Rumor Source Obfuscation

Peter Kairouz

University of Illinois at Urbana-Champaign

Joint work with

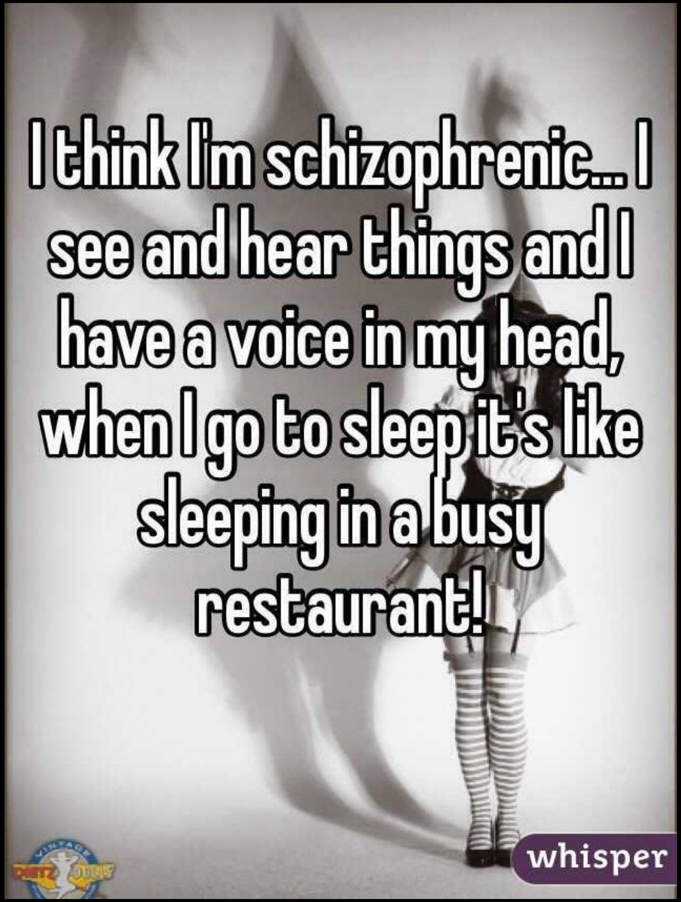
Giulia Fanti, Sewoong Oh, and Pramod Viswanath

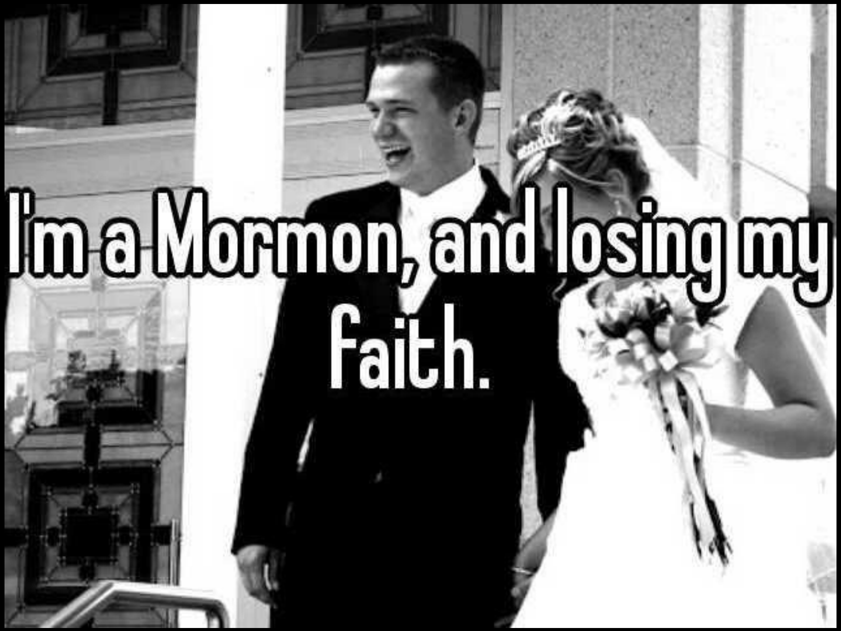
Political activism

Some people have important, sensitive things to say.

Personal confessions

Others have less important, but sensitive things to say.


I think I'm schizophrenic... I see and hear things and I have a voice in my head, when I go to sleep it's like sleeping in a busy restaurant!


I'm a Mormon, and losing my faith.

