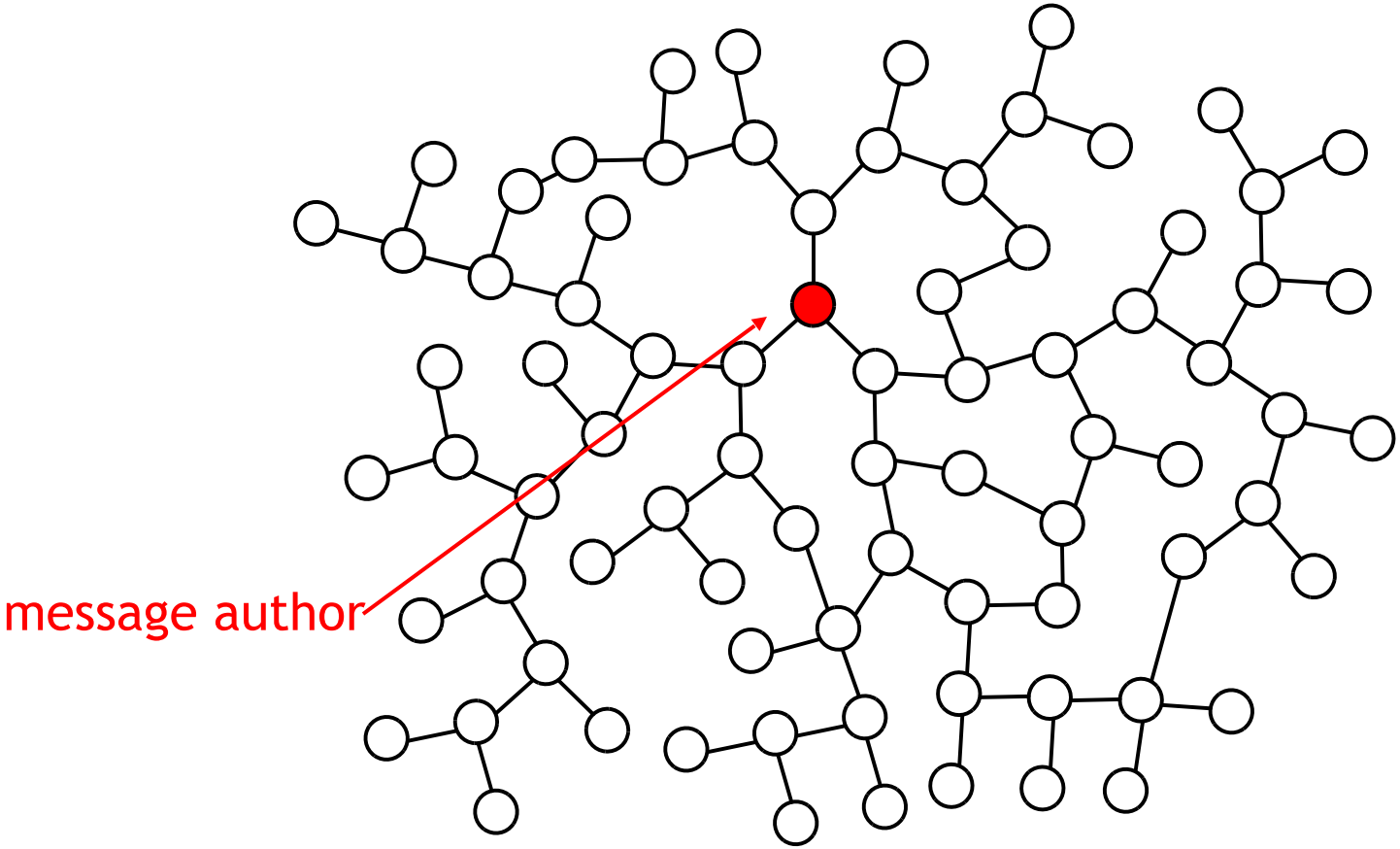


Diffusion of rumor/contagion

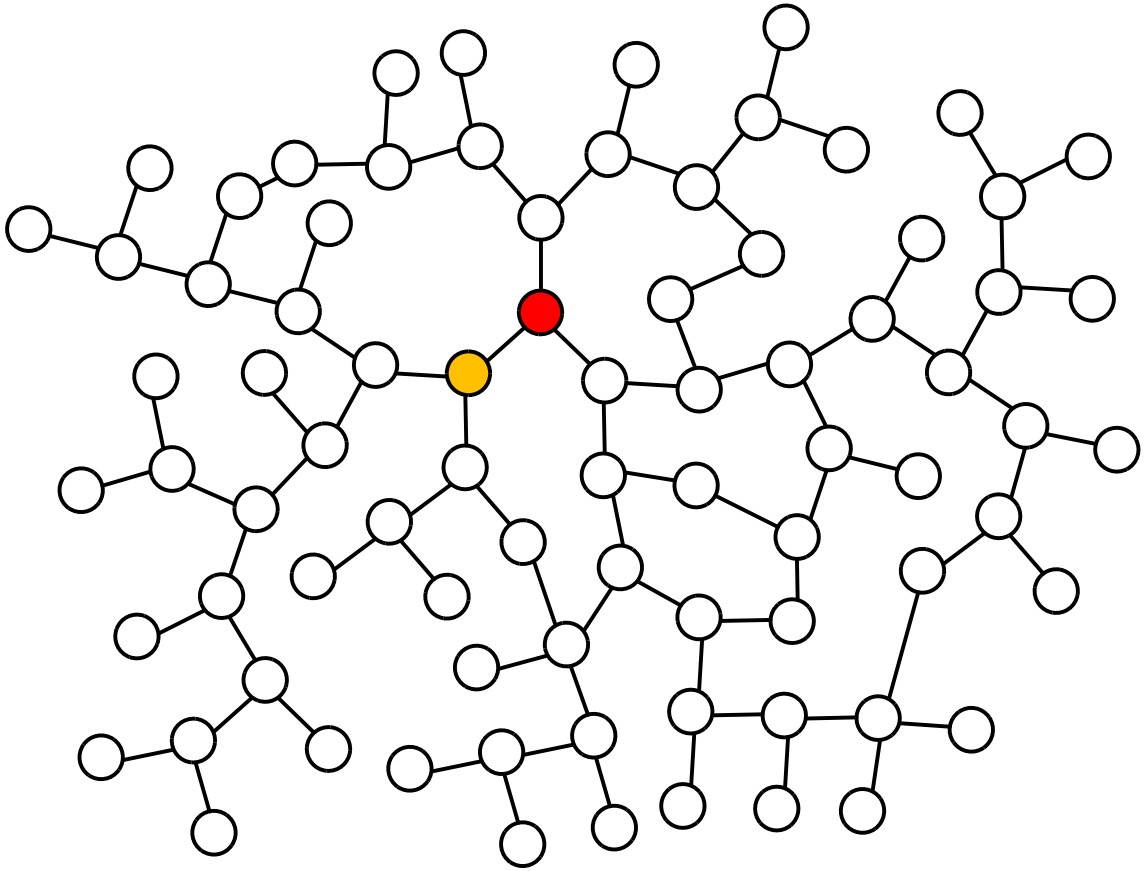


Social network/contact network

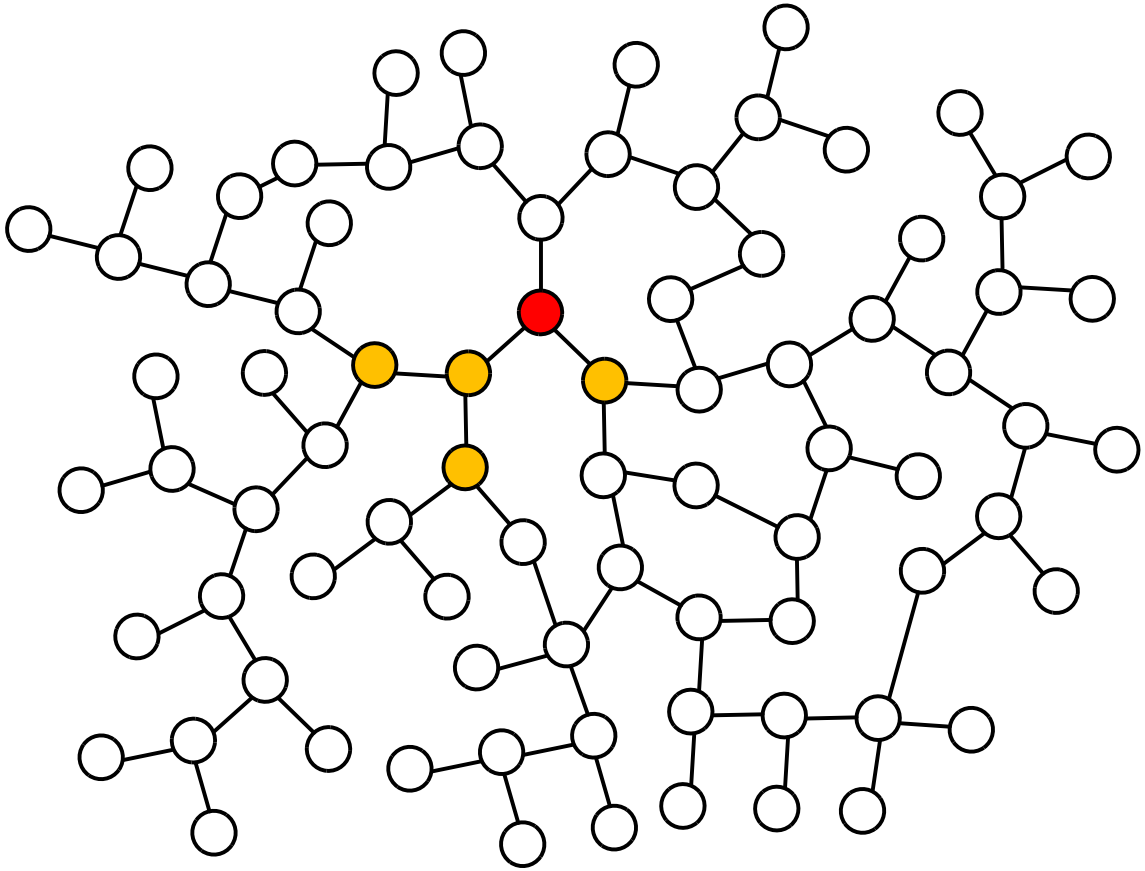
Diffusion of rumor/contagion



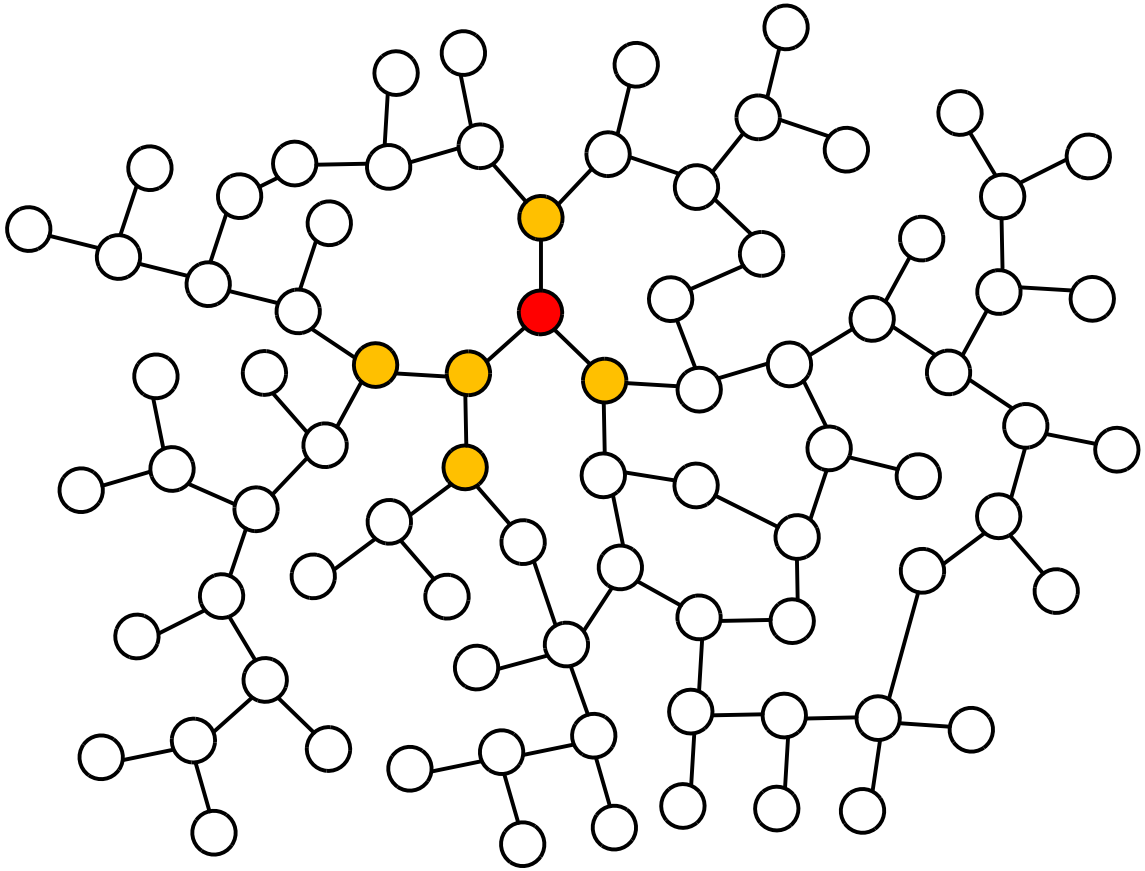
Diffusion of rumor/contagion



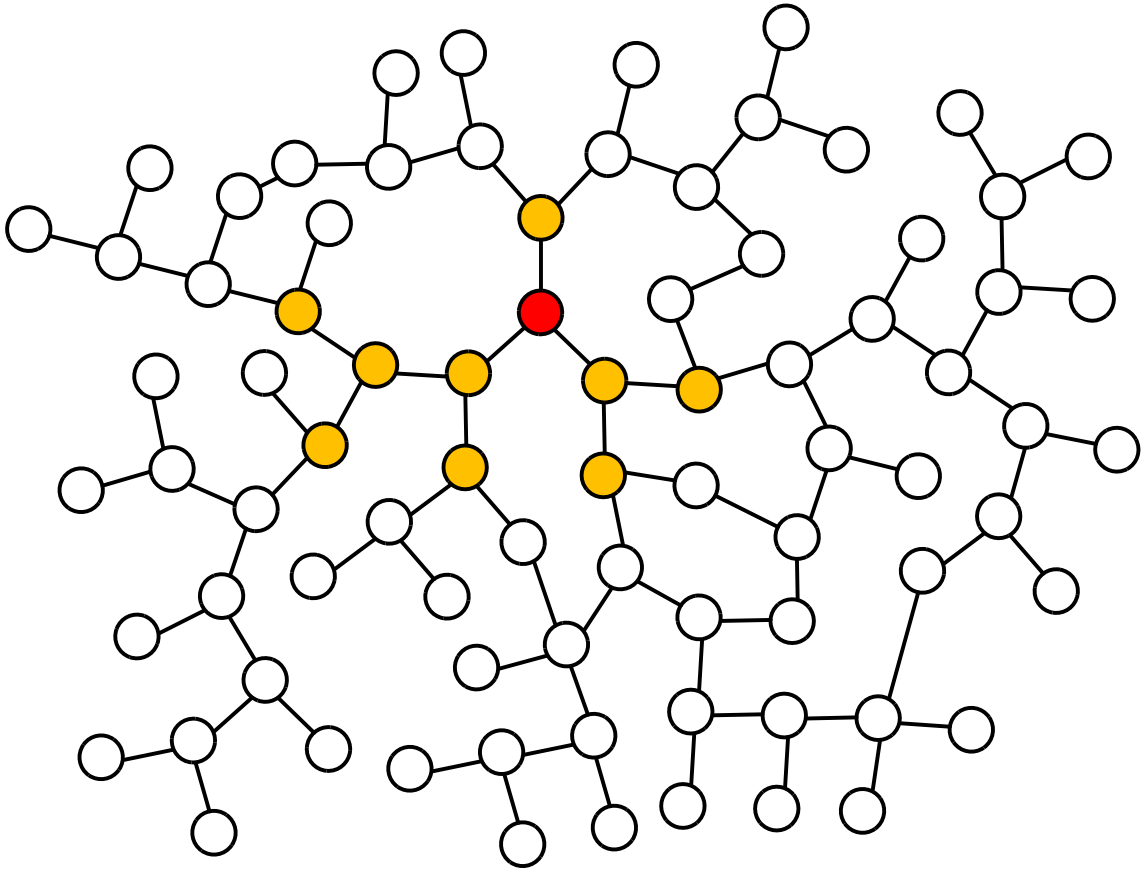
Diffusion of rumor/contagion



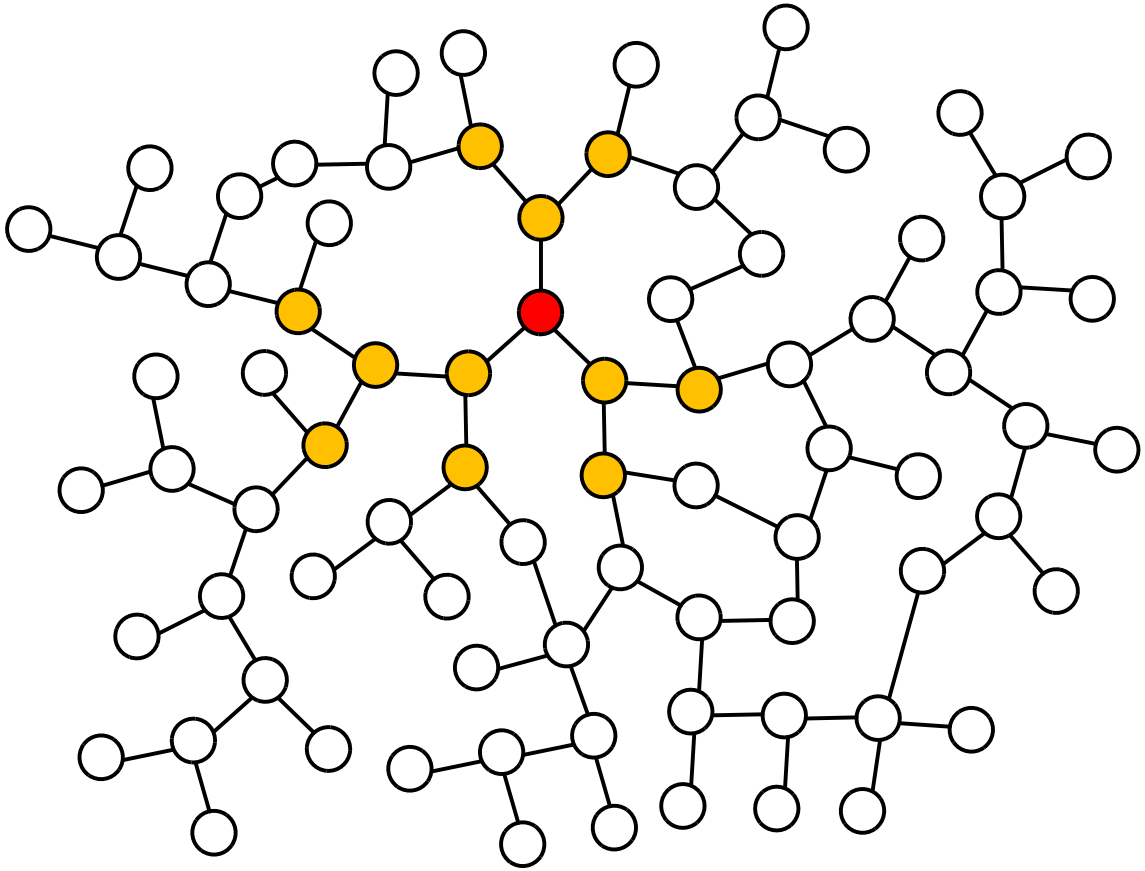
Diffusion of rumor/contagion



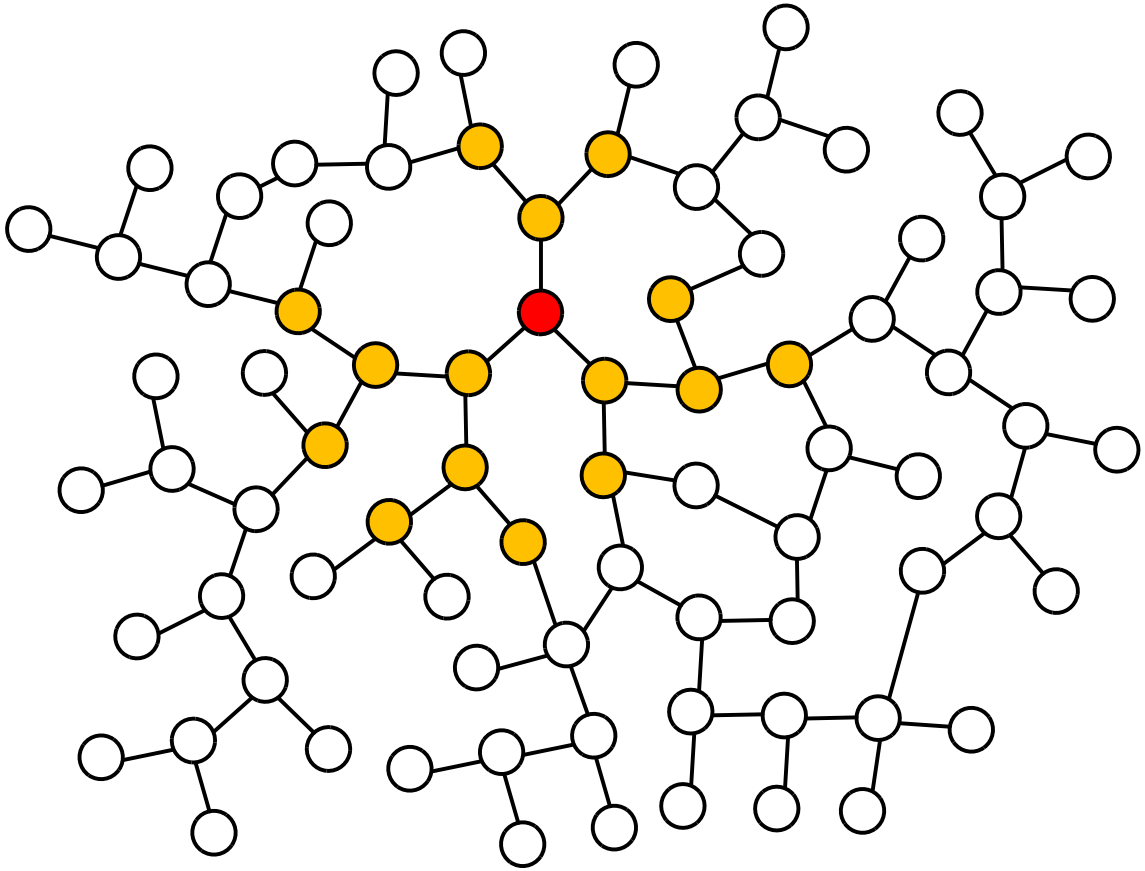