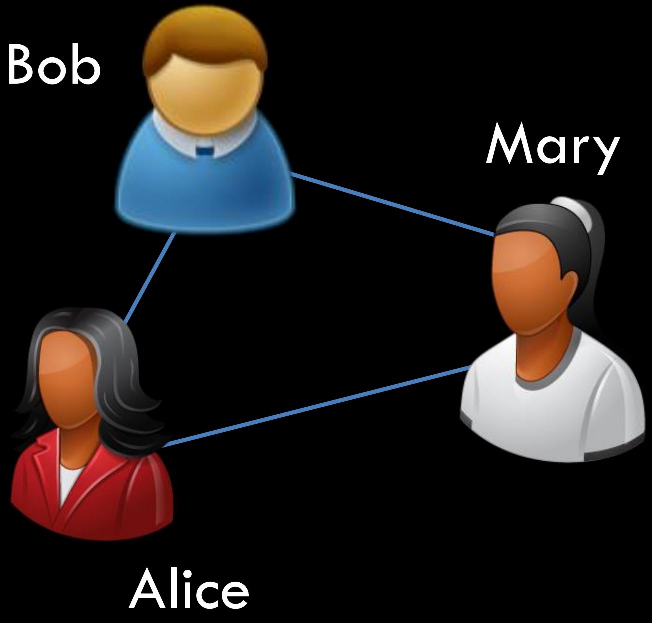
Existing anonymous messaging apps

Existing anonymous messaging apps

secret

whisper

Yik Yak

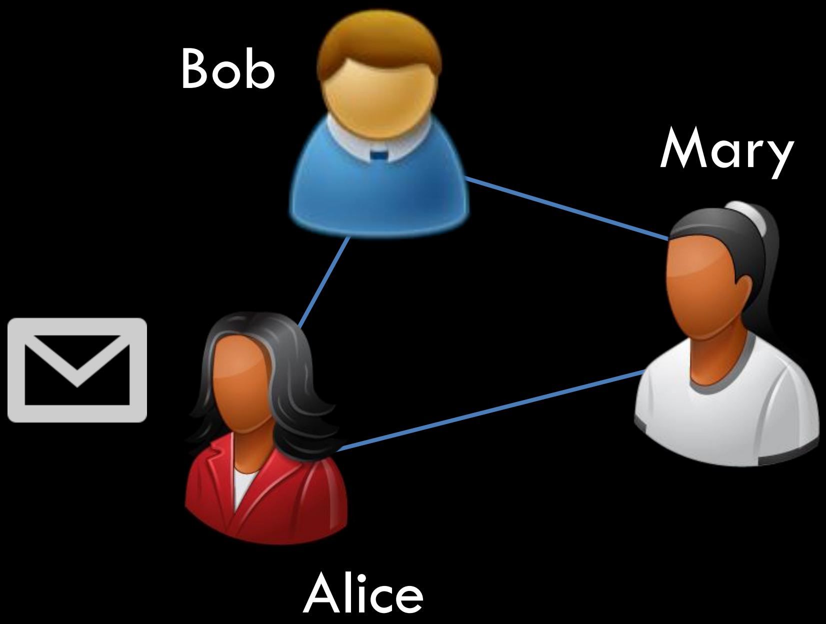


Existing anonymous messaging apps

secret

whisper

Yik Yak

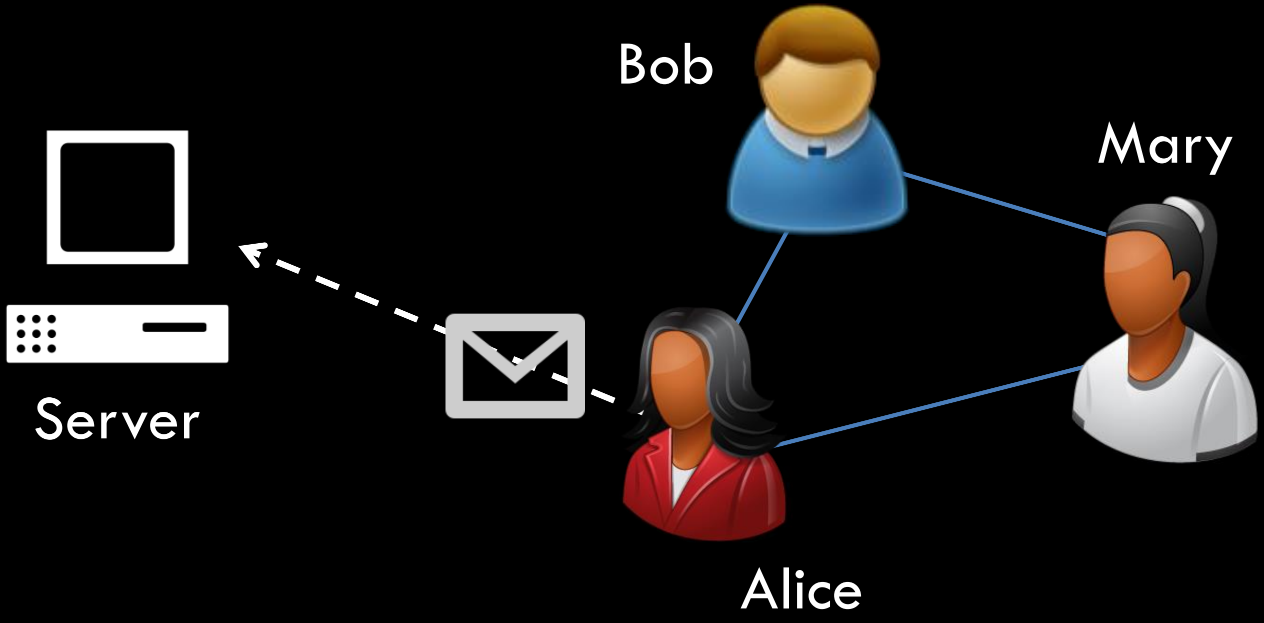


Existing anonymous messaging apps

secret

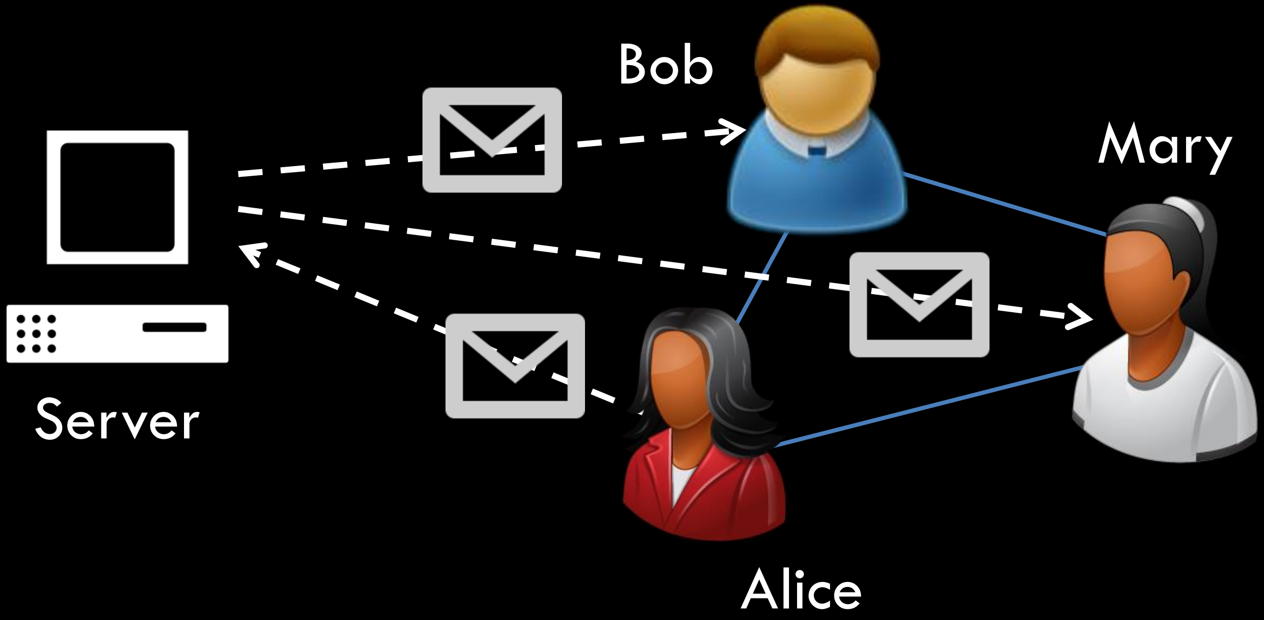
whisper

Yik Yak



Existing anonymous messaging apps

secret whisper Yik Yak

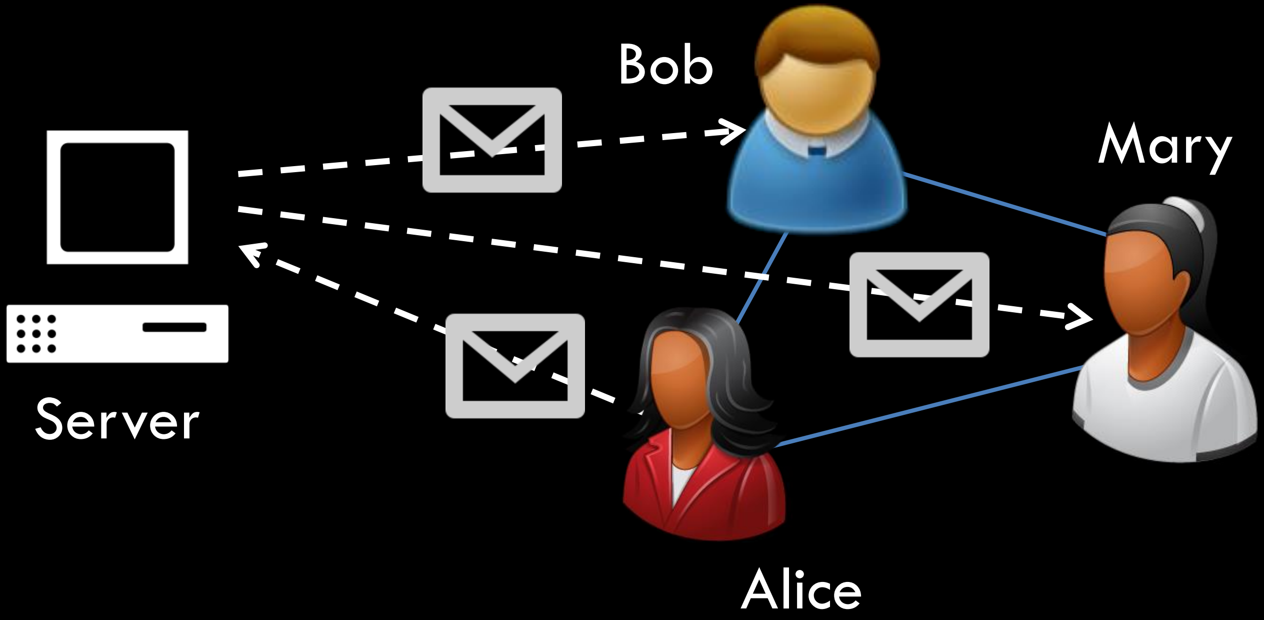


Existing anonymous messaging apps

secret

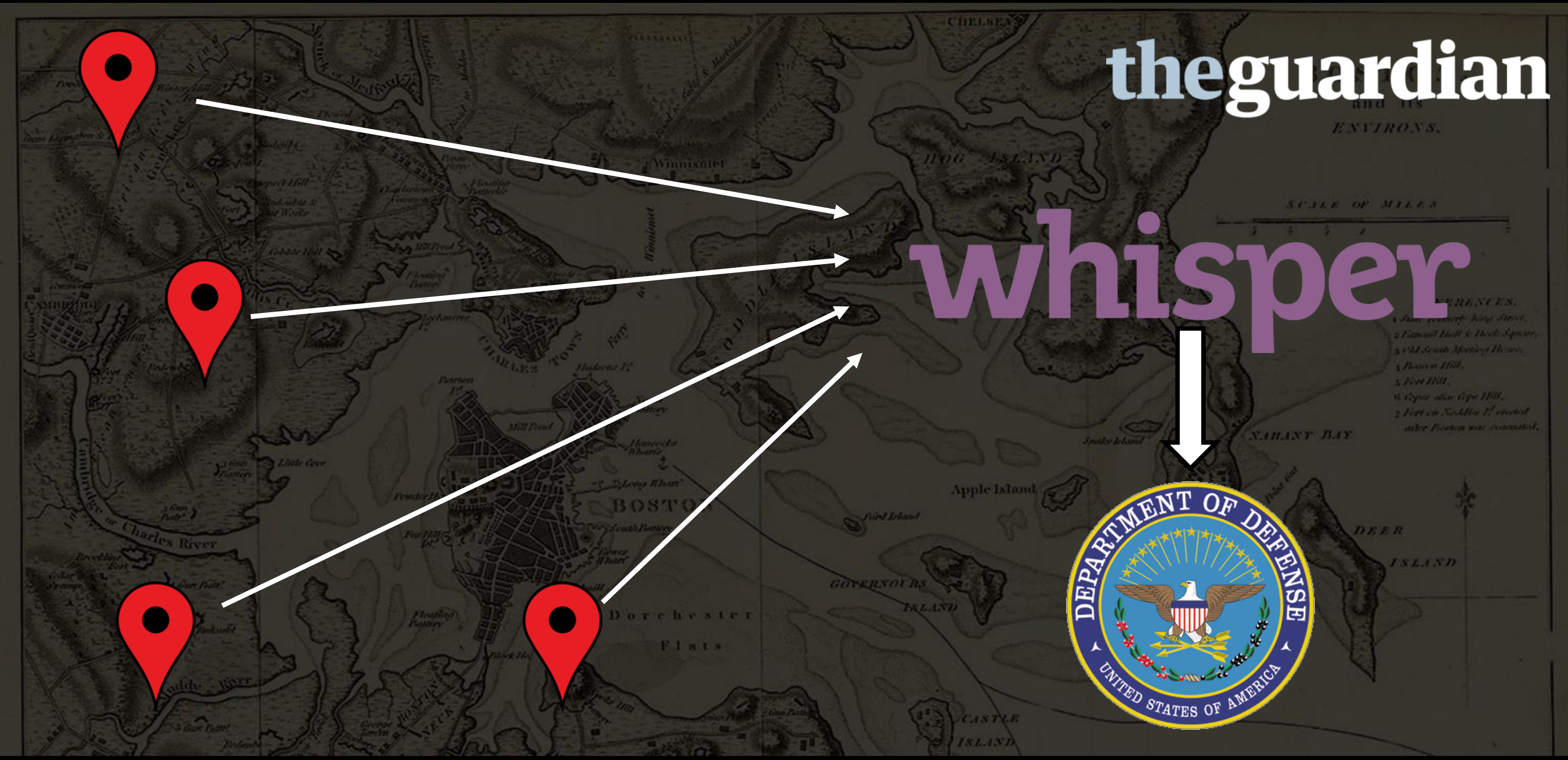
whisper

Yik Yak



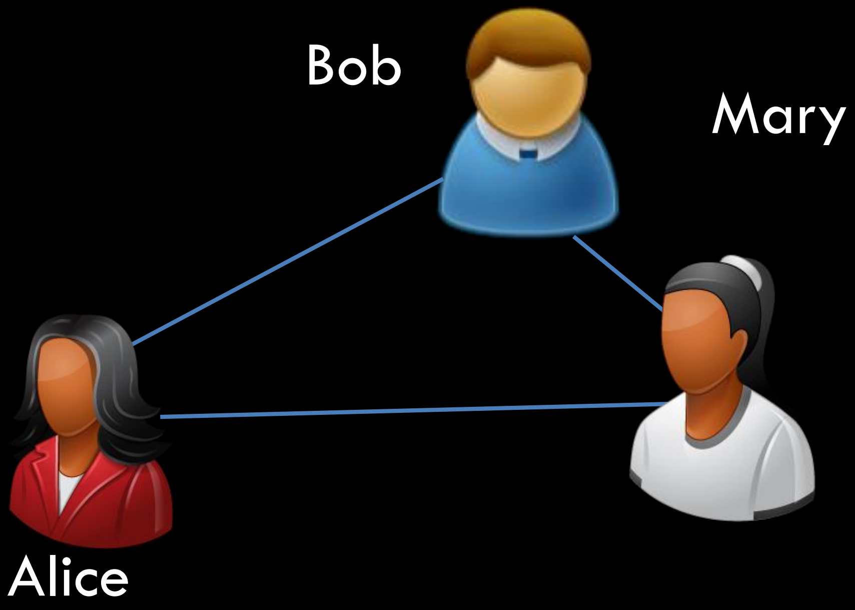
centralized networks are not truly anonymous!