Diffusion of rumor/contagion



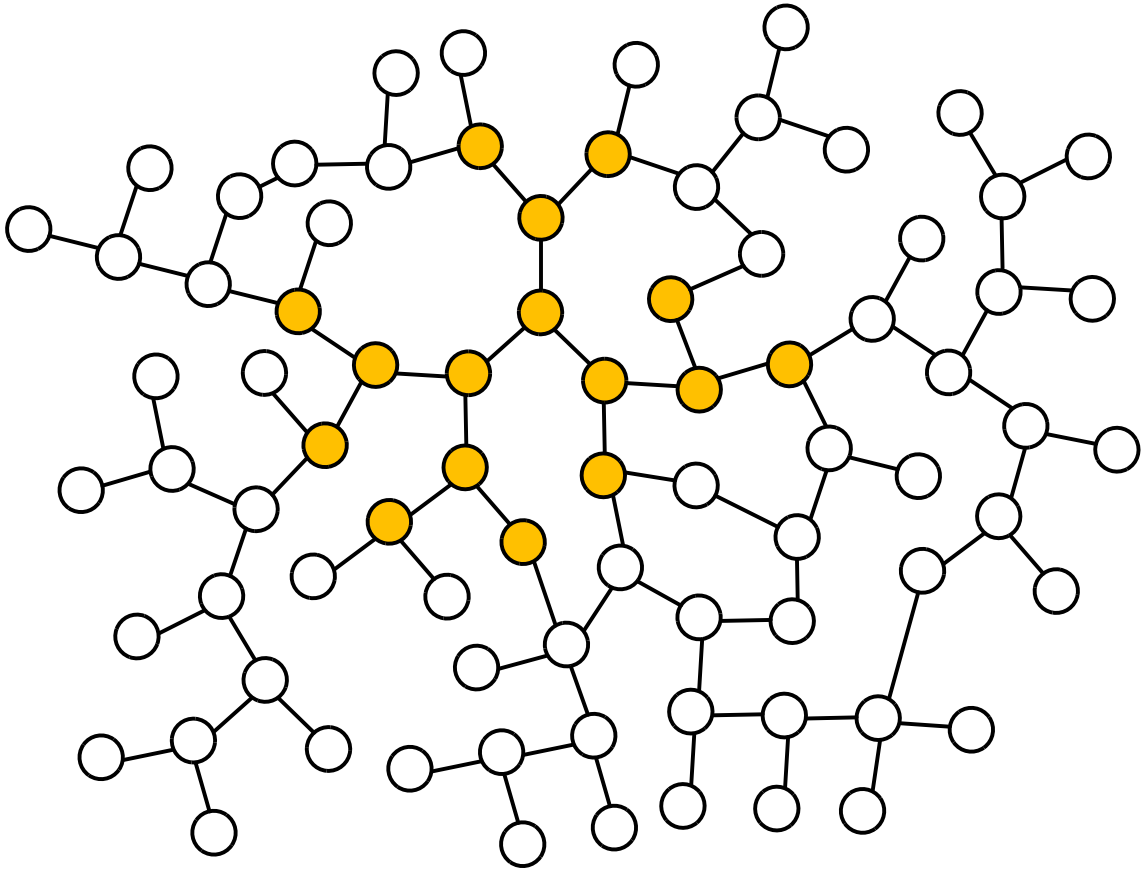
Diffusion of rumor/contagion



Diffusion of rumor/contagion

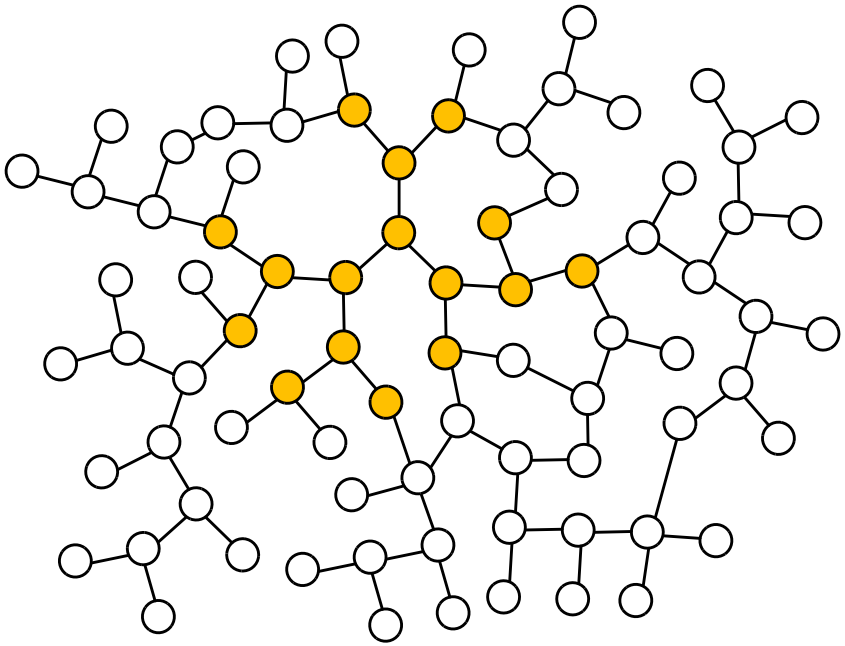


Rumor source detection

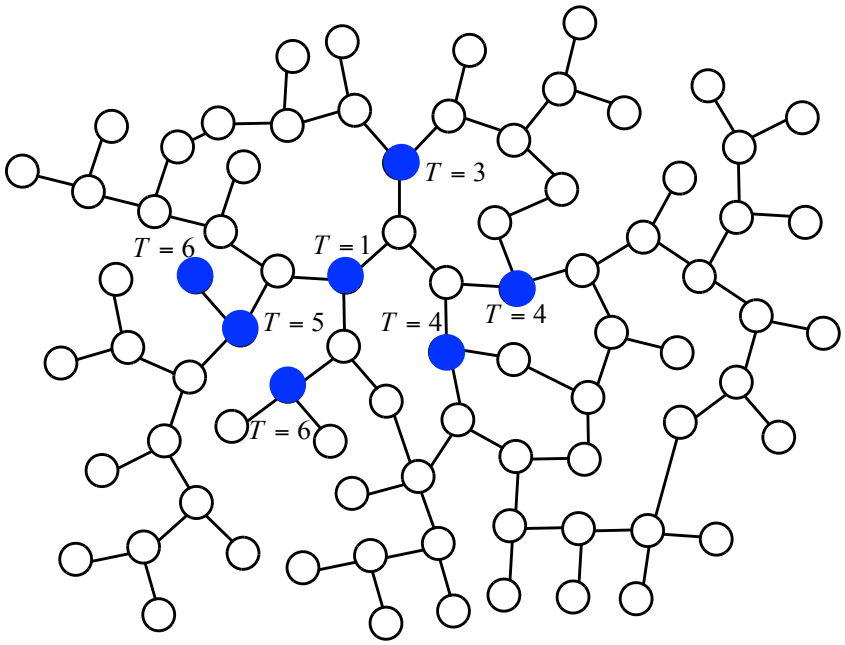


can we locate the message author?

Snapshot



Timing

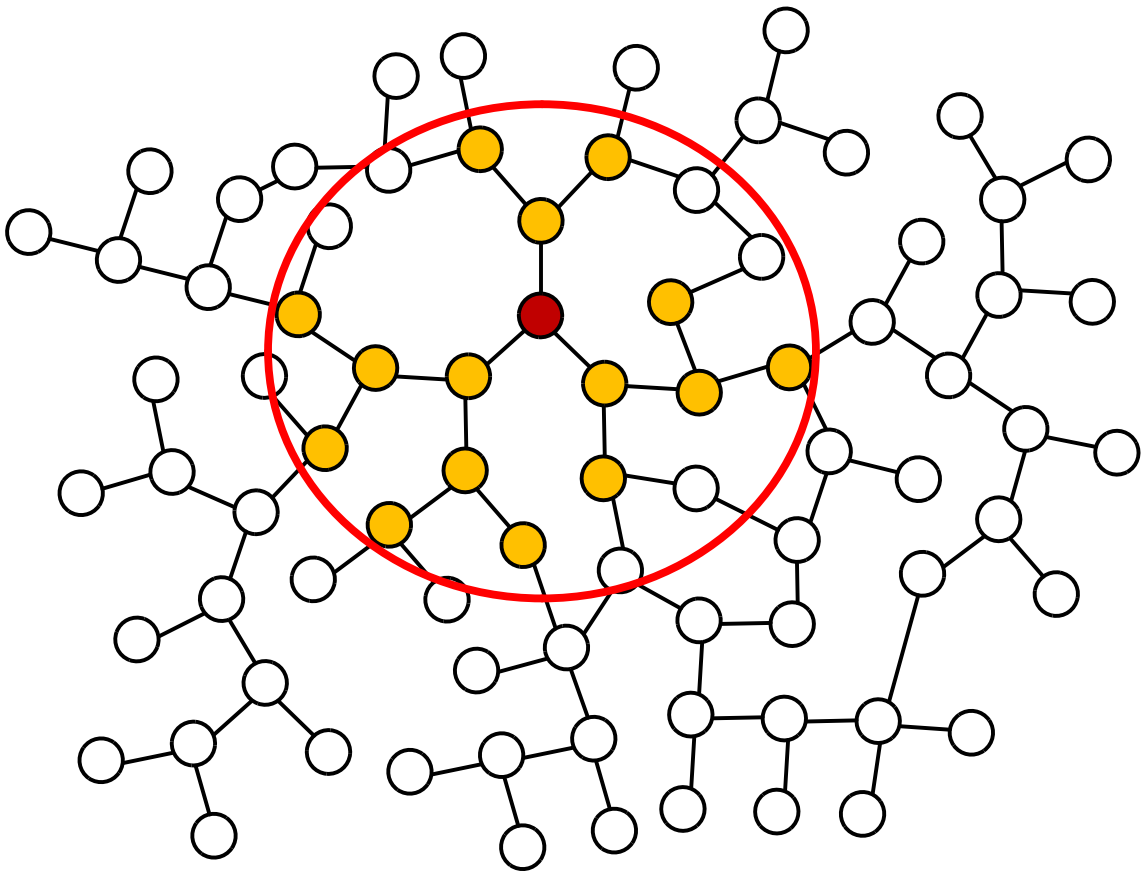


meta-data

time	*from*	*to*	*control*
10:12	***Alice***	***Spy1***	***101101***
10:25	***Bob***	***Spy2***	***101001***
11:01	***Mary***	***Spy3***	***100100***

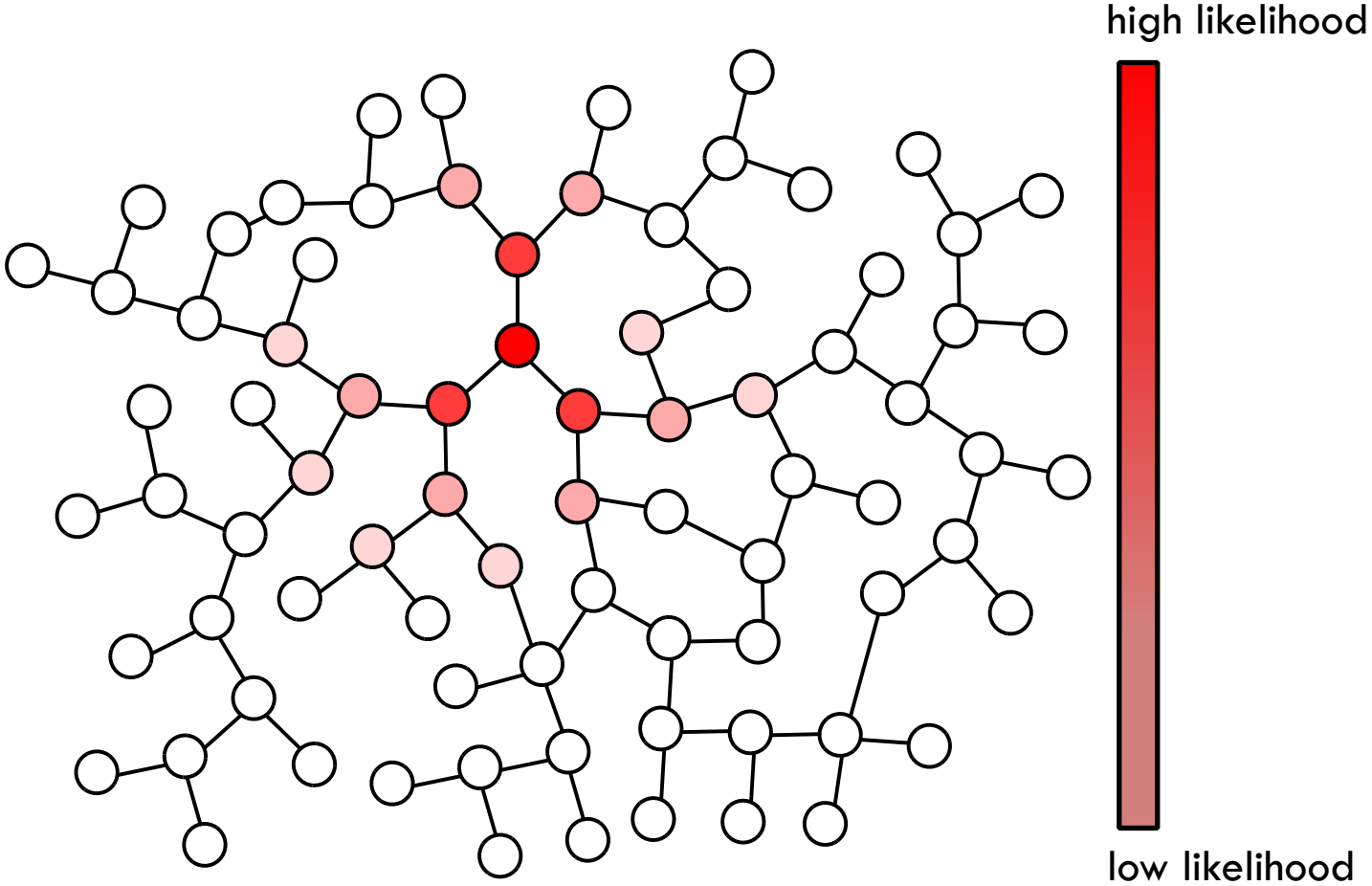
Snapshot-based adversary

Snapshot reveals the source



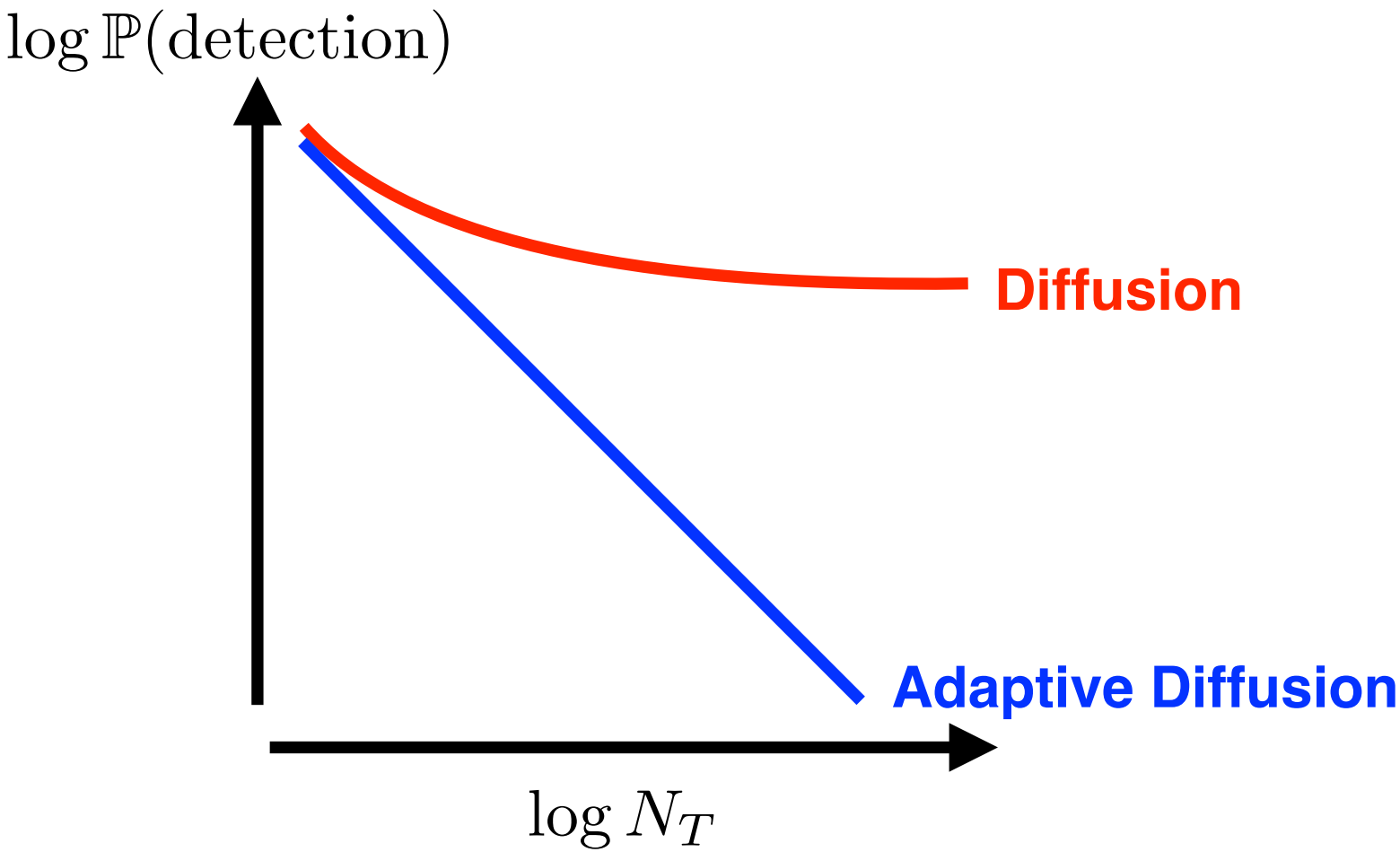
- message author is likely to be in the **“center”**

Rumor Centrality [Shah, Zaman ’11]



Similar performance using Jordan centrality [Zhu, Ying ’13]

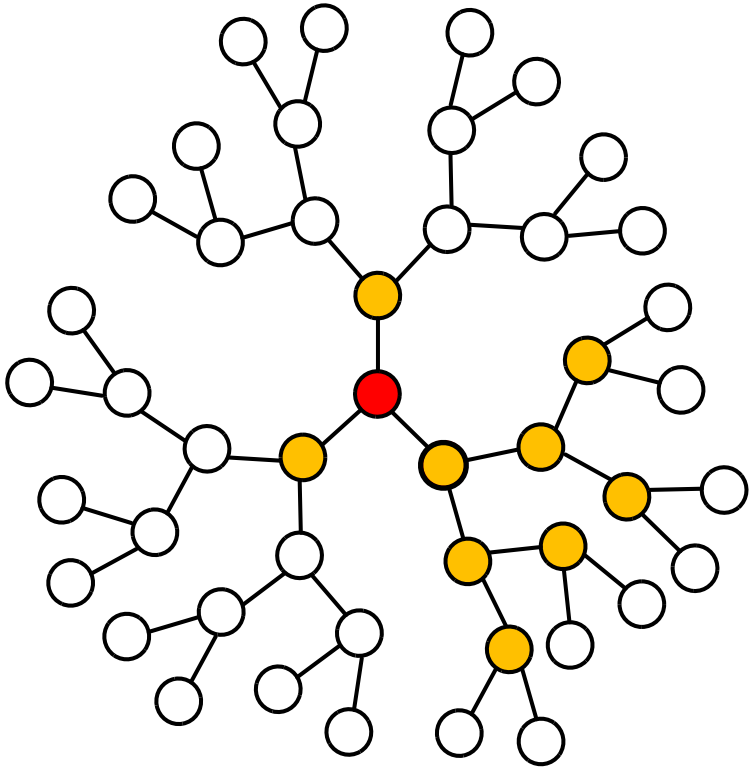
Our Goal (on infinite regular trees)



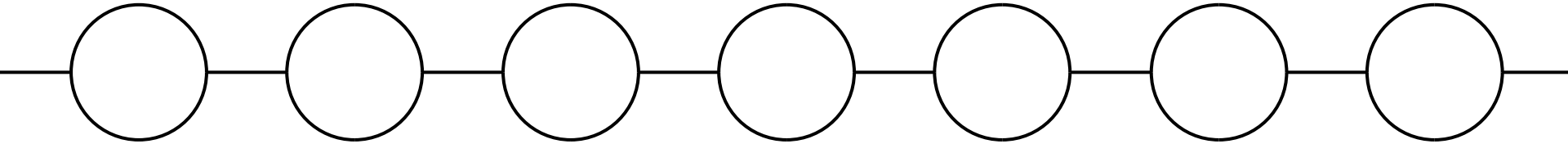
$$\mathbb{P}(\text{Detection}) = \frac{1}{N_T}$$

Setting

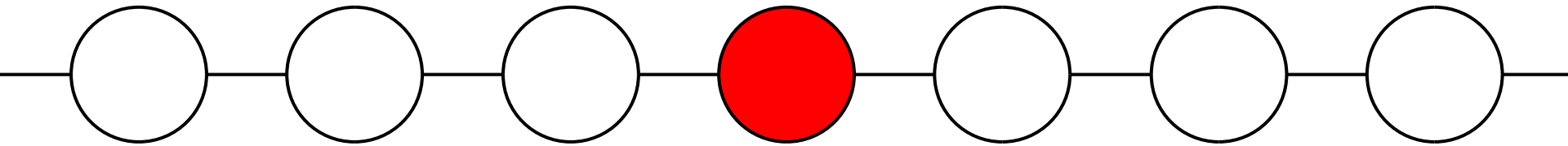
- Contact network is infinite, regular, tree
- Adversary has the contact network and the snapshot
- Protocol is allowed to infect any 1-hop neighbors at each time