Compromises in anonymity

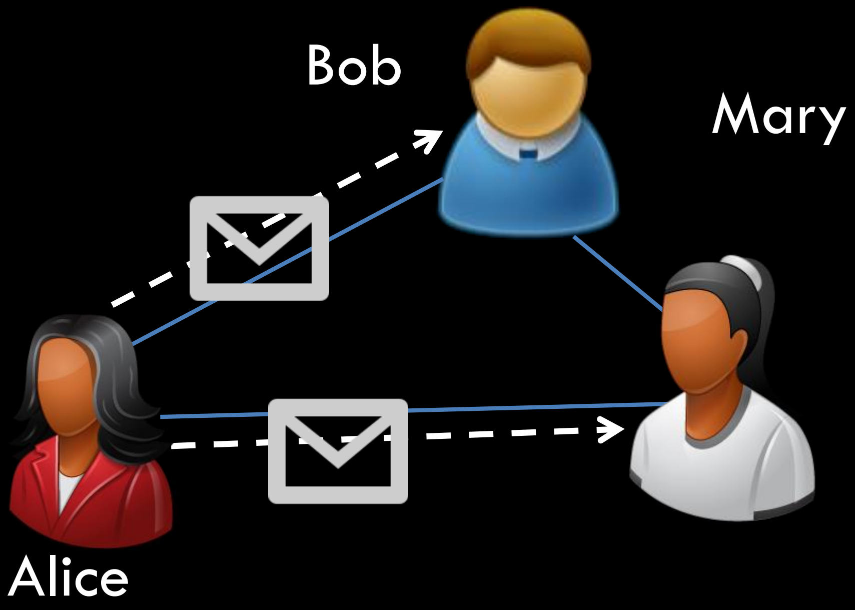


anonymity loss extends beyond the network

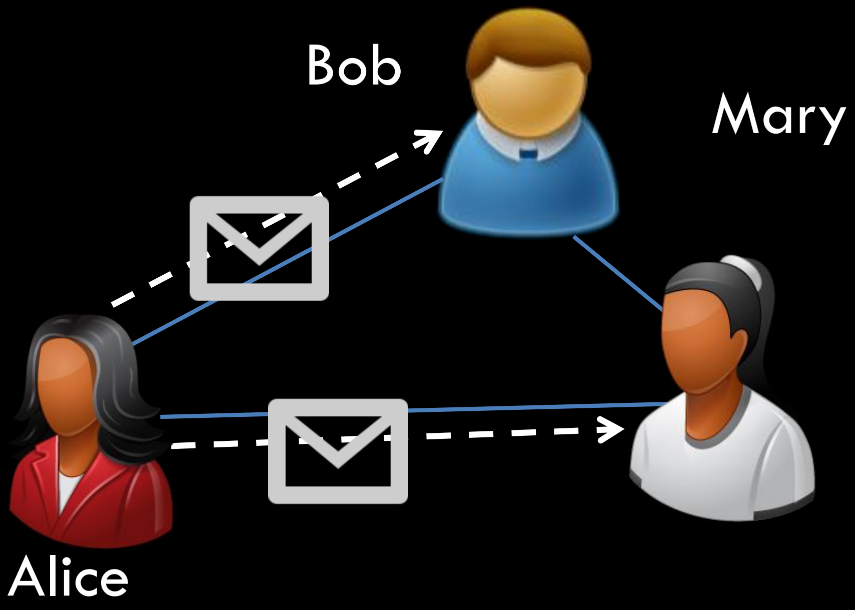
Distributed messaging



Distributed messaging

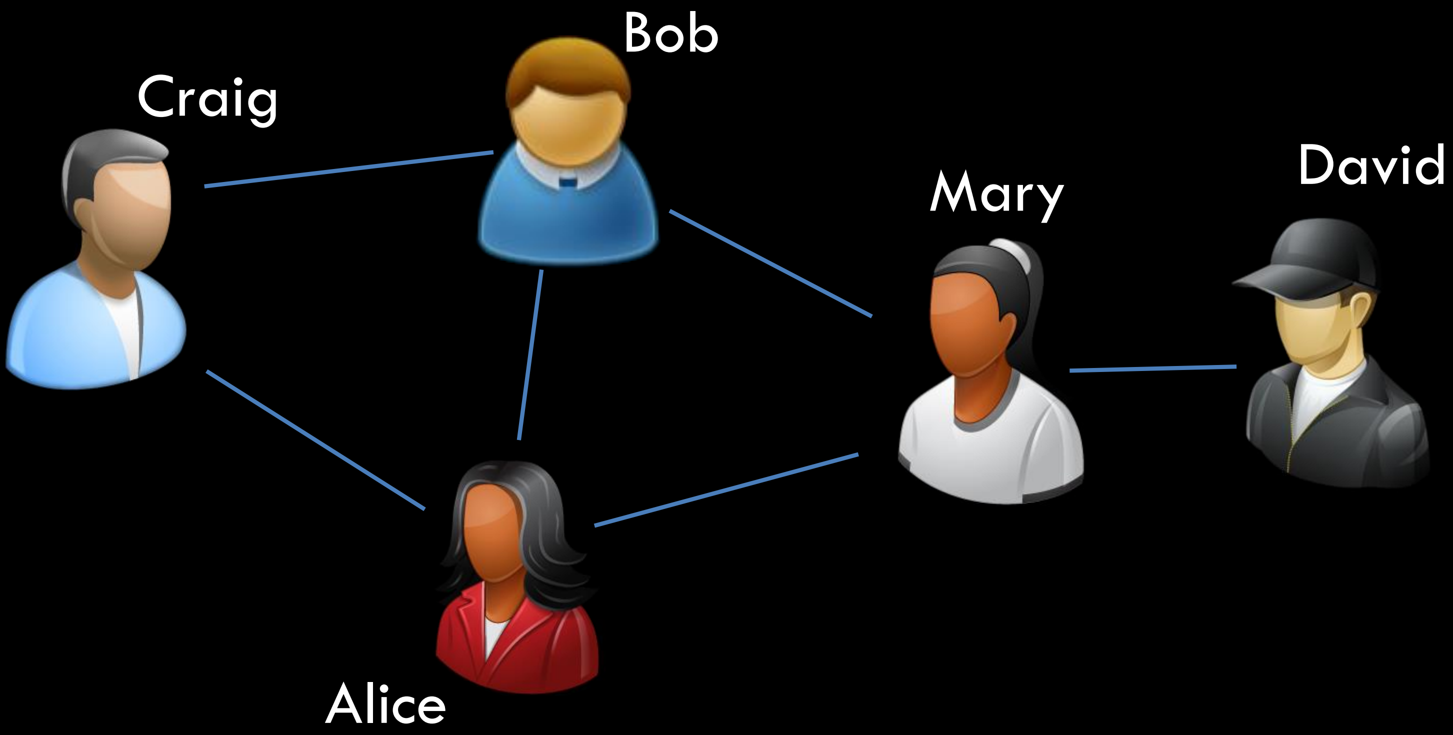


Distributed messaging

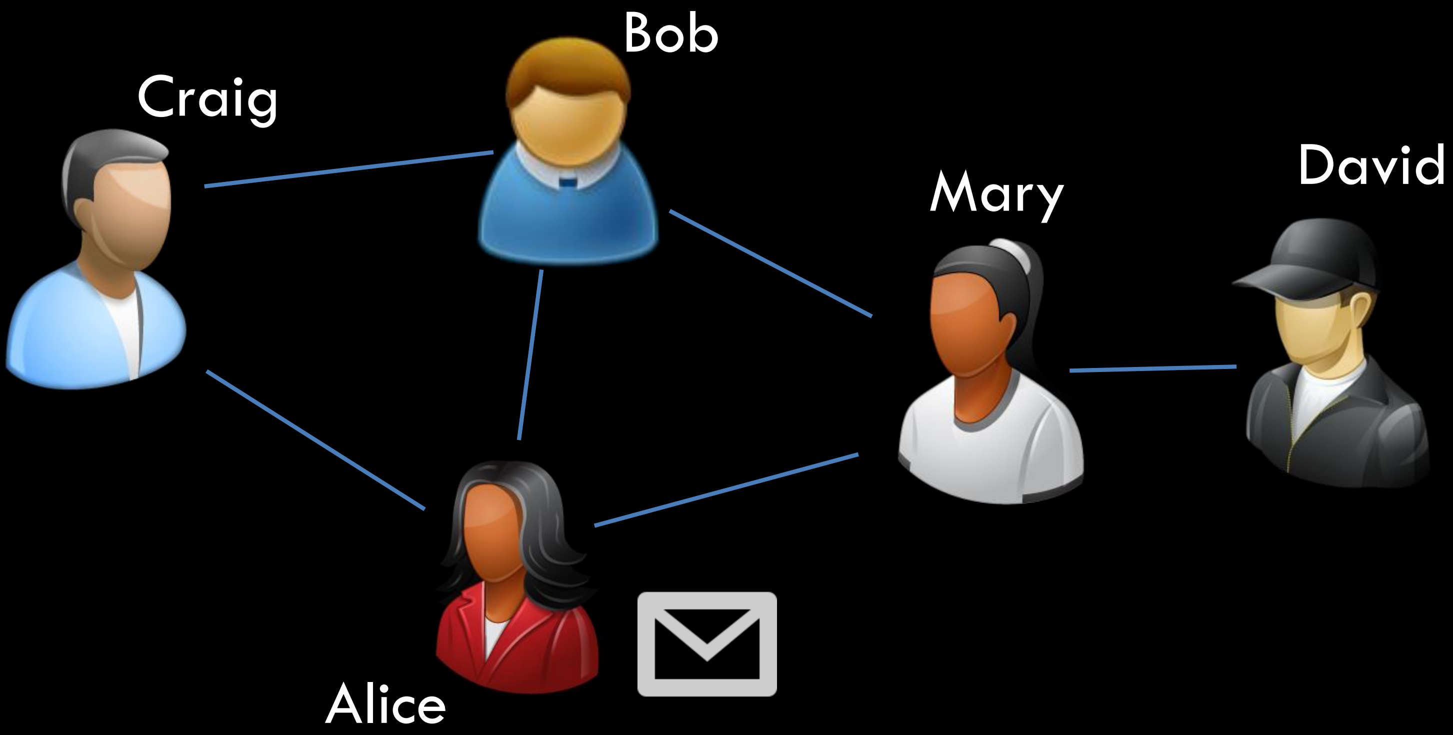


what can an adversary do?