Line graph

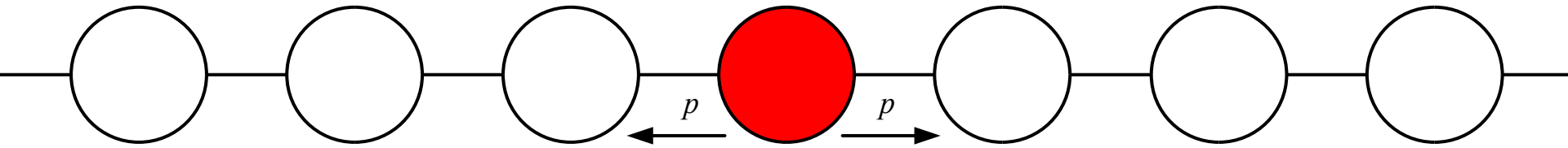


Line graph: diffusion



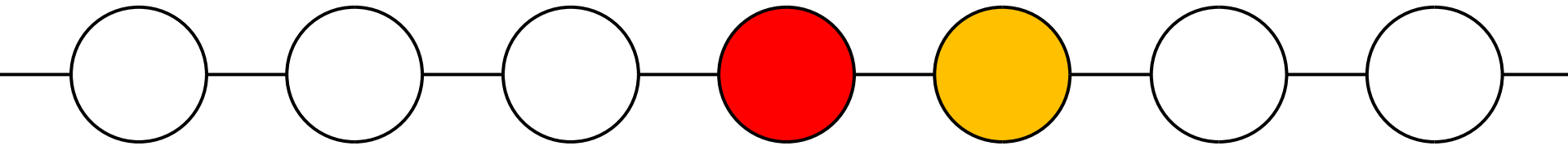
$T = 0$

Line graph: diffusion



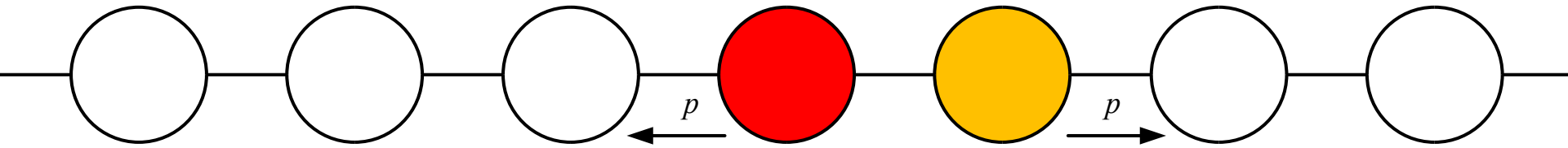
$T = 1$

Line graph: diffusion



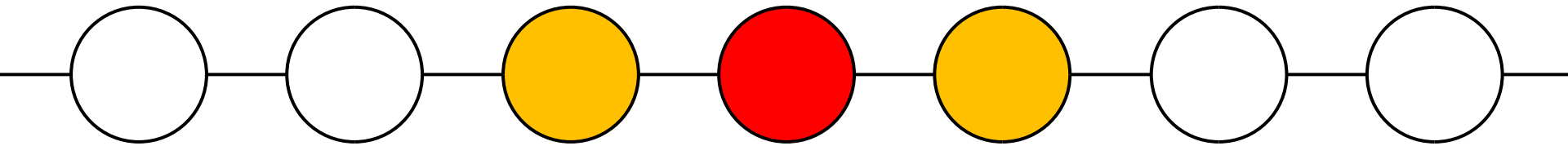
$T = 1$

Line graph: diffusion



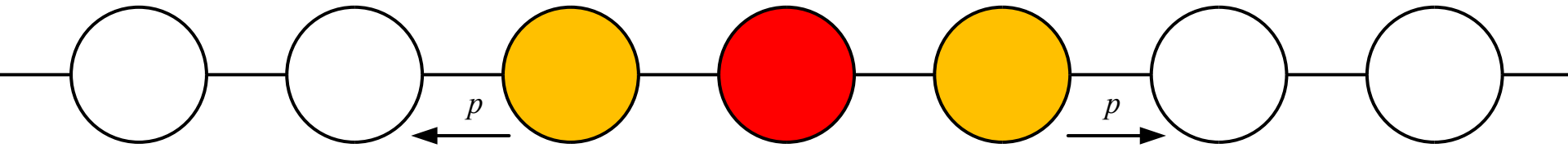
$T = 2$

Line graph: diffusion



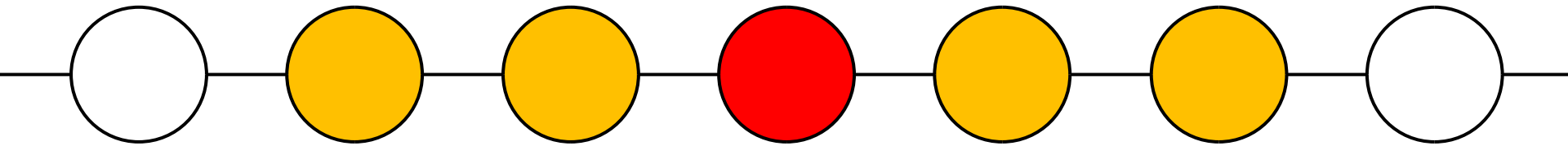
$T = 2$

Line graph: diffusion



$T = 3$

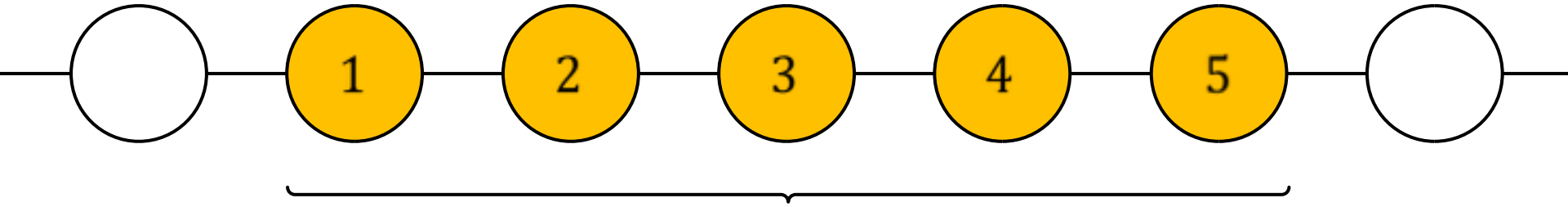
Line graph: diffusion



$T = 3$

- equivalent to **two independent random walks**

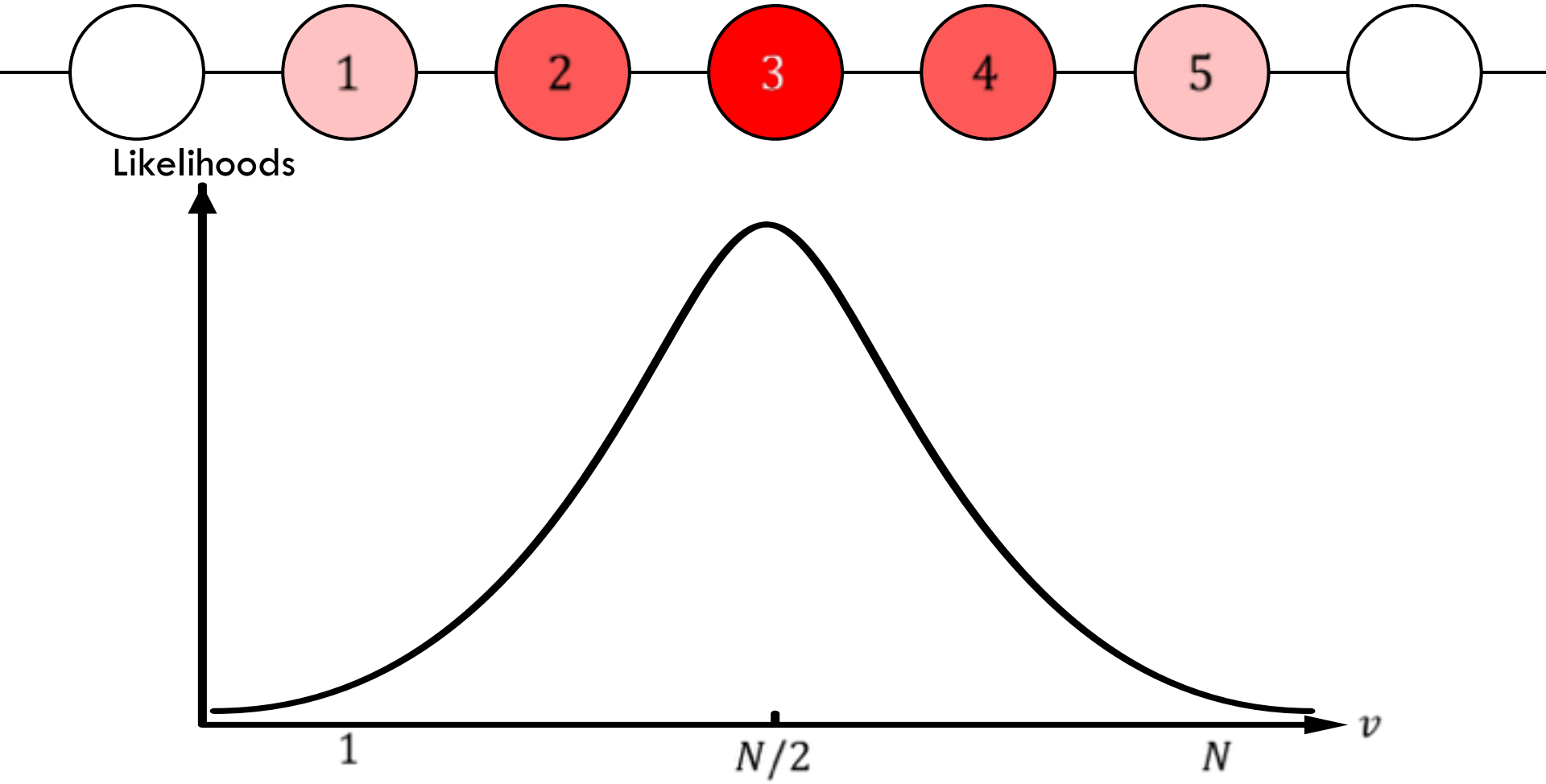
Adversary with snapshot



nodes with the message

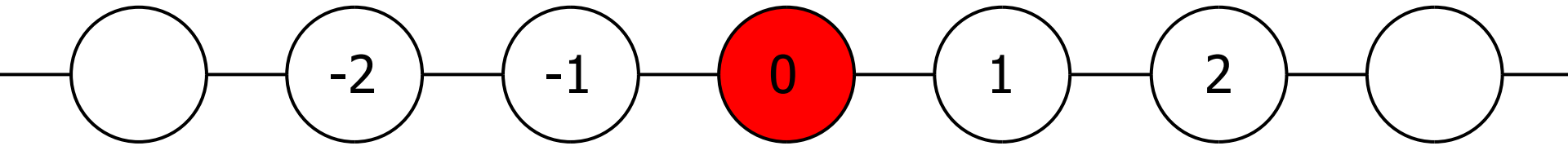
can we locate the message author?

Maximum likelihood detection



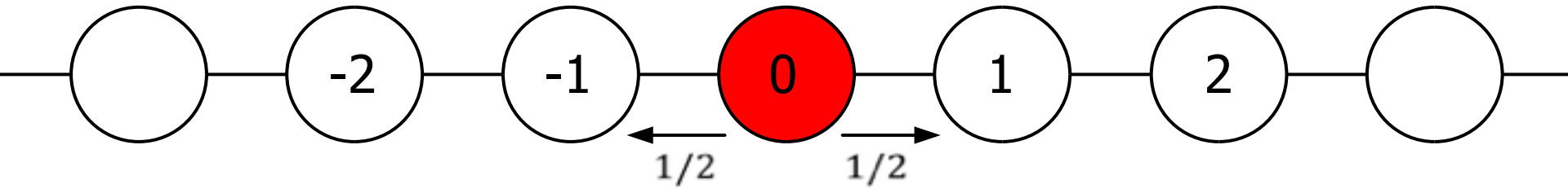
Probability of detection $\approx \frac{1}{\sqrt{N}}$

Line graph: adaptive diffusion



$T = 0$

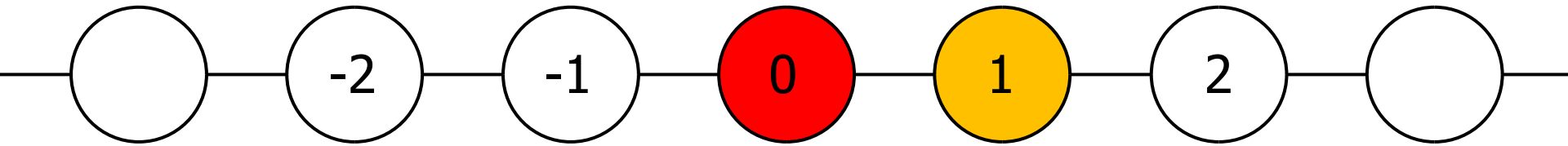
Line graph: adaptive diffusion



$T = 1$

each neighbor is infected w.p. 1/2

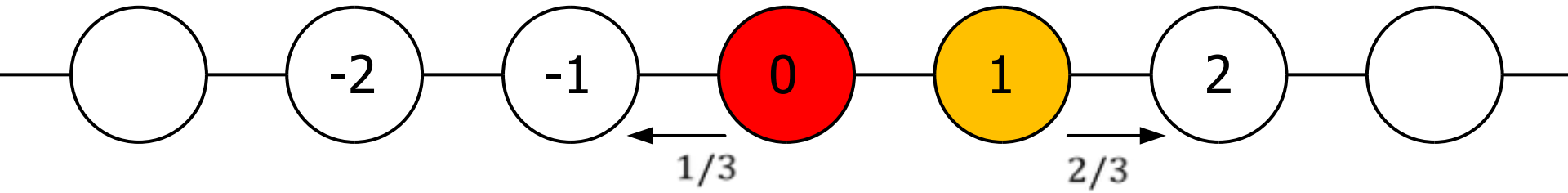
Line graph: adaptive diffusion



$T = 1$

Node 1 receives message at T = 1

Line graph: adaptive diffusion

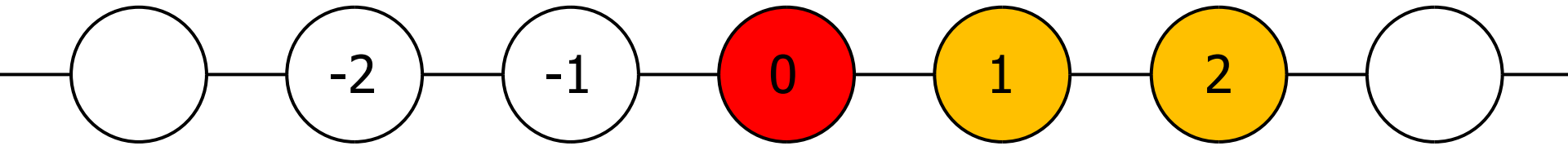


$T = 2$

- Adaptive diffusion prescribes to pass at adaptive rate

$$p_{\text{infect}} = \frac{h+1}{T+1}$$

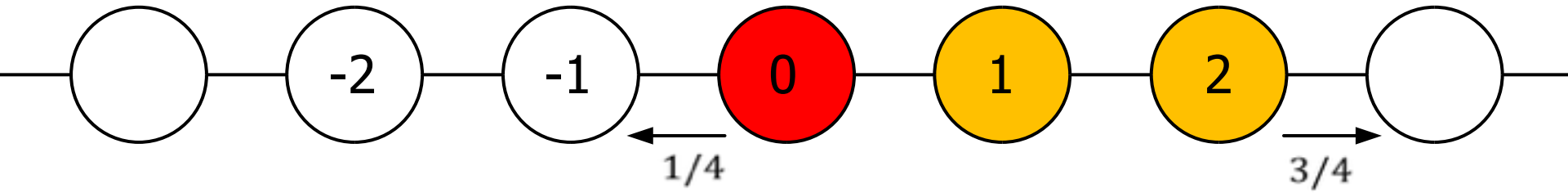
Line graph: adaptive diffusion



$T = 2$

- node 2 receives the message

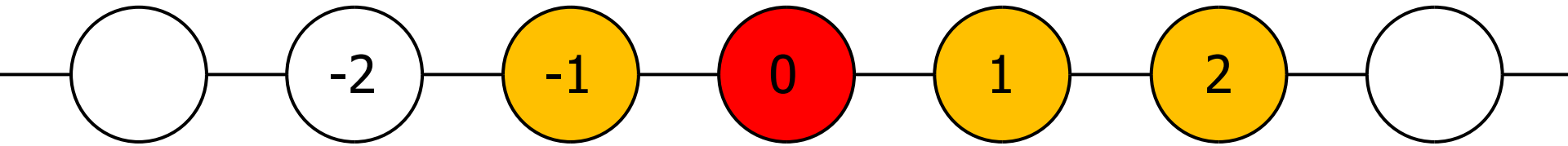
Line graph: adaptive diffusion



$T = 3$

$$p_{\text{infect}} = \frac{h+1}{T+1}$$

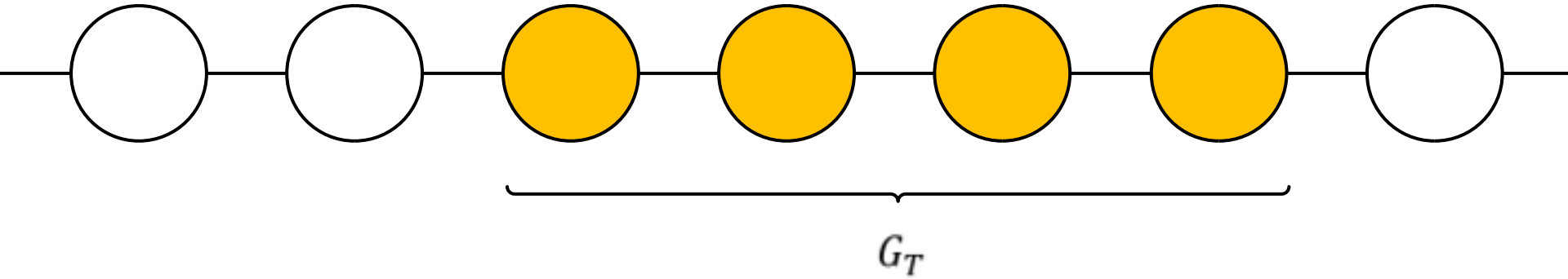
Line graph: adaptive diffusion



$T = 3$

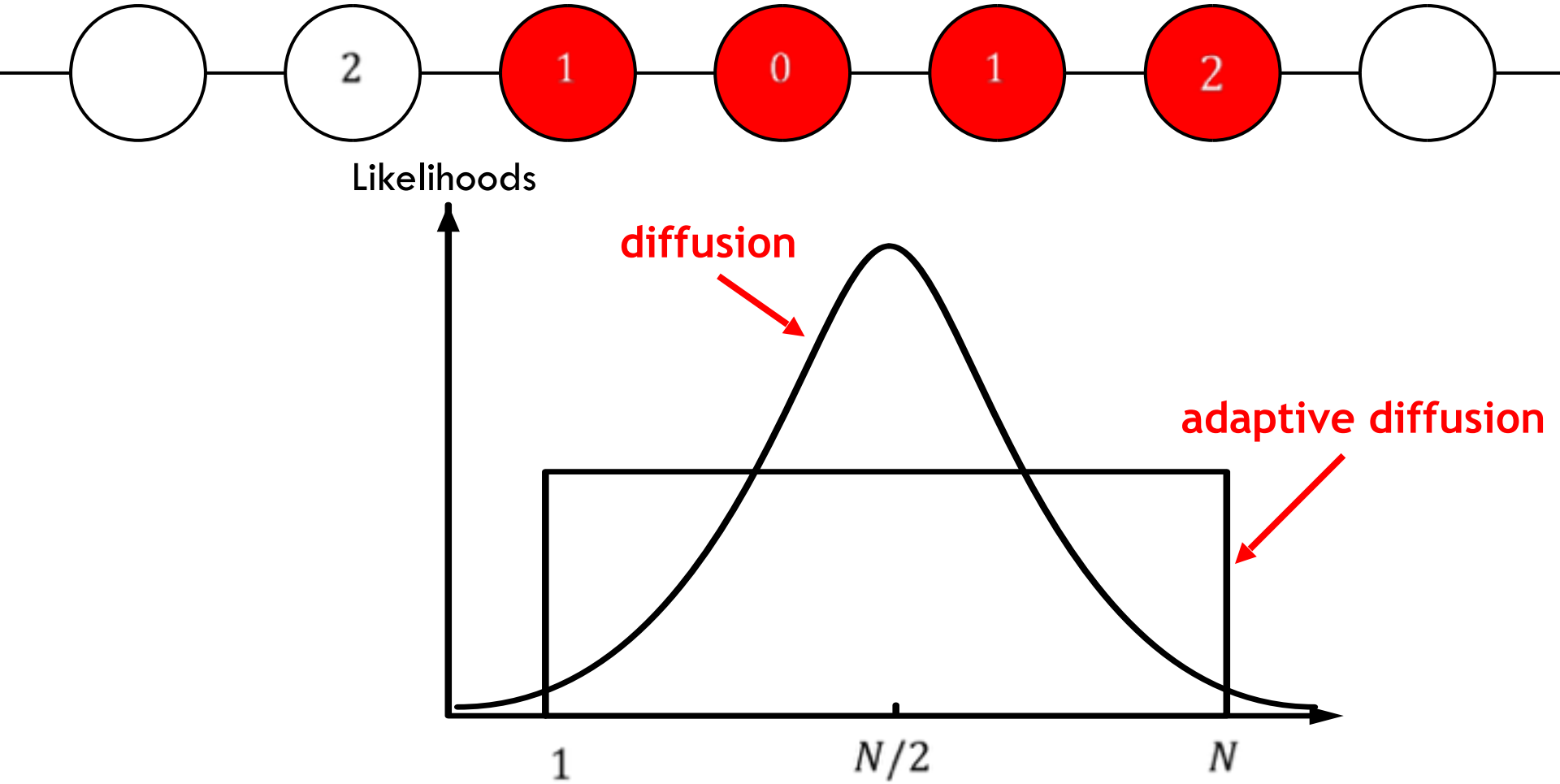
- equivalent to **Polya's urn process**
- **adaptive and asymmetric:** Nodes that are infected earlier, spread faster

Given snapshot



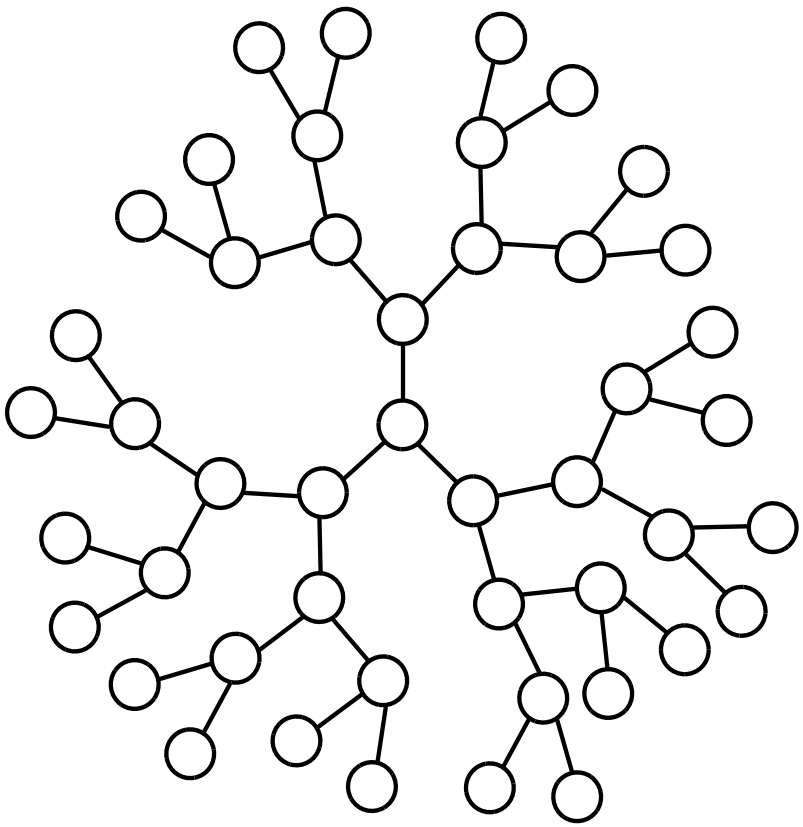
can we locate the **message author?**

Maximum likelihood detection

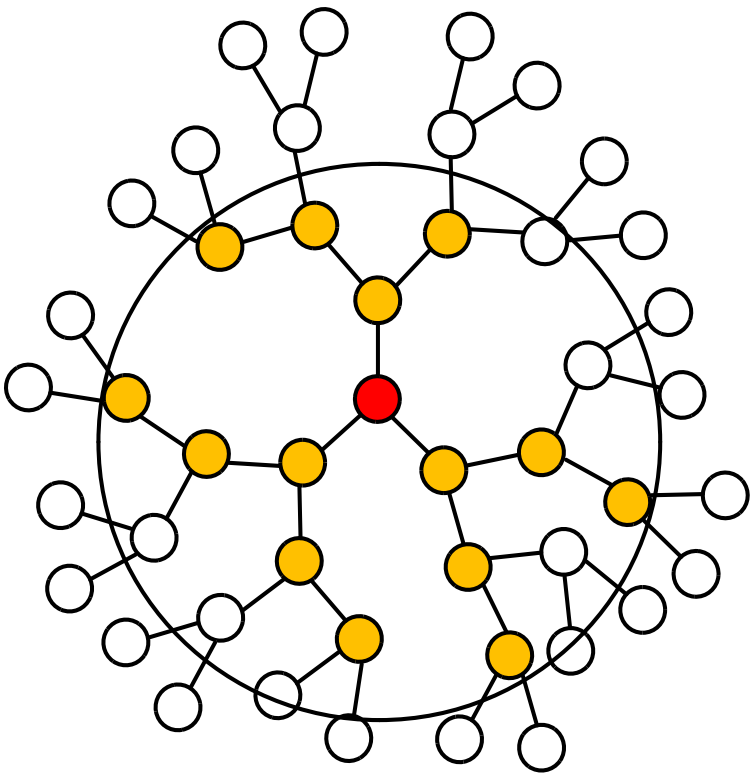


$$\text{probability of detection} \sim \frac{1}{N}$$

d-regular trees

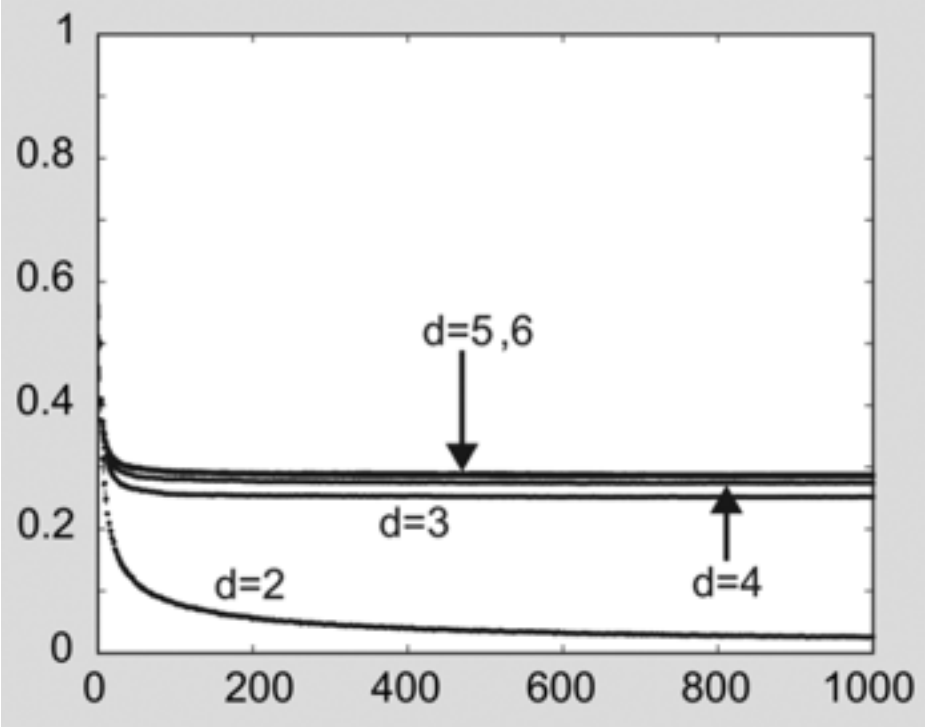


d-regular trees: diffusion



spread with fixed probability p

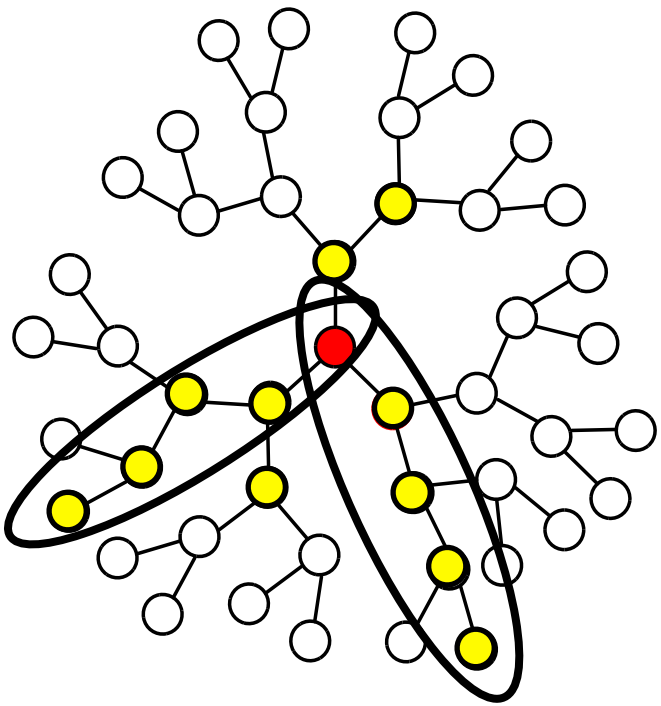
Probability of detection using Rumor Centrality



number of nodes with the message

- [Shah & Zaman '11]

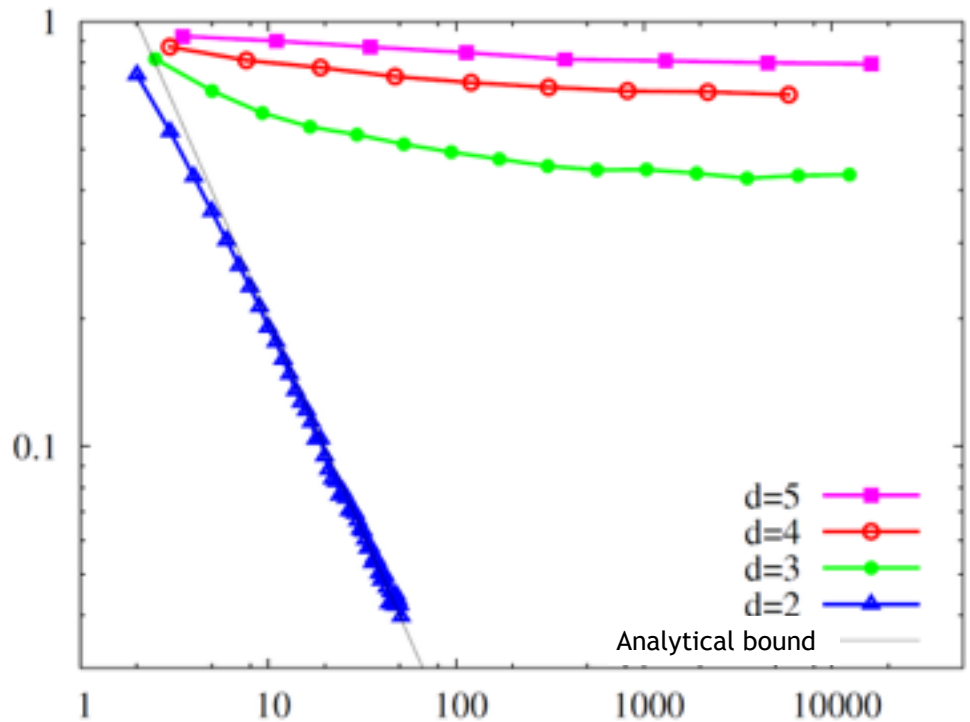
d-regular trees: Pólya's urn processes



spread with probability

$$p(h,t) = \frac{h+1}{t+1}$$

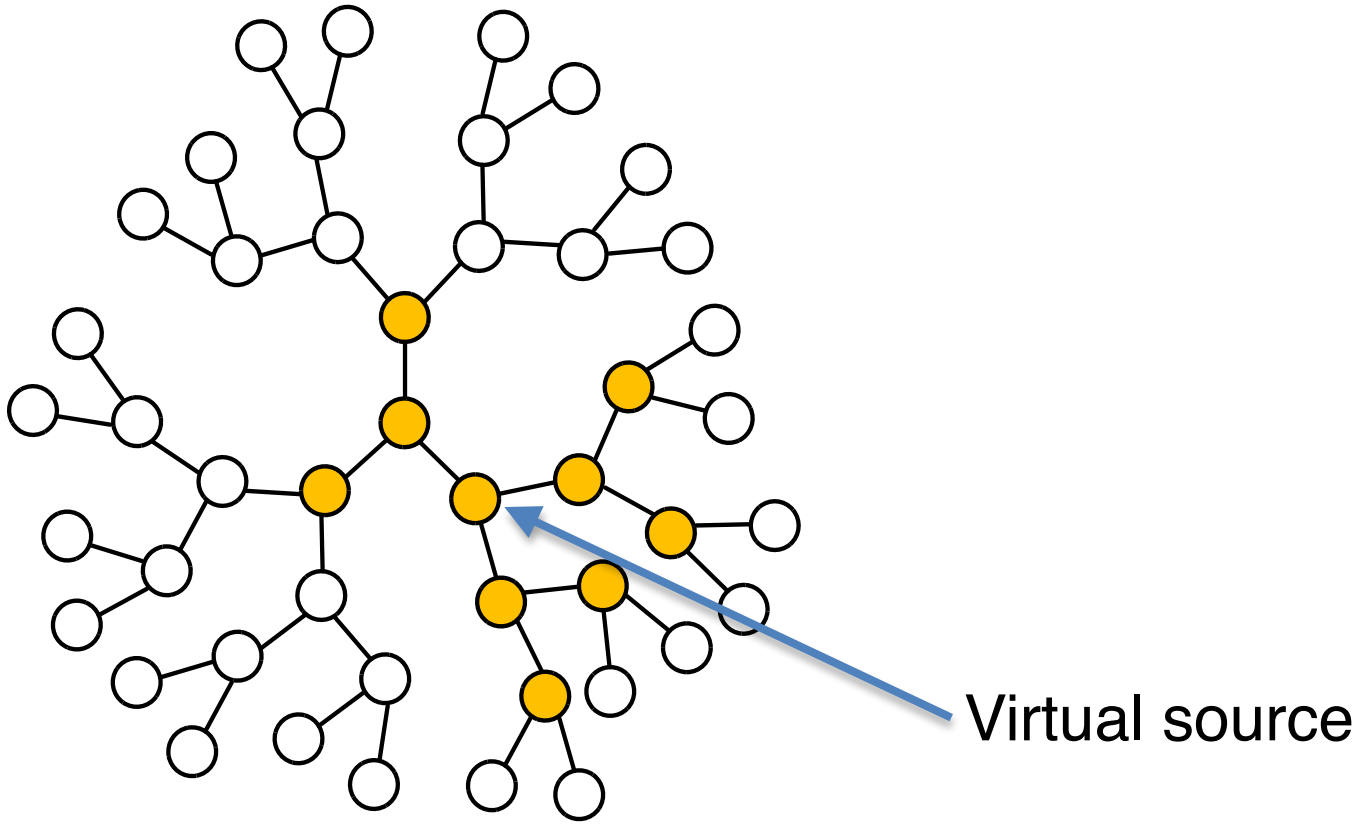
Probability of detection using Jordan centrality



number of nodes with the message

d-regular trees : adaptive diffusion

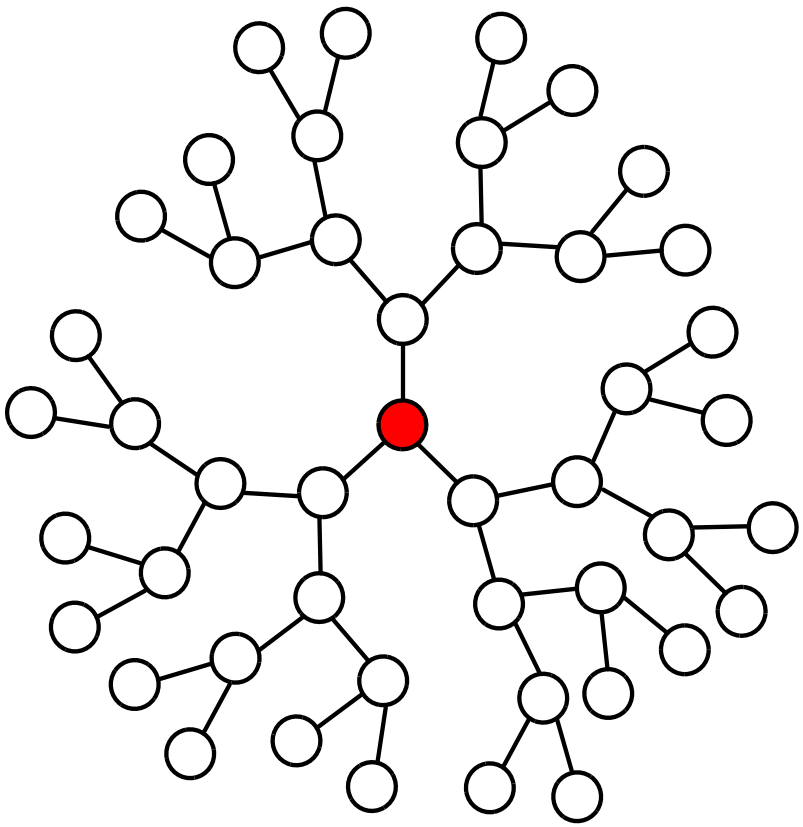
$T = 4$



- Strategy:
  - Design the infection to be a symmetric ball of depth T/2
  - Such that the source is equally likely to be any node at any given time T