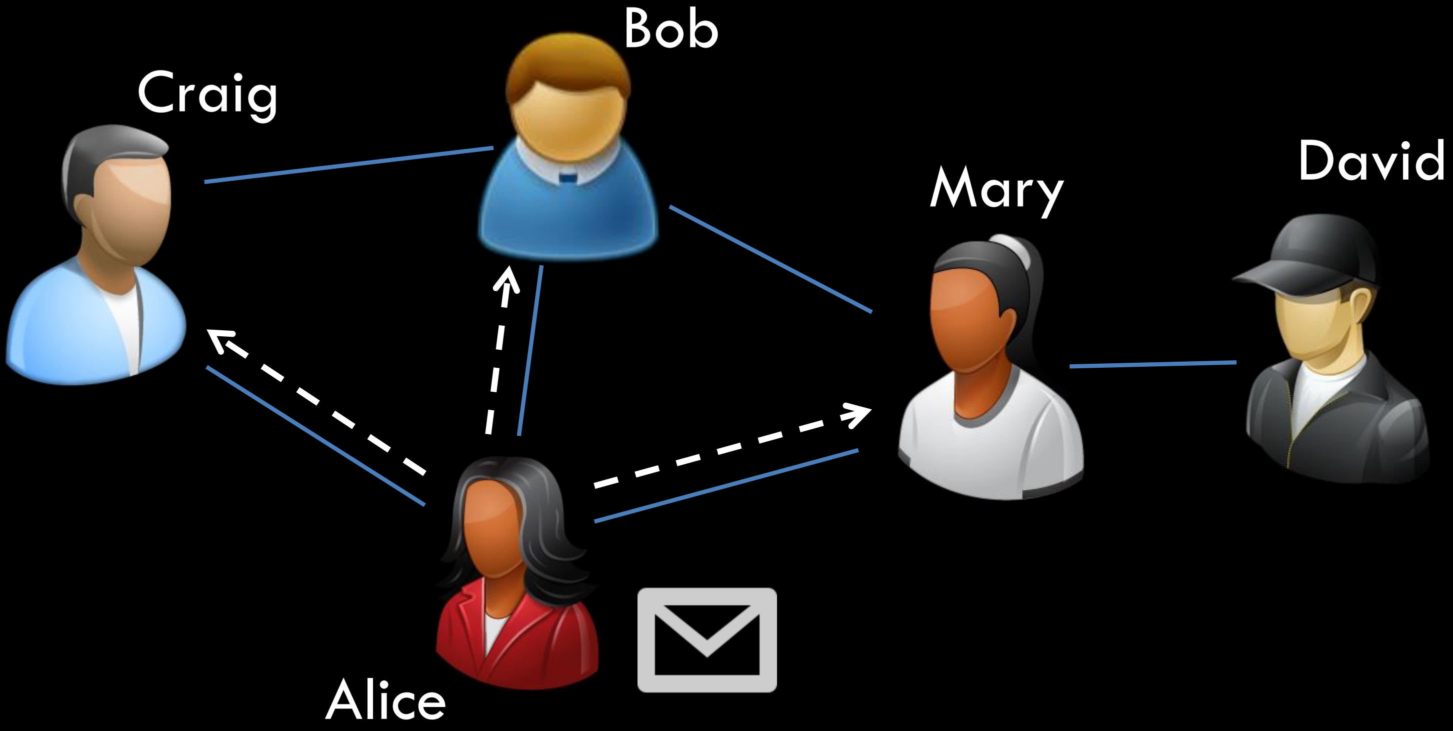
Adversarial model



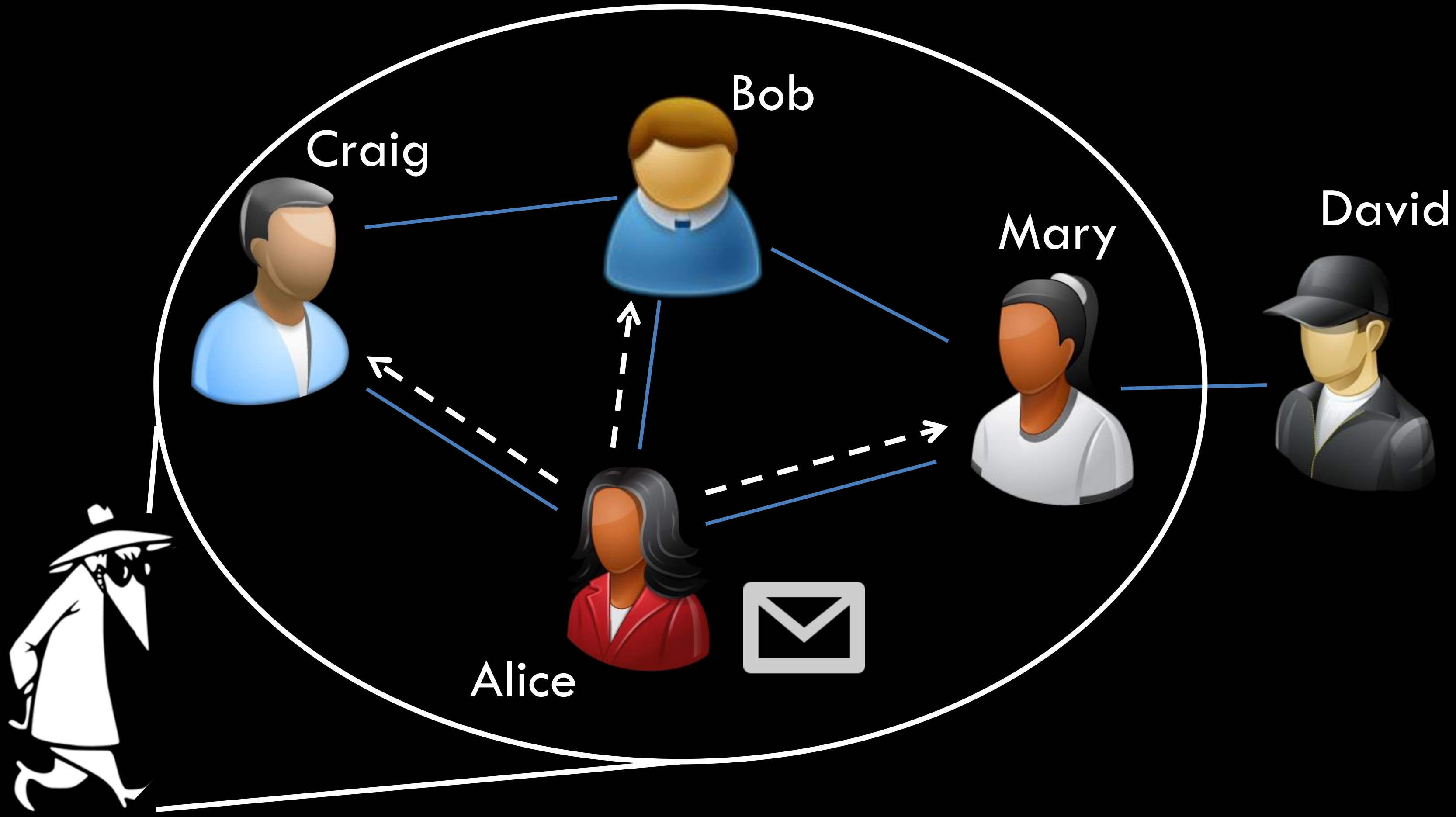
Adversarial model



Adversarial model

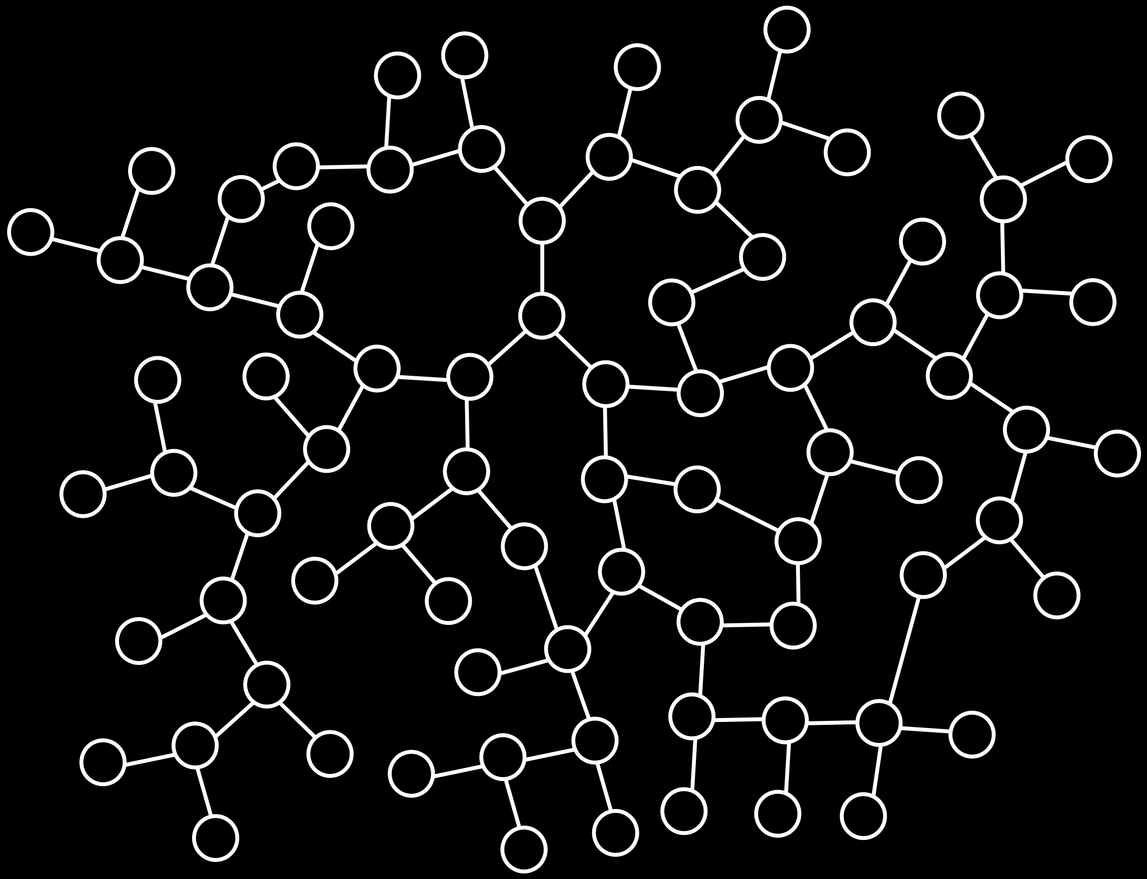


Adversarial model



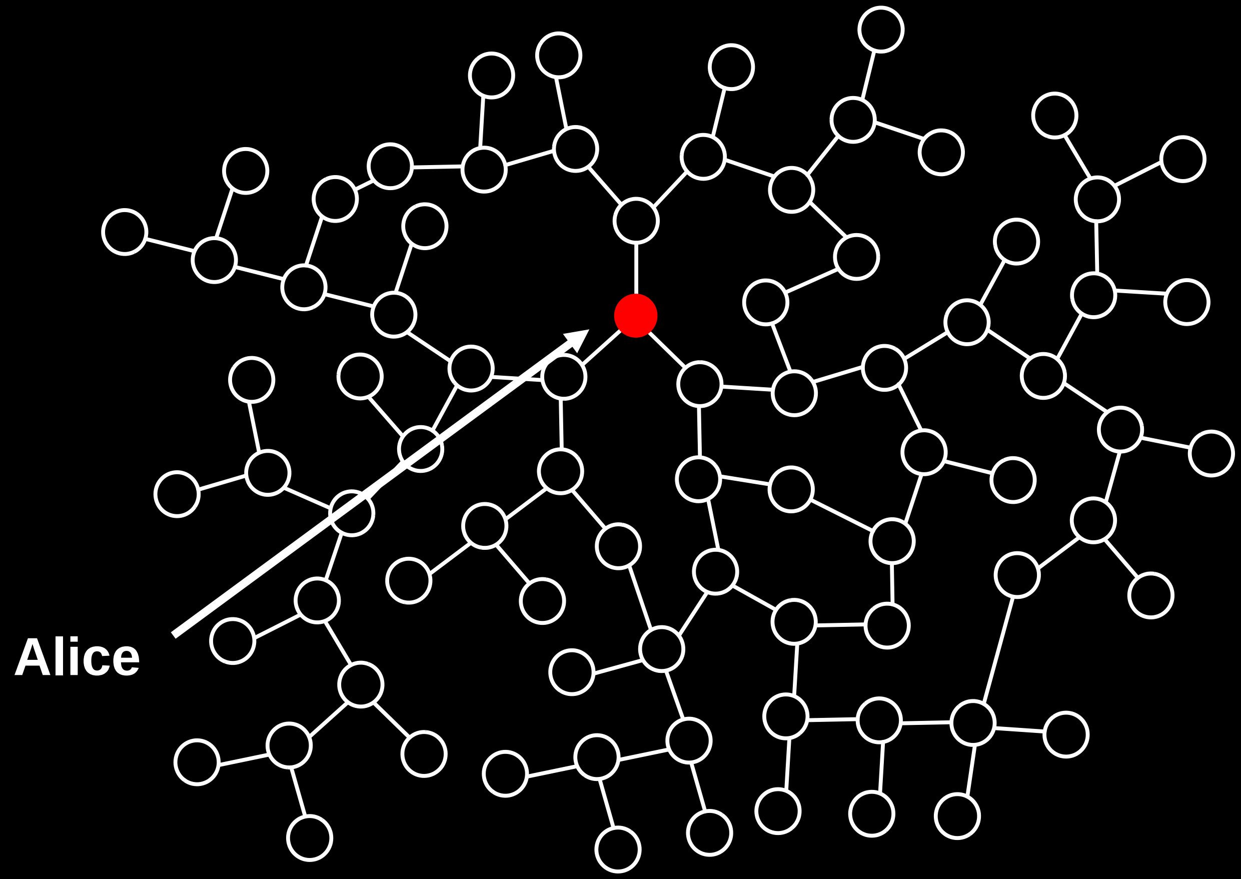
the adversary can figure out who got the message