d-regular trees : adaptive diffusion

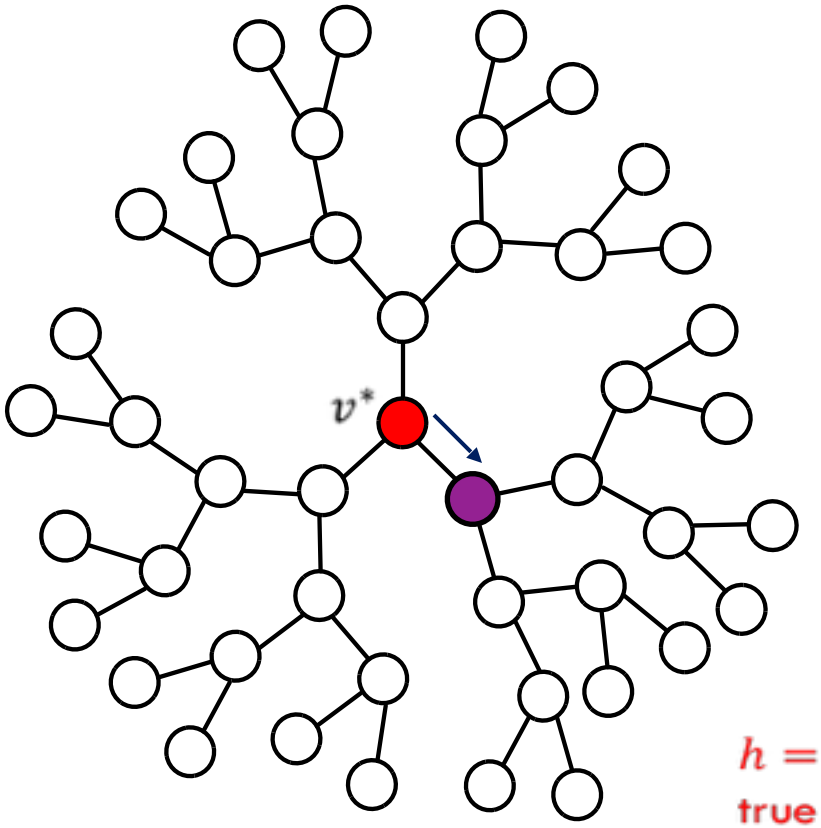
$T = 0$



- initially, the author is also the virtual source
- and randomly selects a neighbor to be next virtual source

d-regular trees : adaptive diffusion

$T = 1$

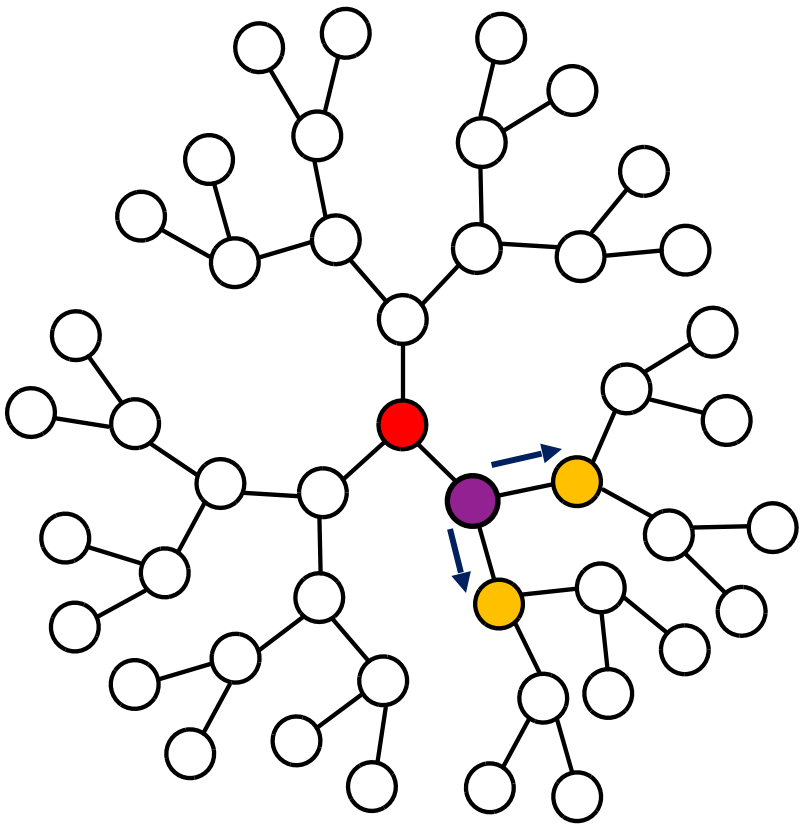


in addition to the message, v^* passes $h = 1$ and $T = 1$ to the chosen neighbor

h = # of hops away from true source

d-regular trees : adaptive diffusion

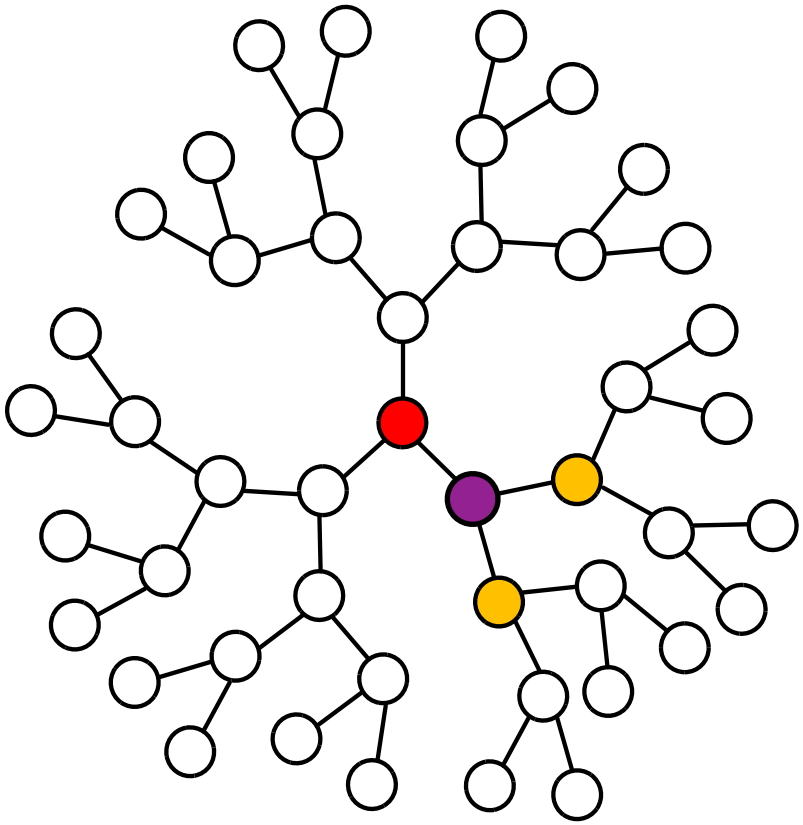
$T = 2$



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

d-regular trees : adaptive diffusion

$T = 2$

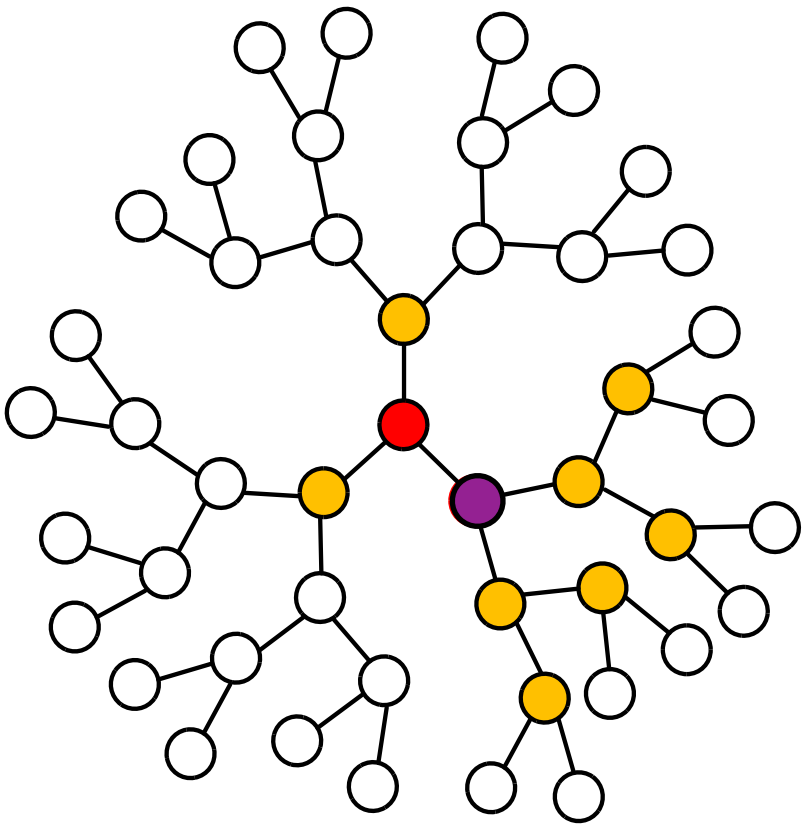


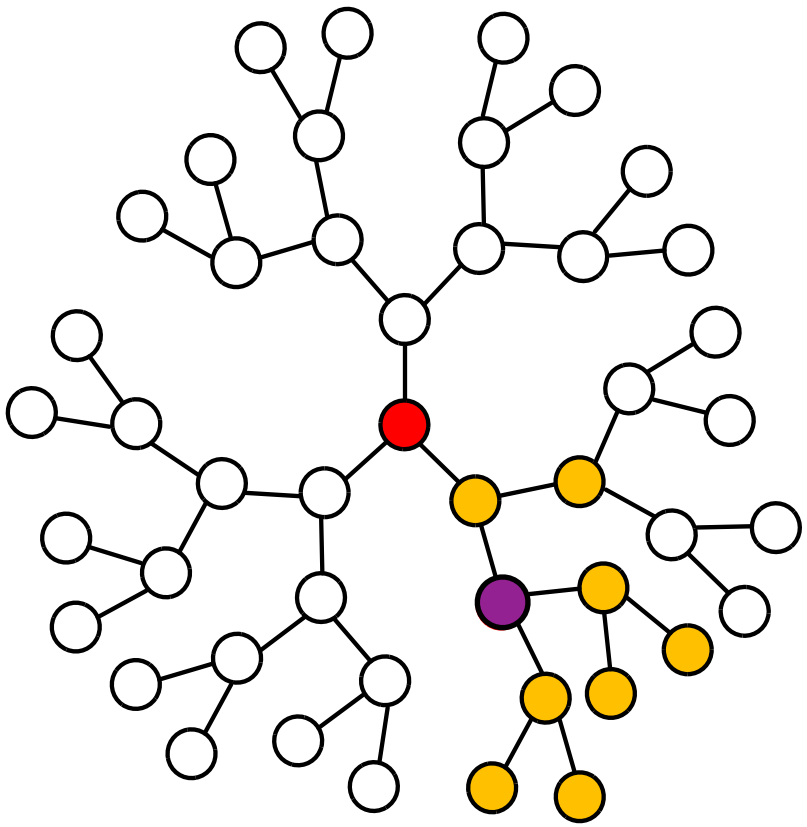
- at T=2, the protocol has two options

keeping the virtual source token

passing the virtual source token

d-regular trees : adaptive diffusion





keeping the virtual source token

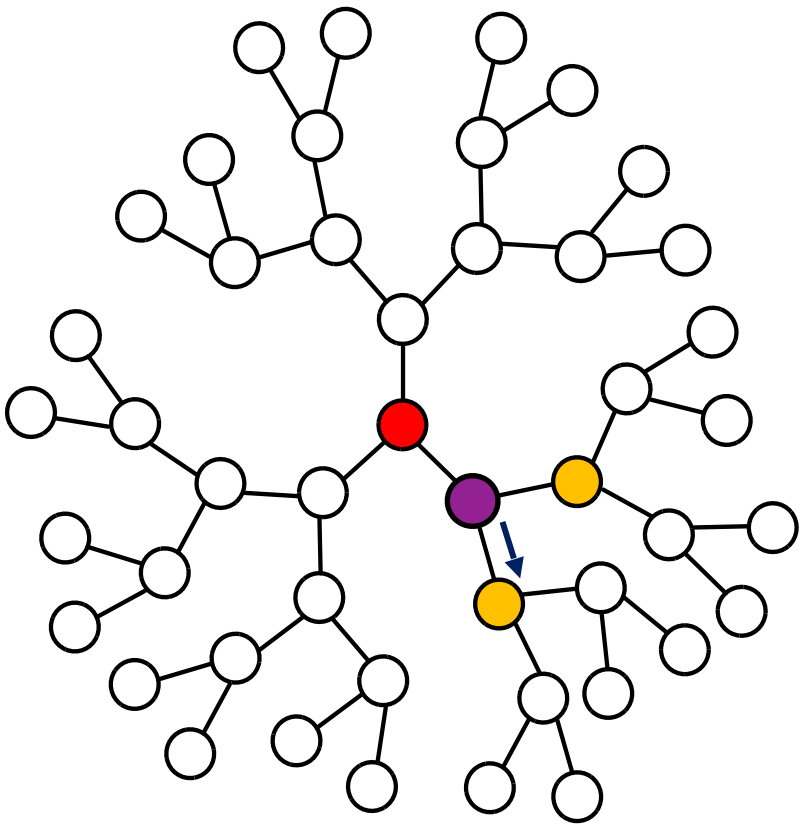
with probability $\alpha_{d,T,h}$

passing the virtual source token

with probability $1 - \alpha_{d,T,h}$

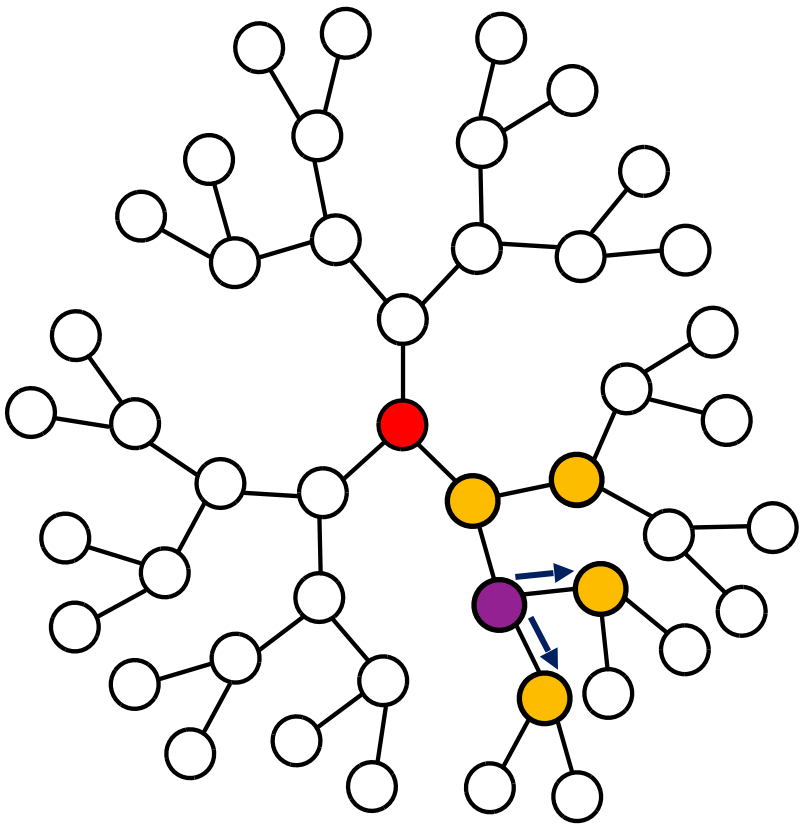
Passing the virtual source token

$T = 2$



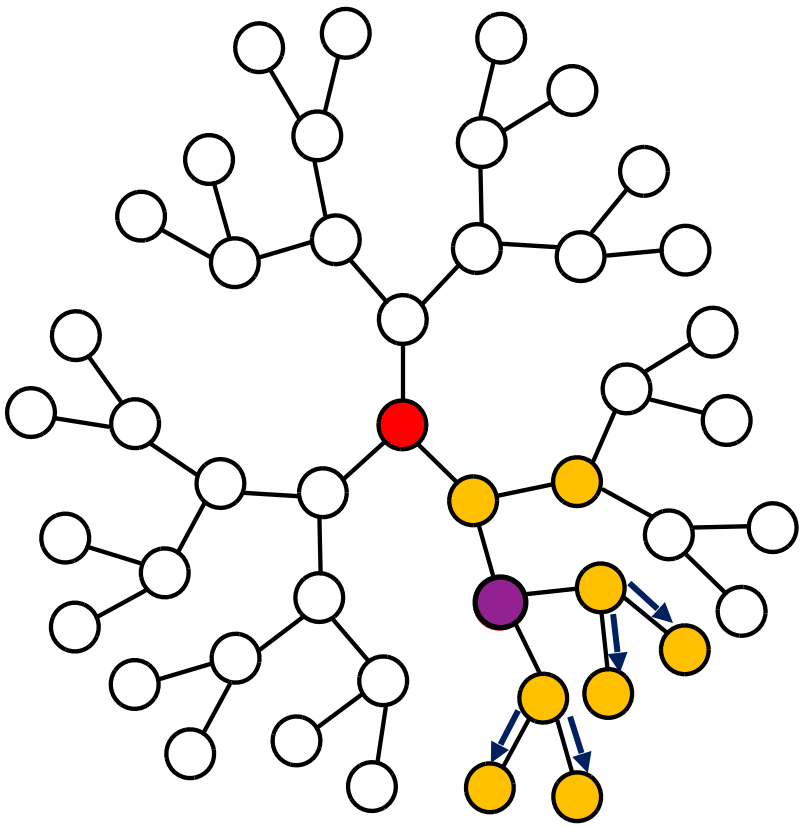
Passing the virtual source token

$T = 3$



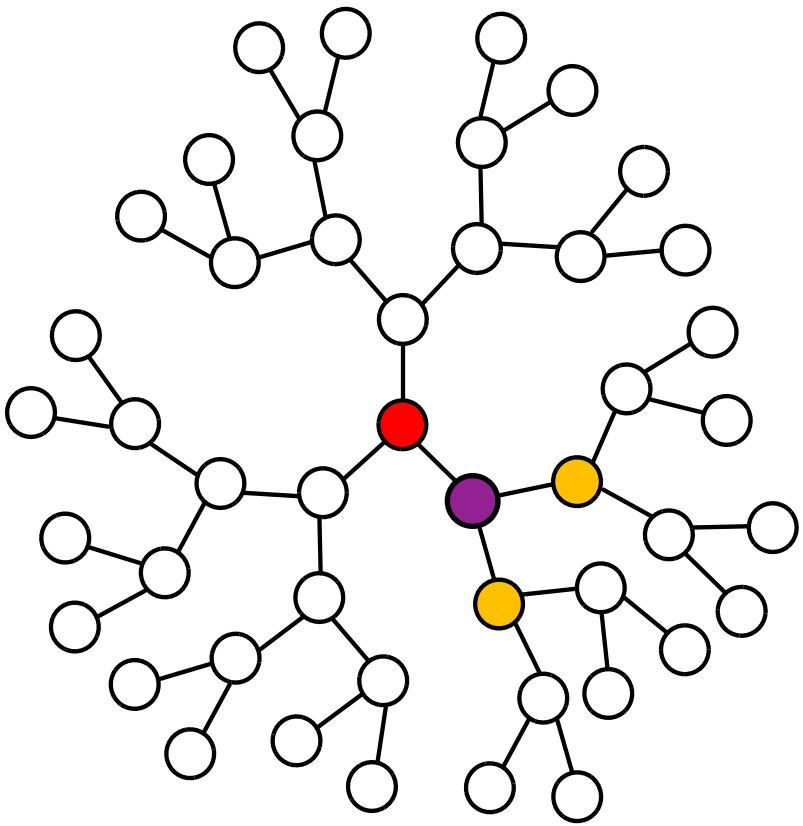
Passing the virtual source token

$$h = 2$$
$$T = 4$$



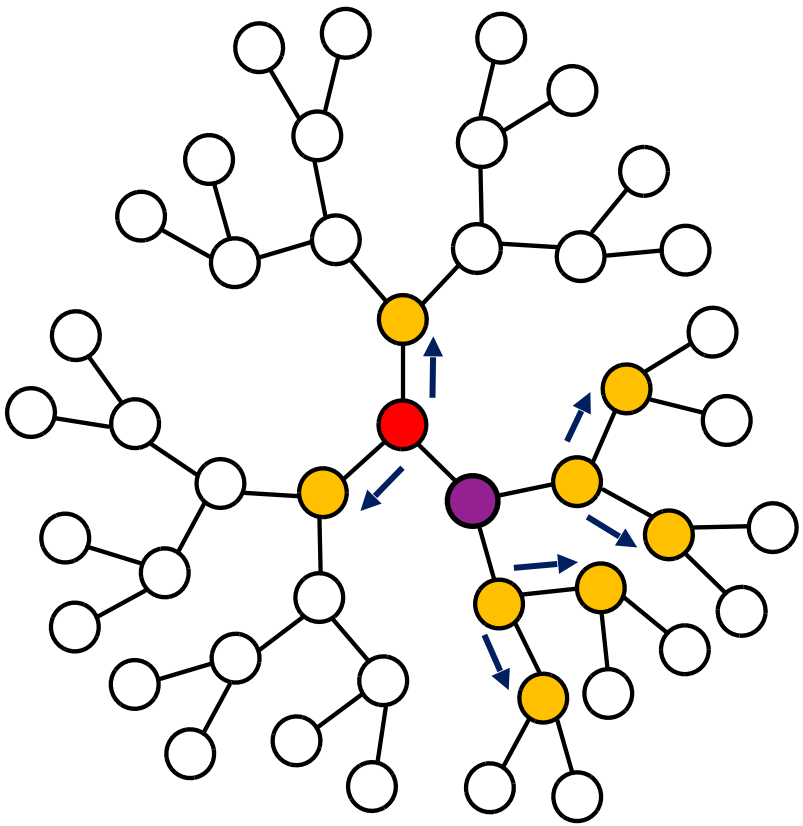
Keeping the virtual source token

$T = 2$



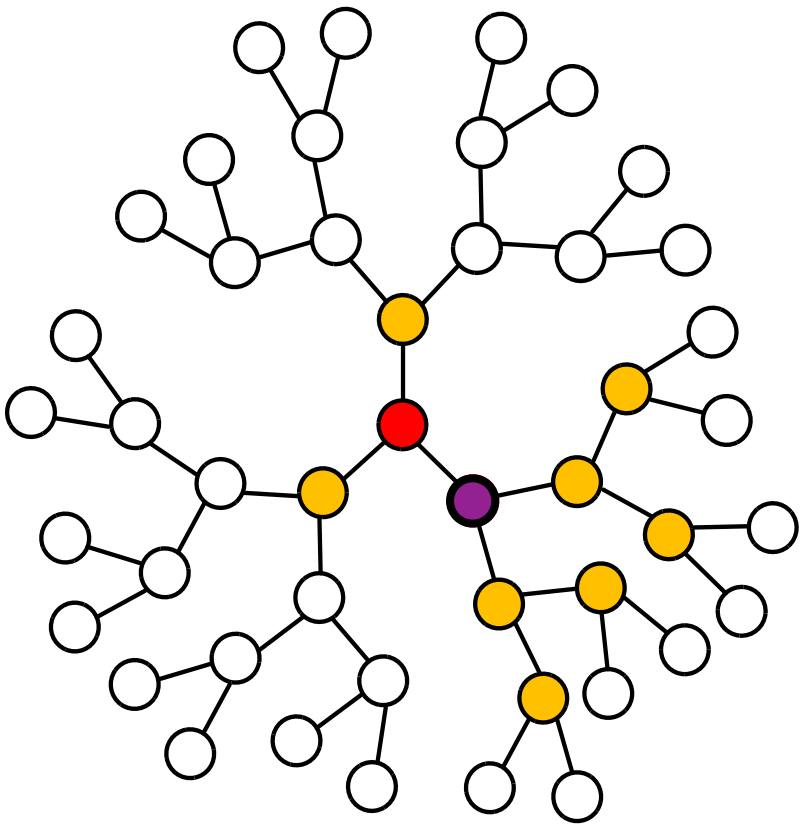
Keeping the virtual source token

$T = 3$

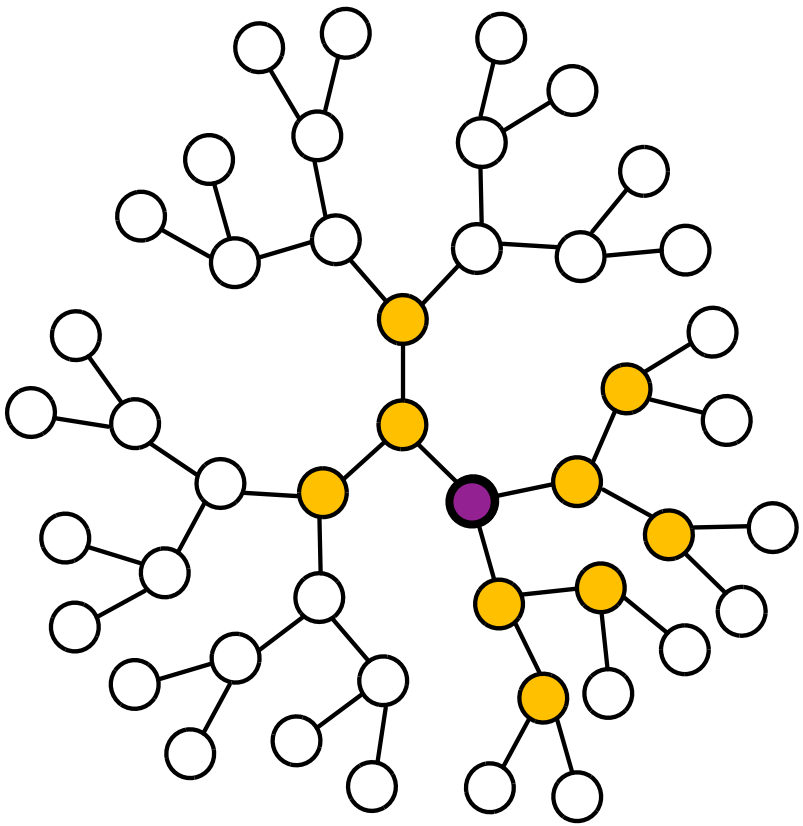


Keeping the virtual source token

$$h = 1$$
$$T = 4$$

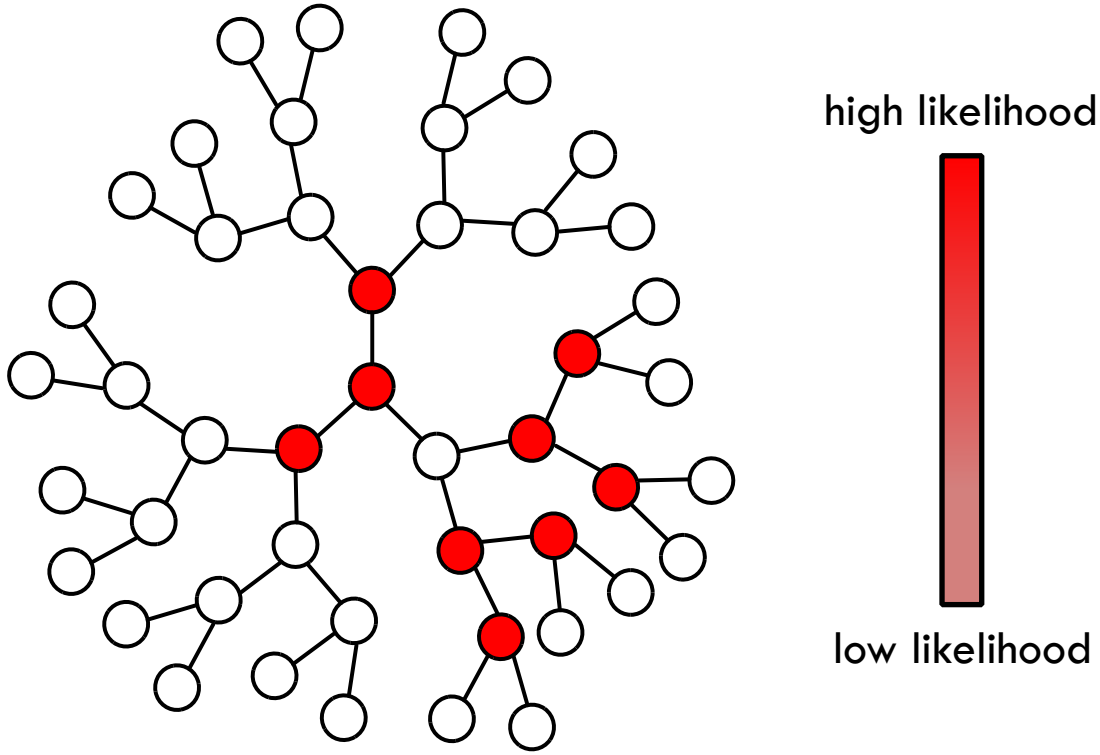


Adversary with snapshot



can the adversary locate the **message author?**

Maximum likelihood estimation



- If virtual source is kept with prob. $\alpha_{d,T,h} = \dfrac{(d-1)^{\frac{T}{2}+1-h} - 1}{(d-1)^{\frac{T}{2}+1} - 1}$
- nodes that receive message faster, spread faster

Maximum likelihood estimation

Theorem. [*Fanti, Kairouz, Oh, Viswanath* 2015]

On an infinite *d*-regular tree,

1. adaptive diffusion spreads fast,

$$N_T \simeq (d-1)^{T/2}$$

2. achieves almost perfect obfuscation

$$\mathbb{P}(\text{Detection}) = \frac{1}{N_T - 1}$$

3. The expected distance between the estimated and the true source is T/2

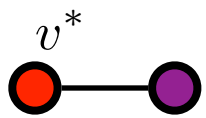
What if the tree is irregular?

$T = 0$

$$d_v = \begin{cases} 3 & \text{w.p. } 0.5 \\ 5 & \text{w.p. } 0.5 \end{cases}$$

G_T

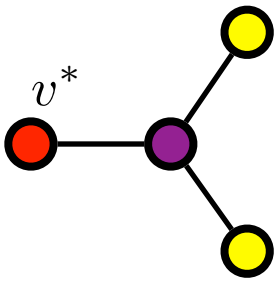
$T = 1$



$$d_v = \begin{cases} 3 & \text{w.p. } 0.5 \\ 5 & \text{w.p. } 0.5 \end{cases}$$

G_T

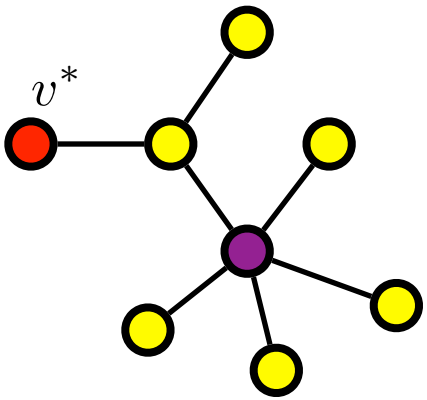
$T = 2$



$$d_v = \begin{cases} 3 & \text{w.p. } 0.5 \\ 5 & \text{w.p. } 0.5 \end{cases}$$

G_T

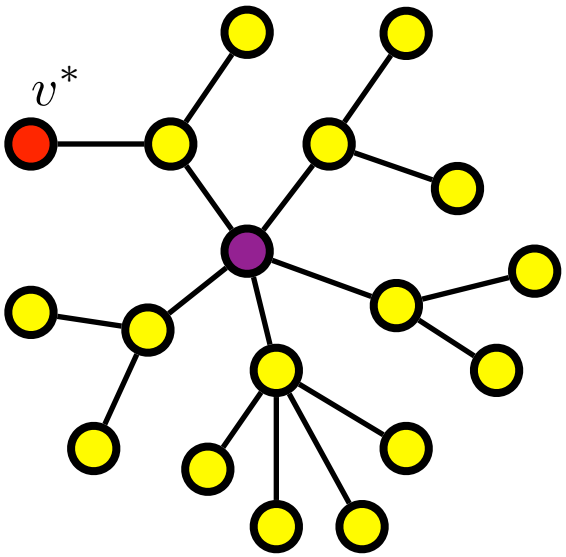
$T = 3$



$$d_v = \begin{cases} 3 & \text{w.p. } 0.5 \\ 5 & \text{w.p. } 0.5 \end{cases}$$

G_T

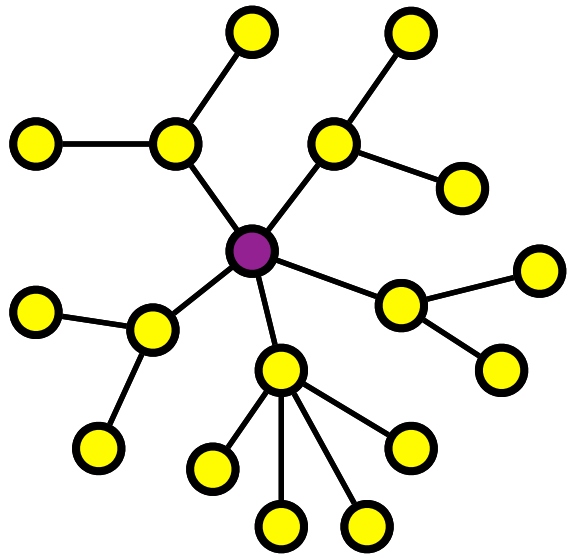
$T = 4$



G_T

$$d_v = \begin{cases} 3 & \text{w.p. } 0.5 \\ 5 & \text{w.p. } 0.5 \end{cases}$$

$T = 4$

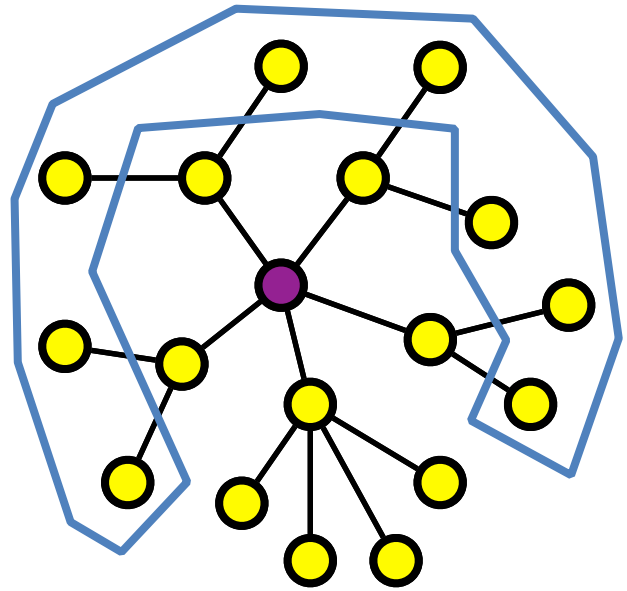


G_T

$$d_v = \begin{cases} 3 & \text{w.p. } 0.5 \\ 5 & \text{w.p. } 0.5 \end{cases}$$

$$\hat{v}_{\mathrm{ML}} = \arg\max_{v \in \partial G_T} \underbrace{\frac{1}{d_{vs} \prod_{w \in \phi(vs,v) \setminus \{vs,v\}} (d_w - 1)}}_{F(G_T)}$$