Information flow in social networks

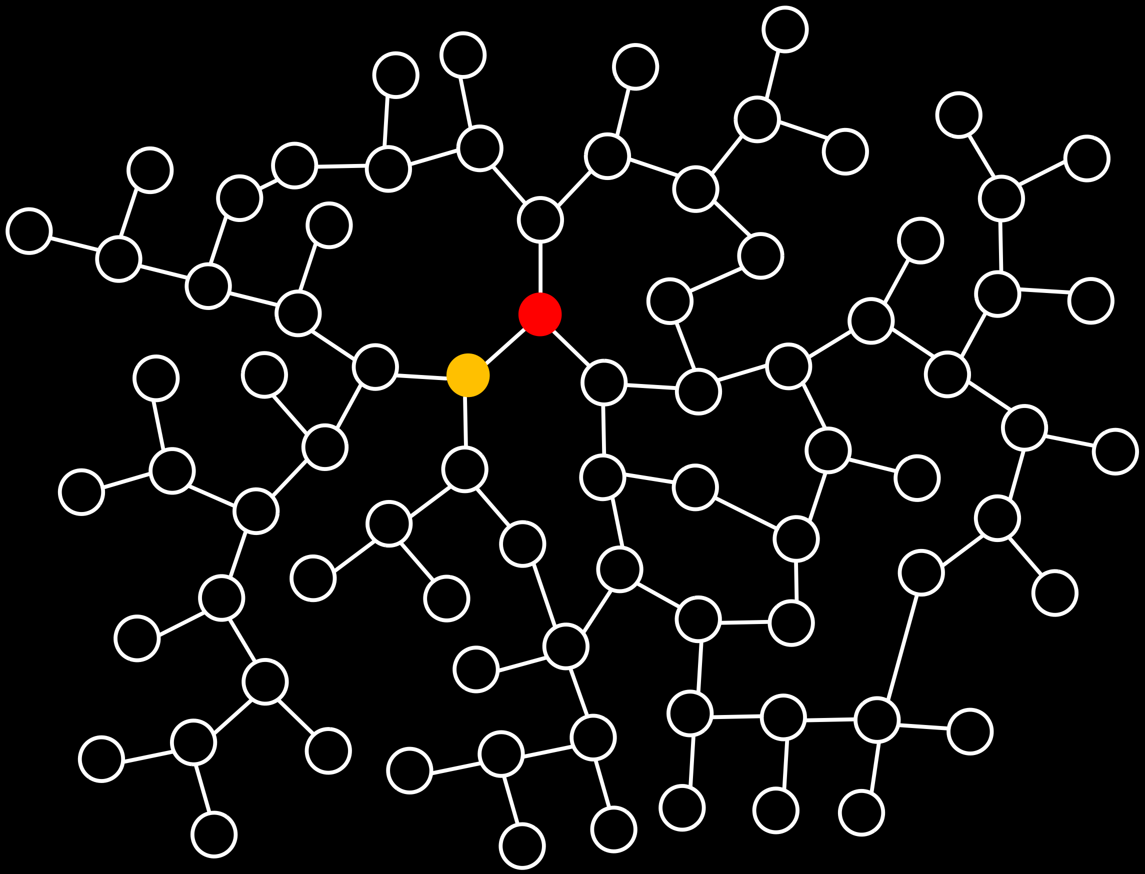


- G is the graph representing the social network

Information flow in social networks

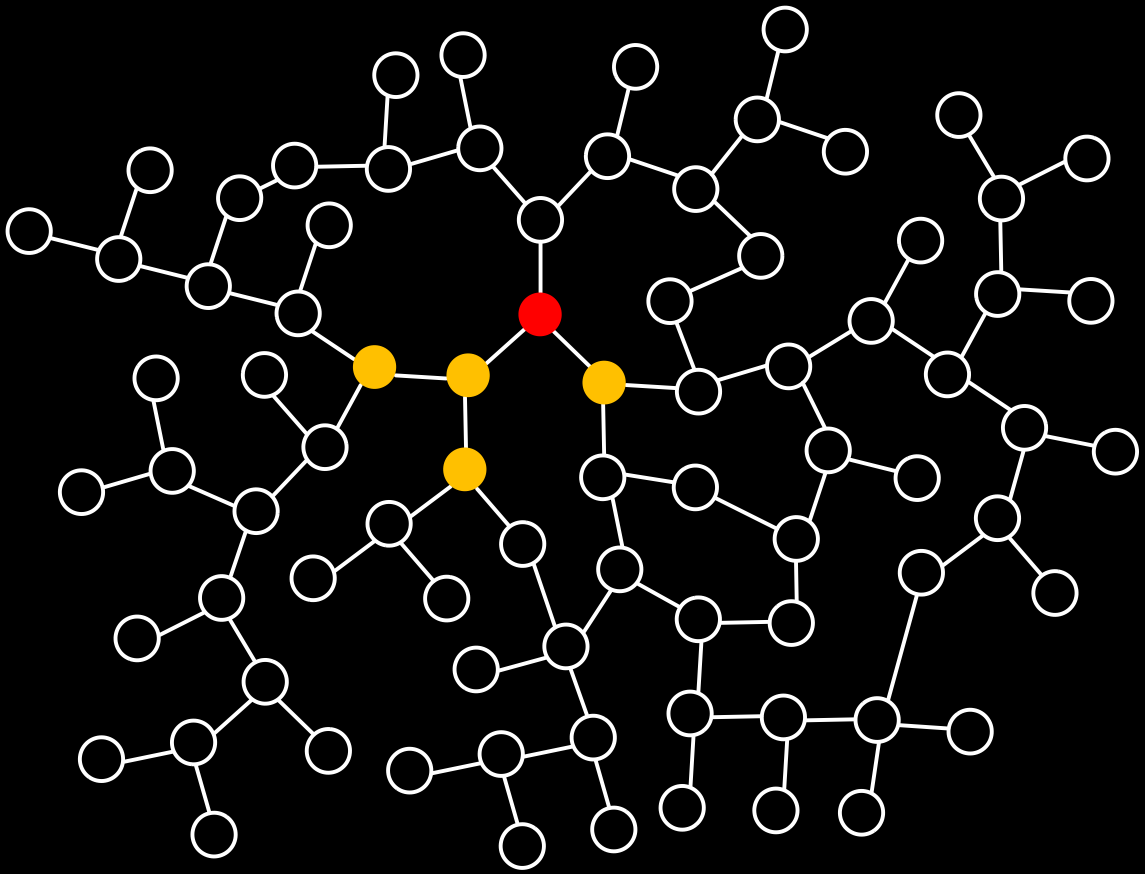


Information flow in social networks



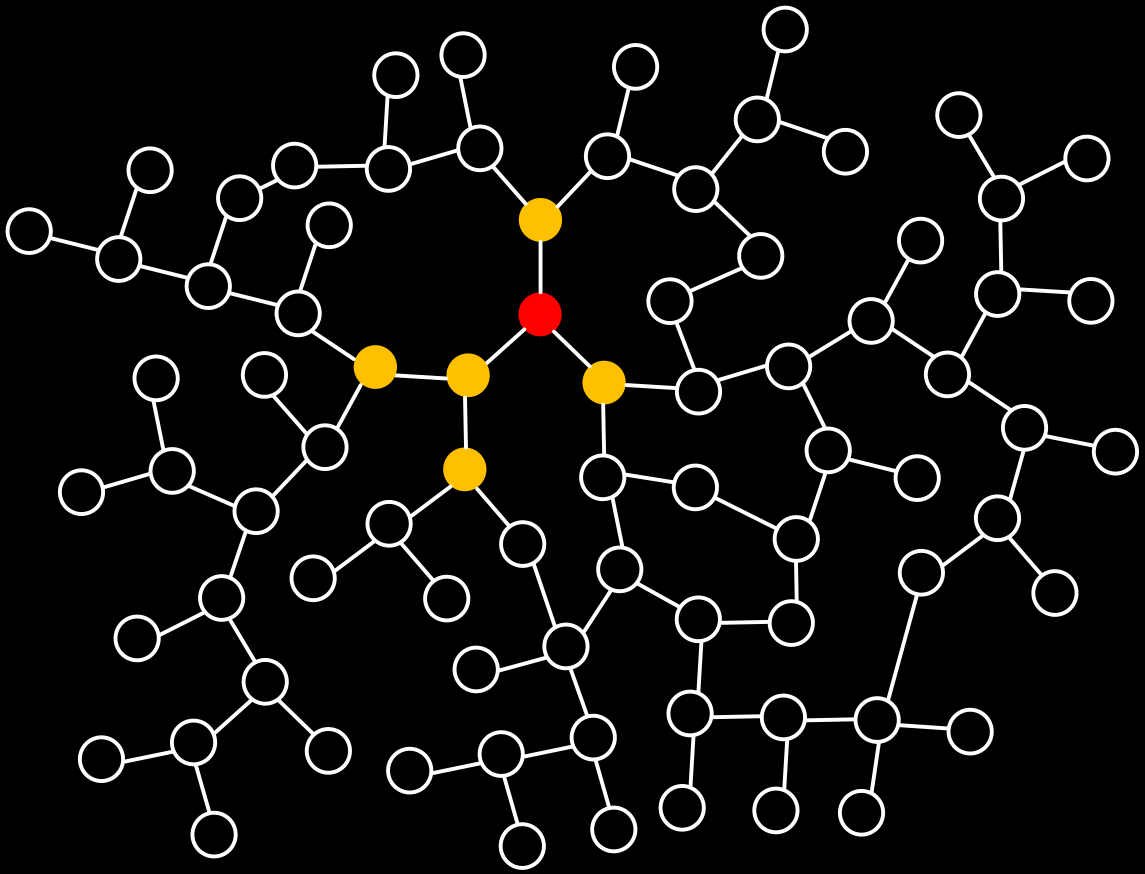
- Alice passes the message to her friends

Information flow in social networks



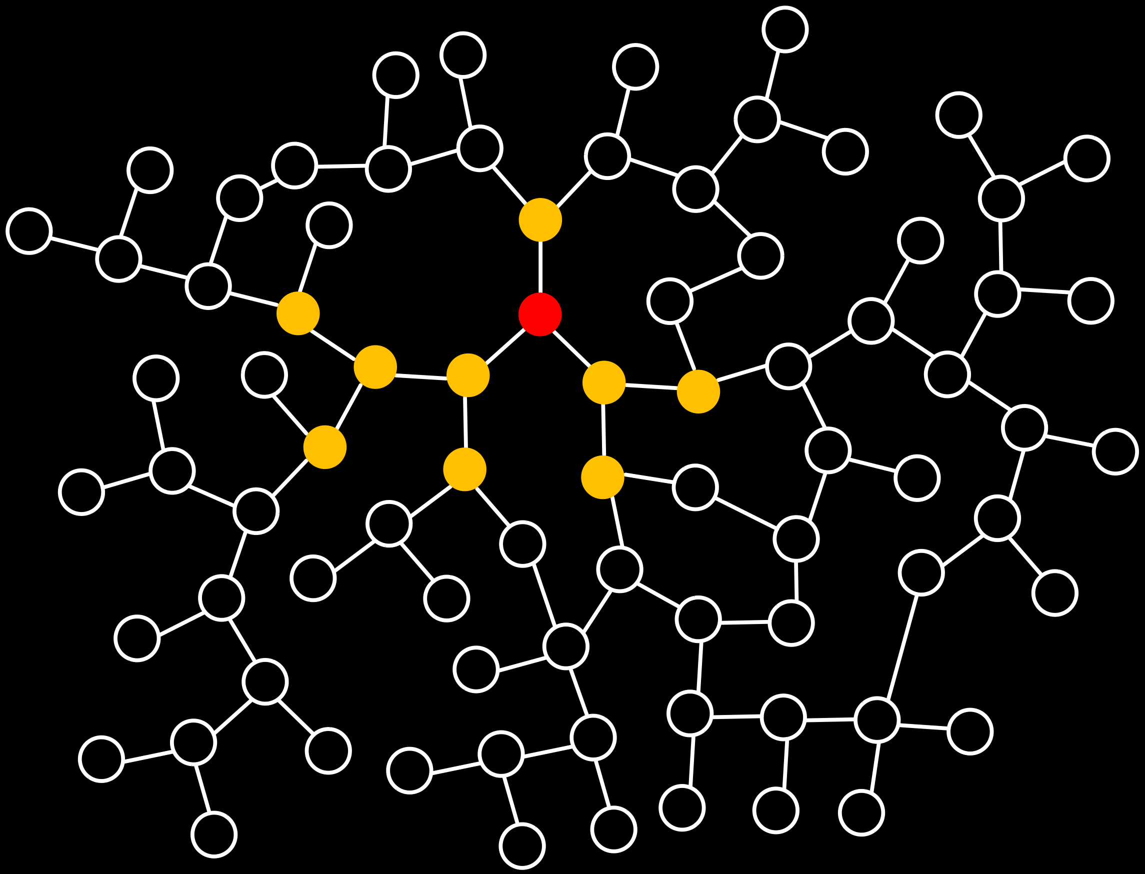
- her friends pass the message to theirs

Information flow in social networks



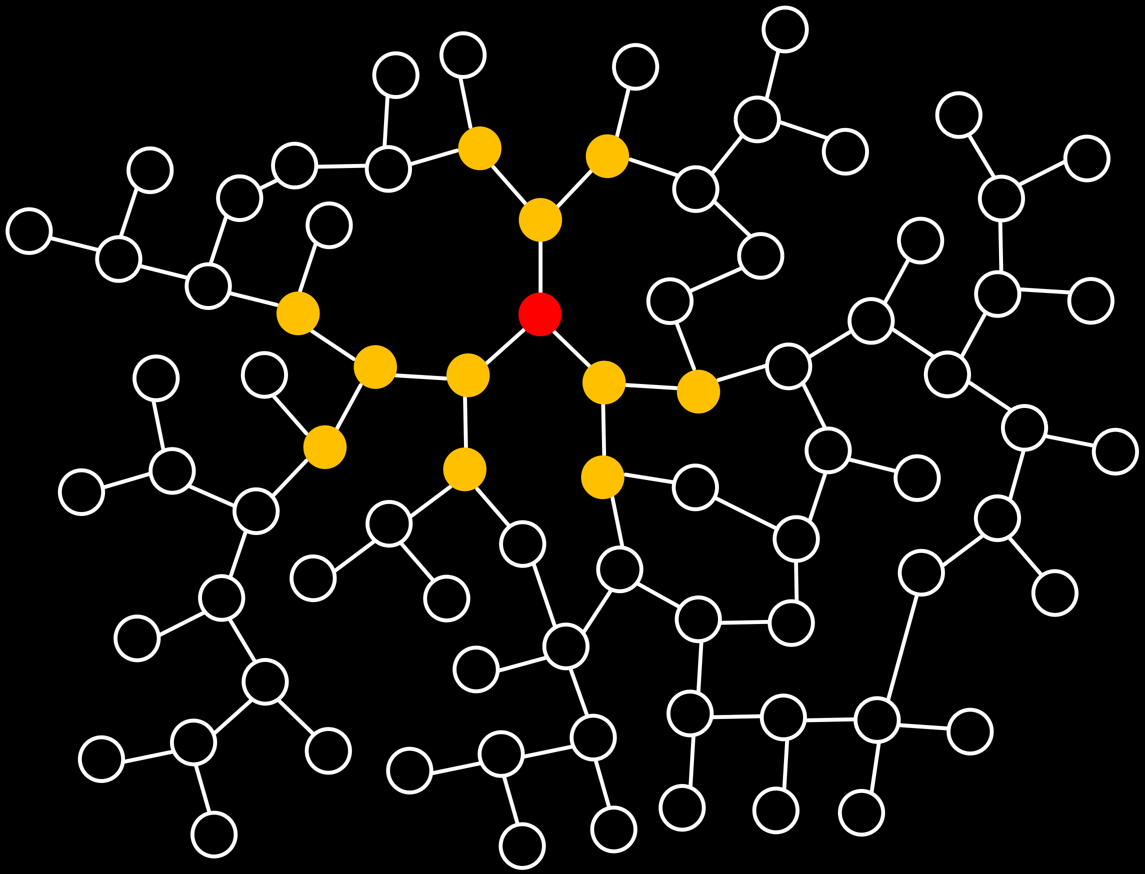
- the message spreads in **all directions** at the **same rate**