$T = 4$



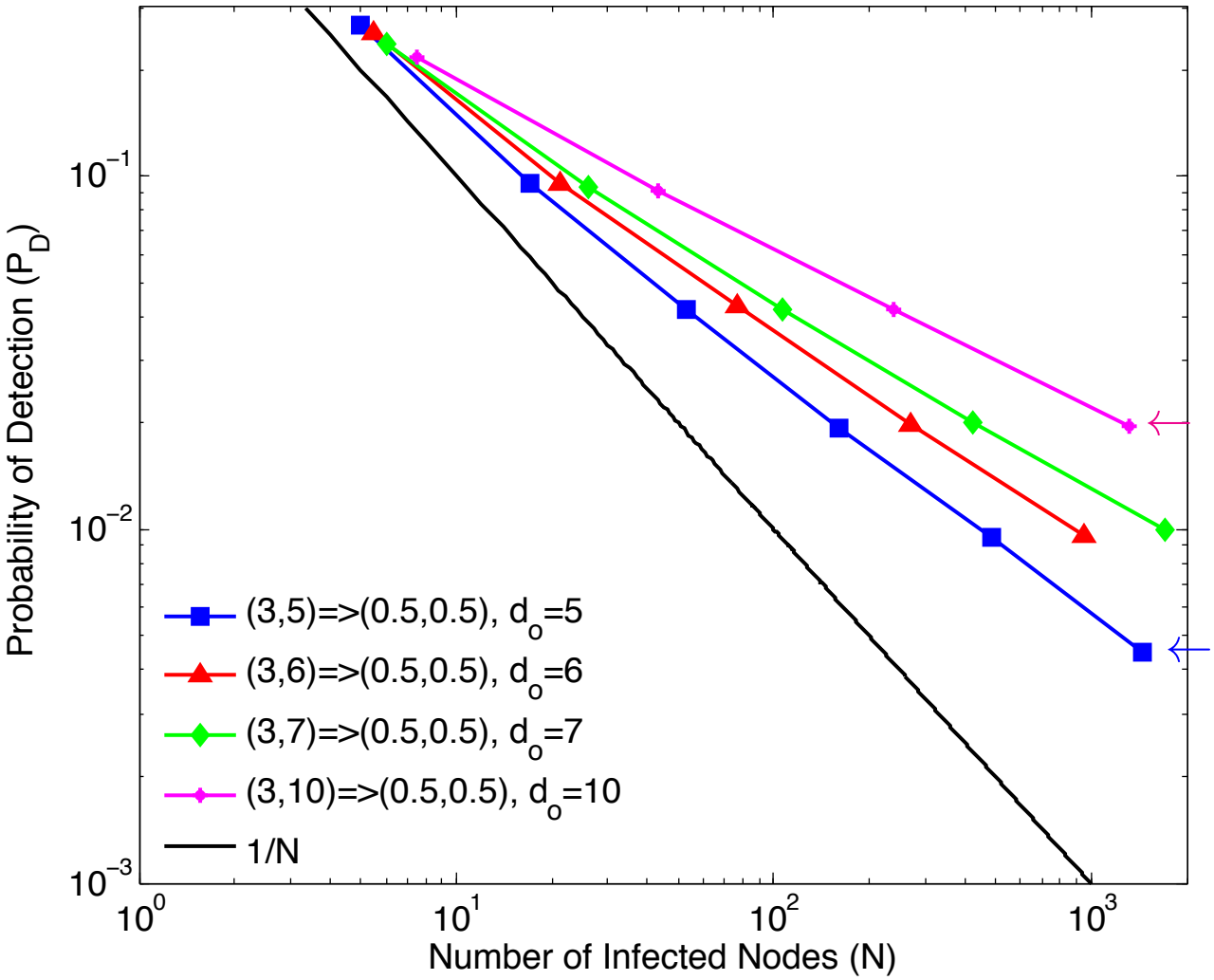
G_T

$$d_v = \begin{cases} 3 & \text{w.p. } 0.5 \\ 5 & \text{w.p. } 0.5 \end{cases}$$

$$\hat{v}_{\mathrm{ML}} = \arg\max_{v \in \partial G_T} \underbrace{\frac{1}{d_{vs} \prod_{w \in \phi(vs,v) \setminus \{vs,v\}} (d_w - 1)}}_{F(G_T)}$$

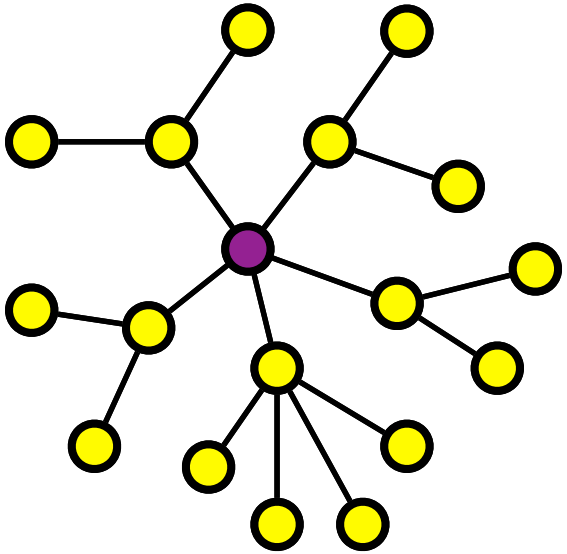
$$F(G_T) = \frac{1}{5 \times 2}$$

Does adaptive diffusion still achieve perfect obfuscation?



$$d_v = \begin{cases} 3 & \text{w.p. } 0.5 \\ 10 & \text{w.p. } 0.5 \end{cases}$$

$$d_v = \begin{cases} 3 & \text{w.p. } 0.5 \\ 5 & \text{w.p. } 0.5 \end{cases}$$



Galton-Watson tree G_T

$$F(G_T) = \max_{v \in \partial G_T} \frac{1}{d_{vs} \prod_{w \in \phi(vs,v) \setminus \{vs,v\}} (d_w - 1)}$$

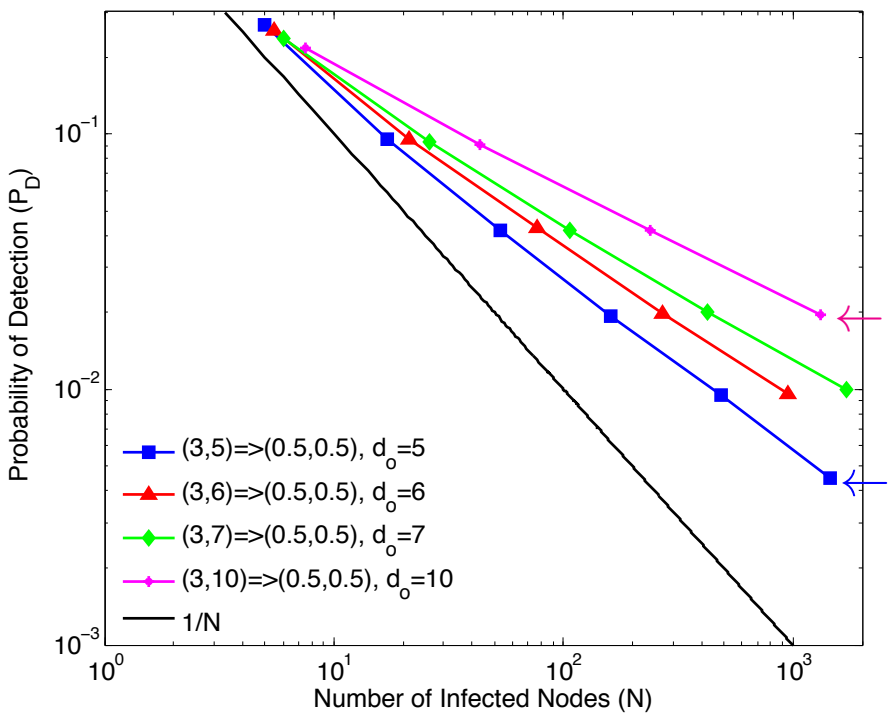
- Probability of detection:

$$\begin{aligned} \mathbb{P}(\hat{v}_{\mathrm{ML}} = v^*) &= \sum_{G_T} \mathbb{P}(G_T) \underbrace{\mathbb{P}(\hat{v}_{\mathrm{ML}} = v^* | G_T)}_{F(G_T)} \\ &= \mathbb{E}(F(G_T)) \end{aligned}$$

Corollary

$$\text{If } d_v = \begin{cases} d_{\min} & \text{w.p. } p_{\min} \\ \vdots & \vdots \end{cases}, \text{ and } (d_{\min}-1)p_{\min} \geq 1, \text{ then}$$

$$\begin{aligned} \mathbb{P}(\hat{v}_{\mathrm{ML}} = v^*) &= \mathbb{E}_{G_T}\left[\max_{v \in \partial G_T} \frac{1}{d_{vs} \prod_{w \in \phi(vs,v)\setminus\{vs,v\}} (d_w - 1)}\right] \\ &= (d_{\min}-1)^{-T+o(T)} \end{aligned}$$



$$\mathrm{d_v} = \begin{cases} 3 & \text{w.p. } 0.5 \\ 10 & \text{w.p. } 0.5 \end{cases}$$

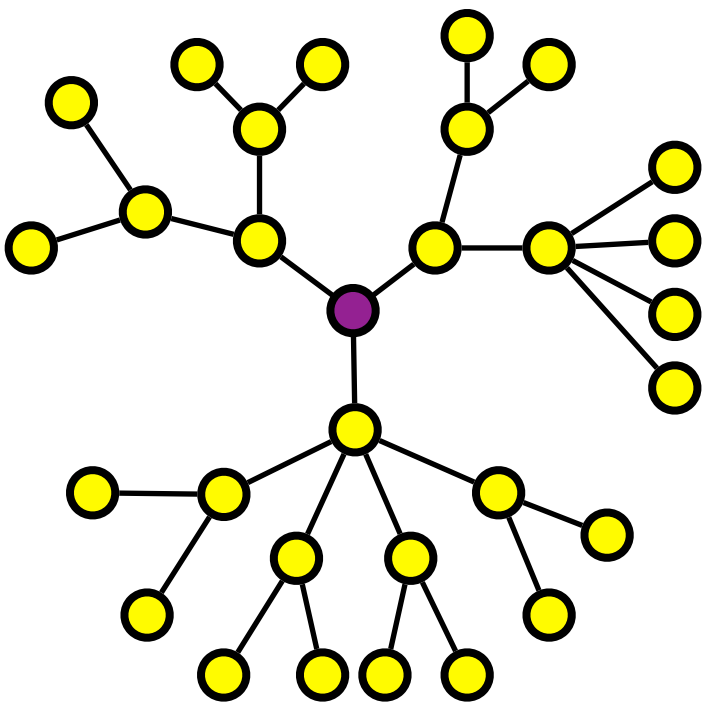
$$\mathrm{d_v} = \begin{cases} 3 & \text{w.p. } 0.5 \\ 5 & \text{w.p. } 0.5 \end{cases}$$

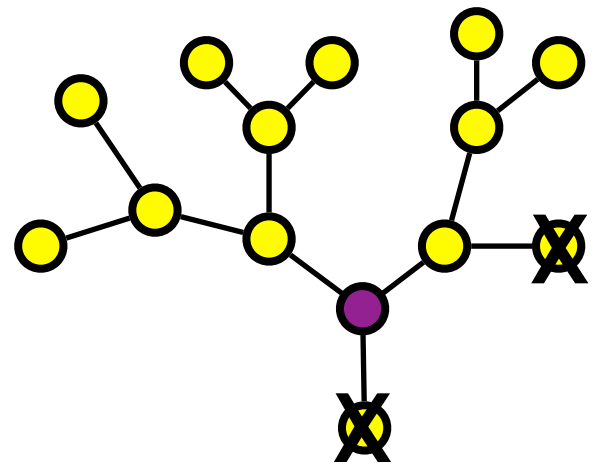
Proof idea for

$$\min_{v \in \partial G_T} \prod_{w \in \phi(vs,v) \setminus \{vs,v\}} (d_w - 1) \quad = \quad (d_{\min} - 1)^{T + o(T)}$$

$$d_v = \begin{cases} 3 & \text{w.p. } 0.5 \\ 5 & \text{w.p. } 0.5 \end{cases}$$

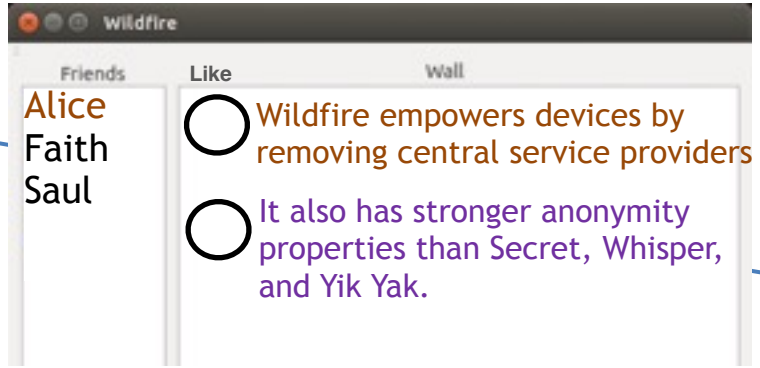
$$d_v = \begin{cases} 3 & \text{w.p. } 0.5 \\ 1 & \text{w.p. } 0.5 \end{cases}$$



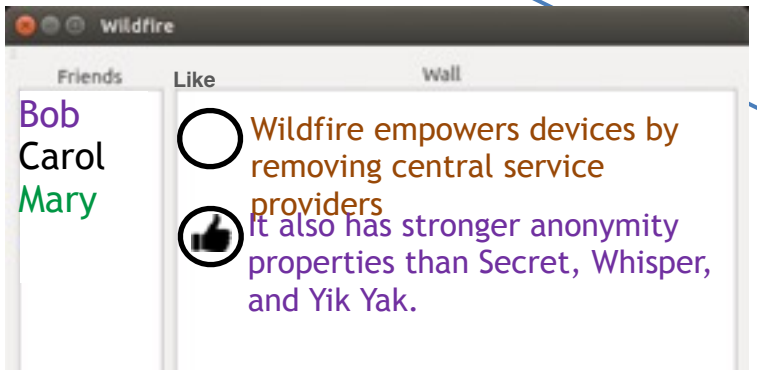


Messaging App: Wildfire

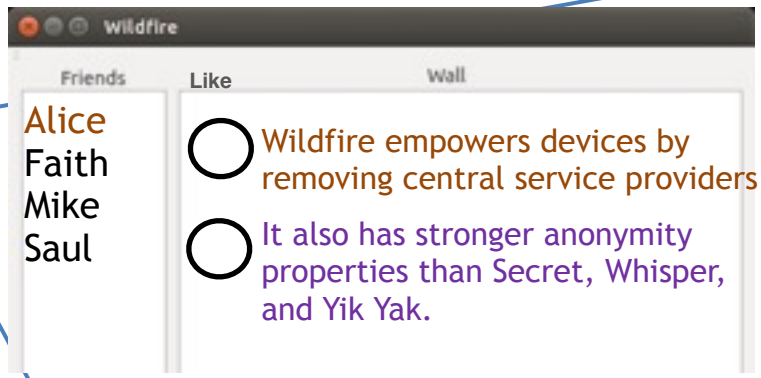
Bob



Alice



Mary



anonymous, distributed, secure implementation