Information flow in social networks



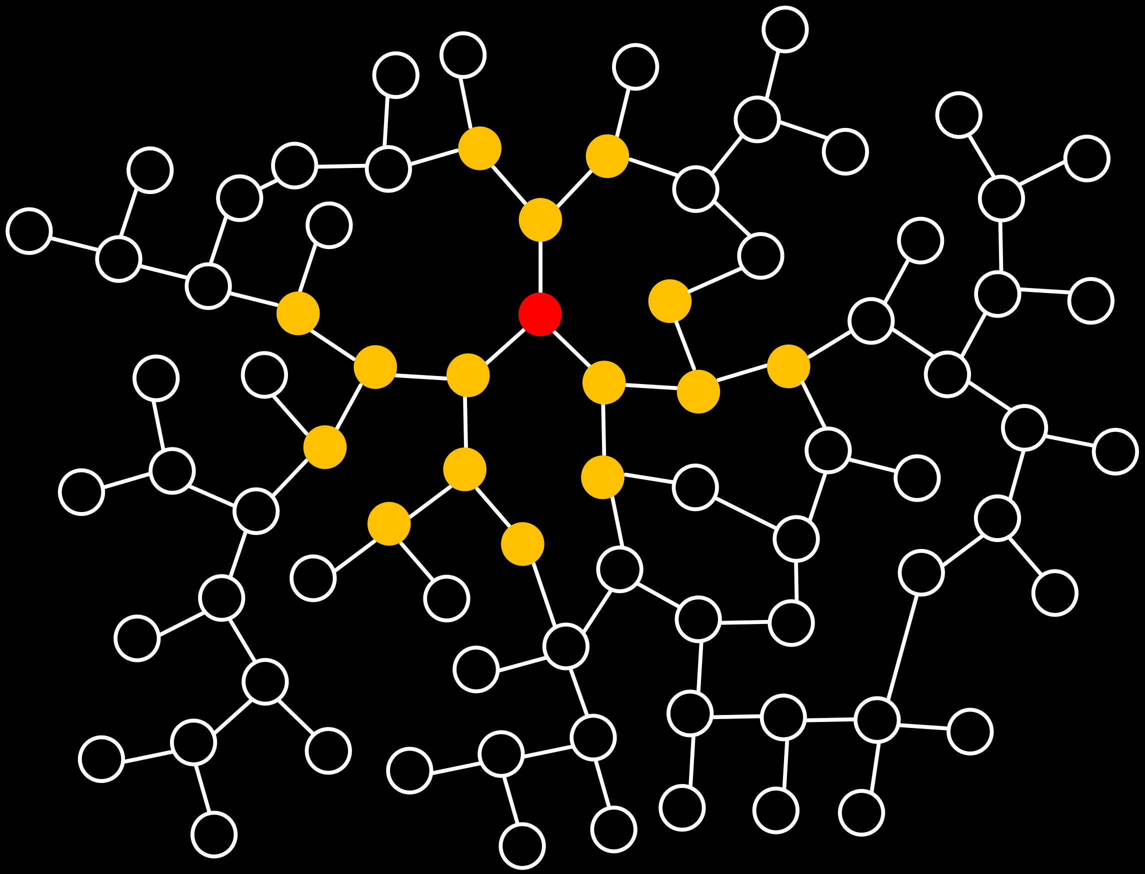
- the message spreads in **all directions** at the **same rate**

Information flow in social networks



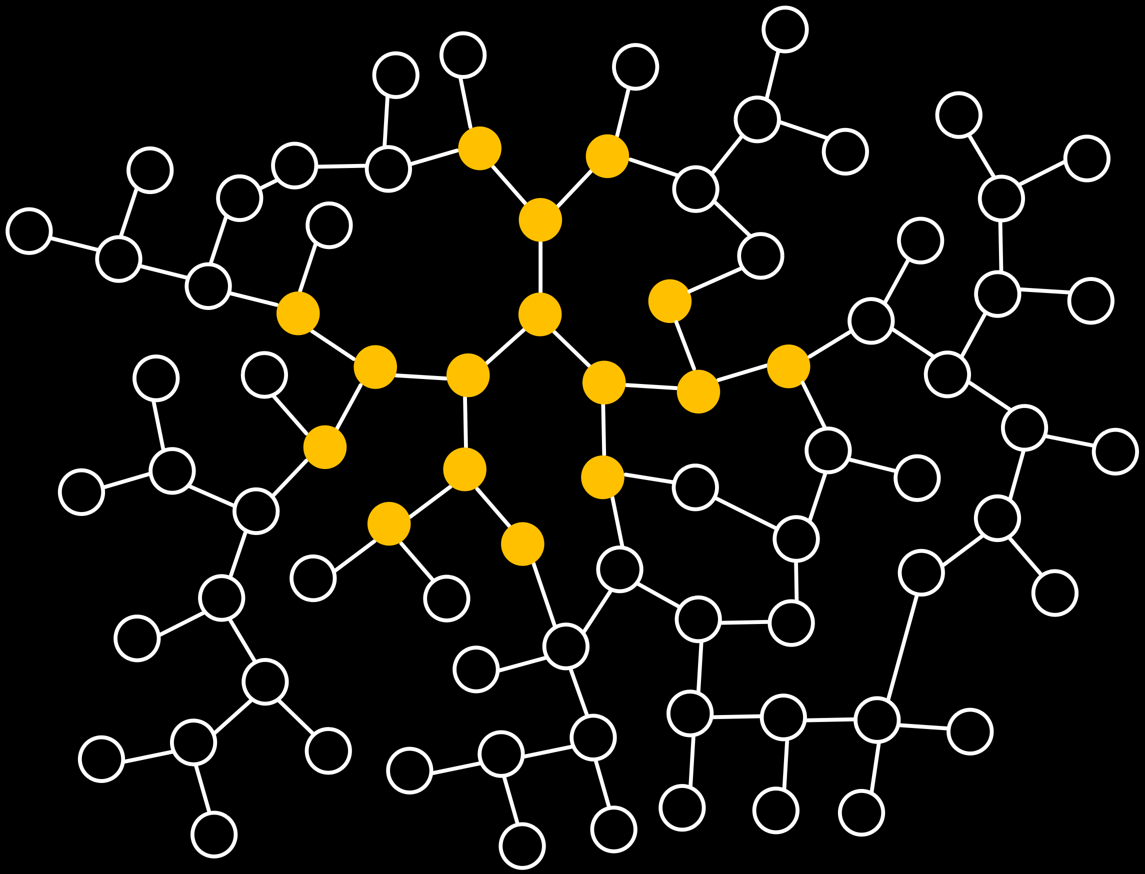
- the message spreads in **all directions** at the **same rate**

Information flow in social networks



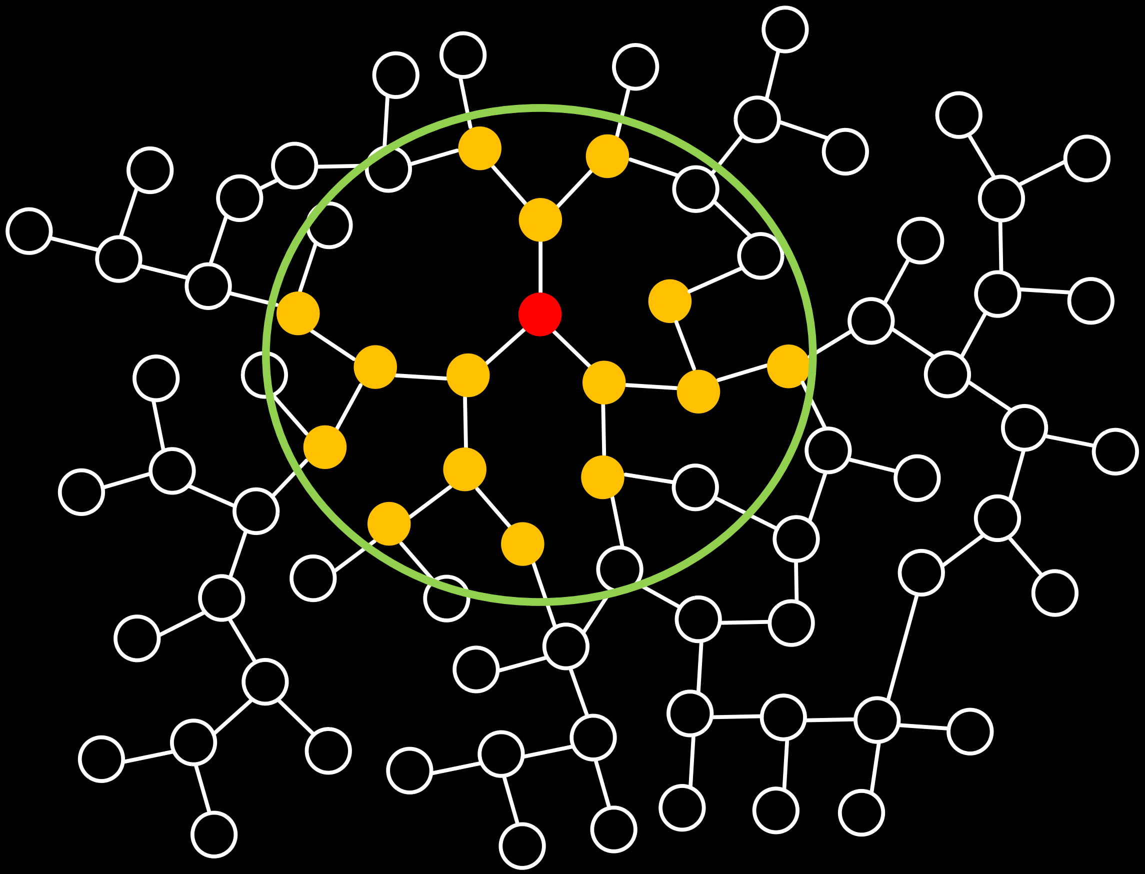
- this **spreading model** is known as the **diffusion model**

Adversary’s observation



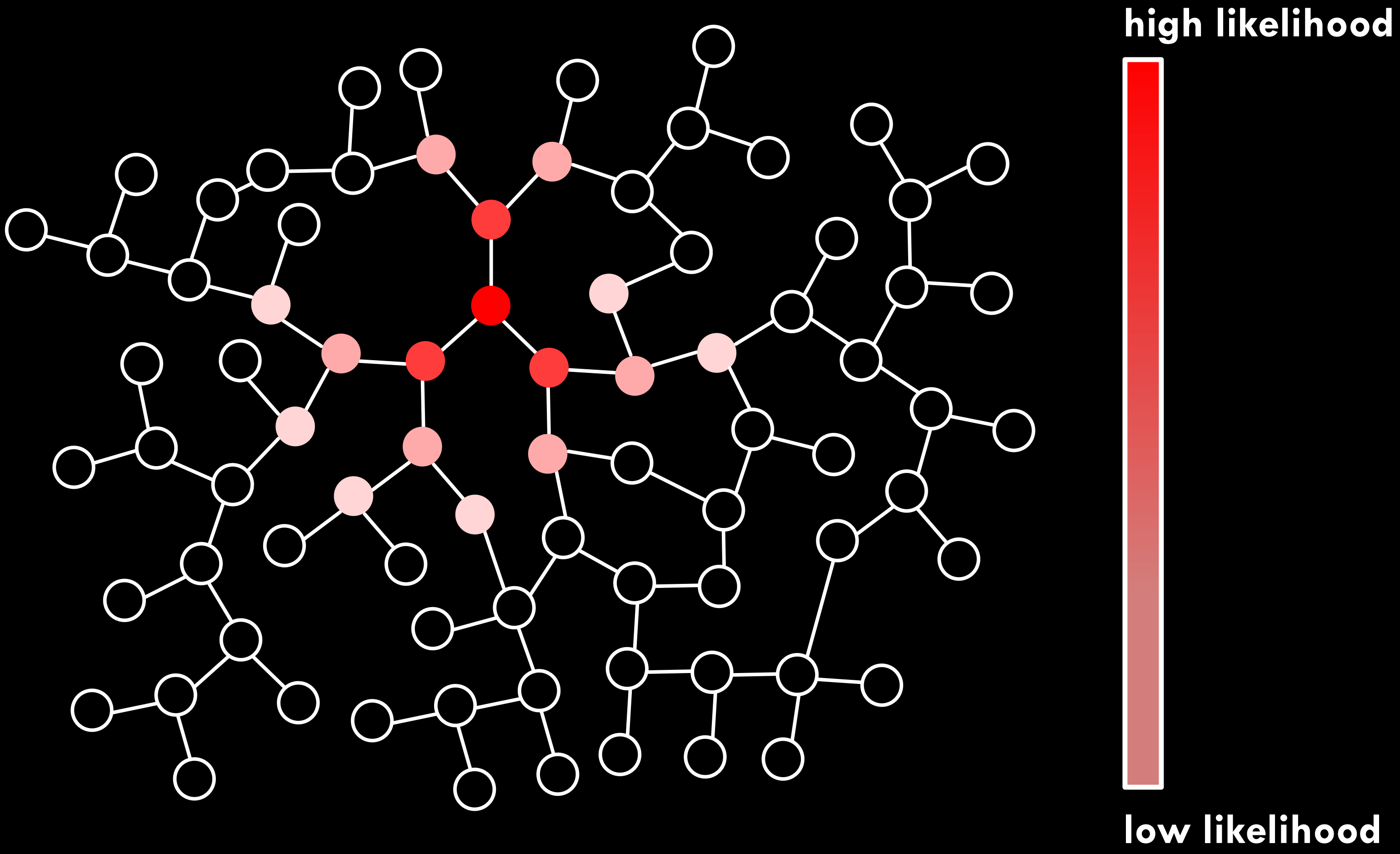
can the adversary locate the message author?

Concentration around the center



- the **message author** is in the **“center”** with high probability

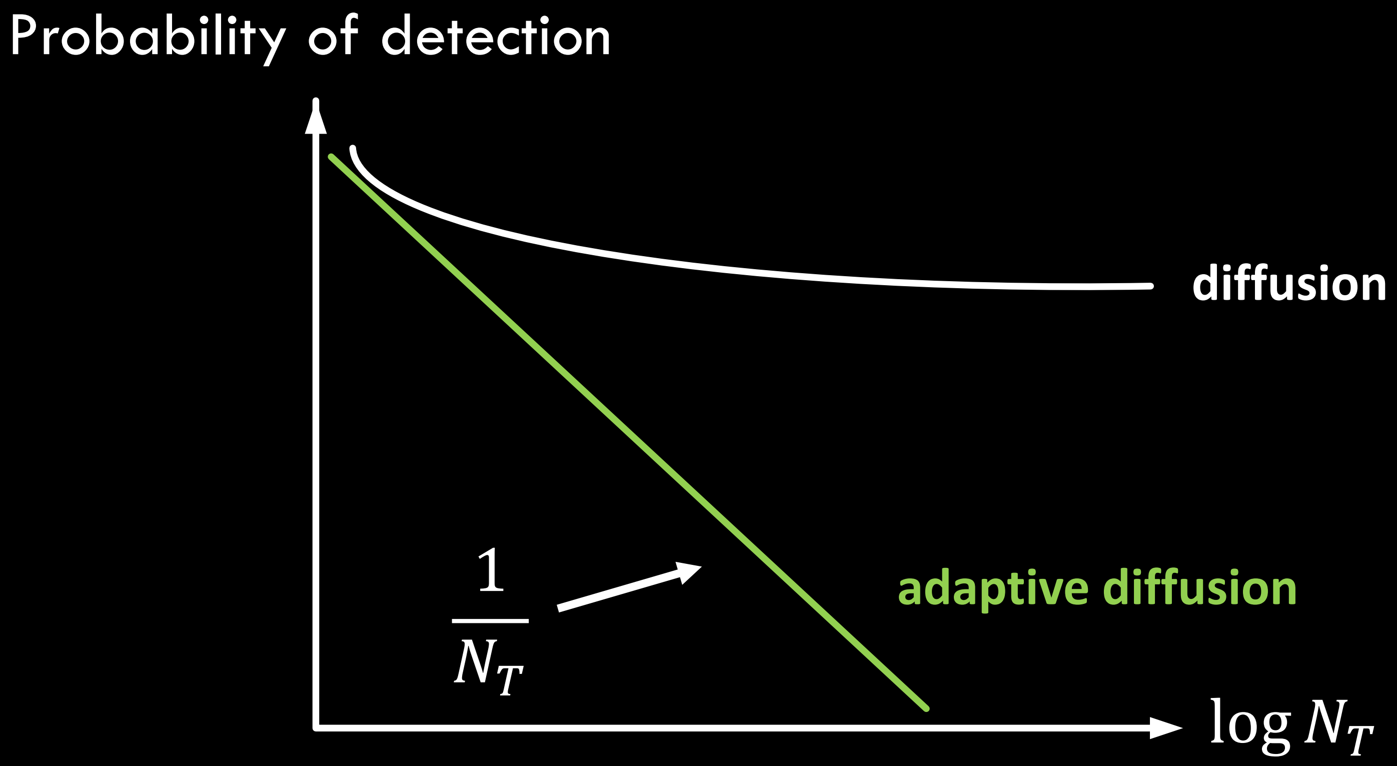
Rumor source identification



diffusion does not provide anonymity

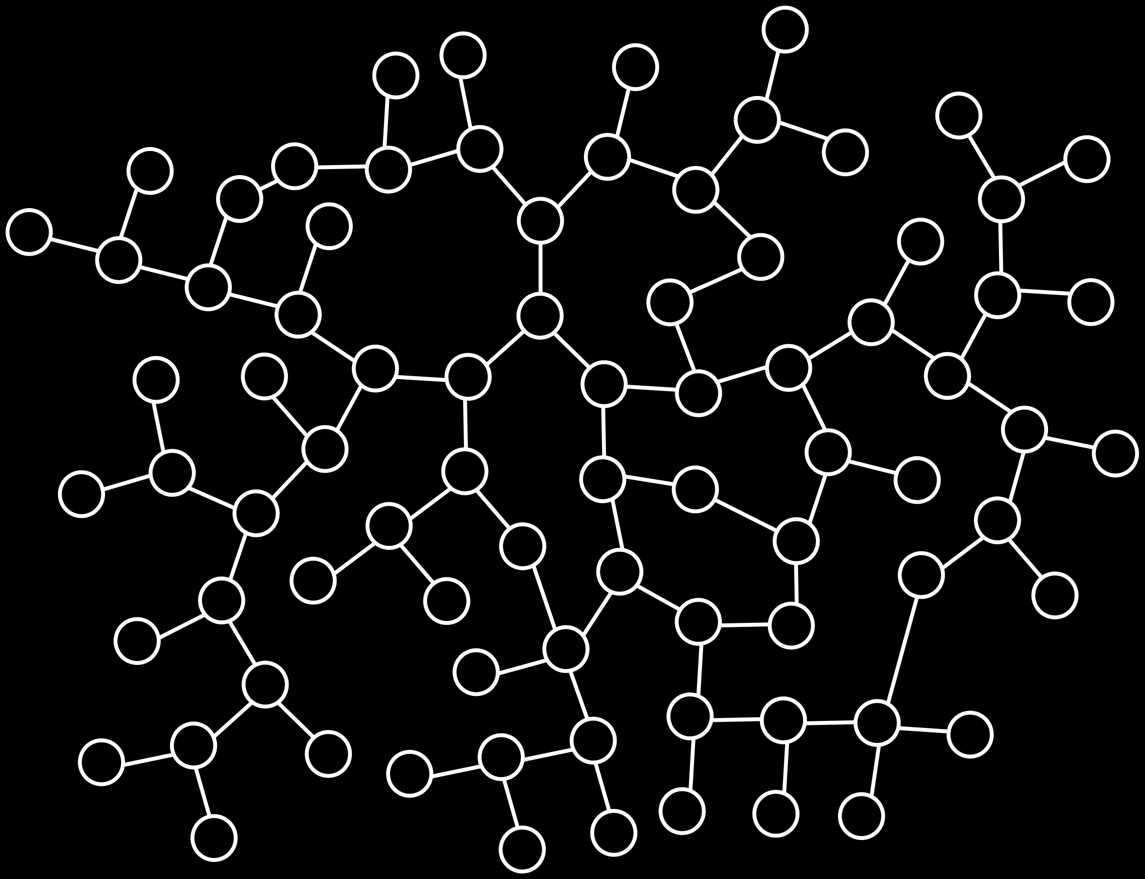
[*Shah, Zaman* 2011]

Our goal

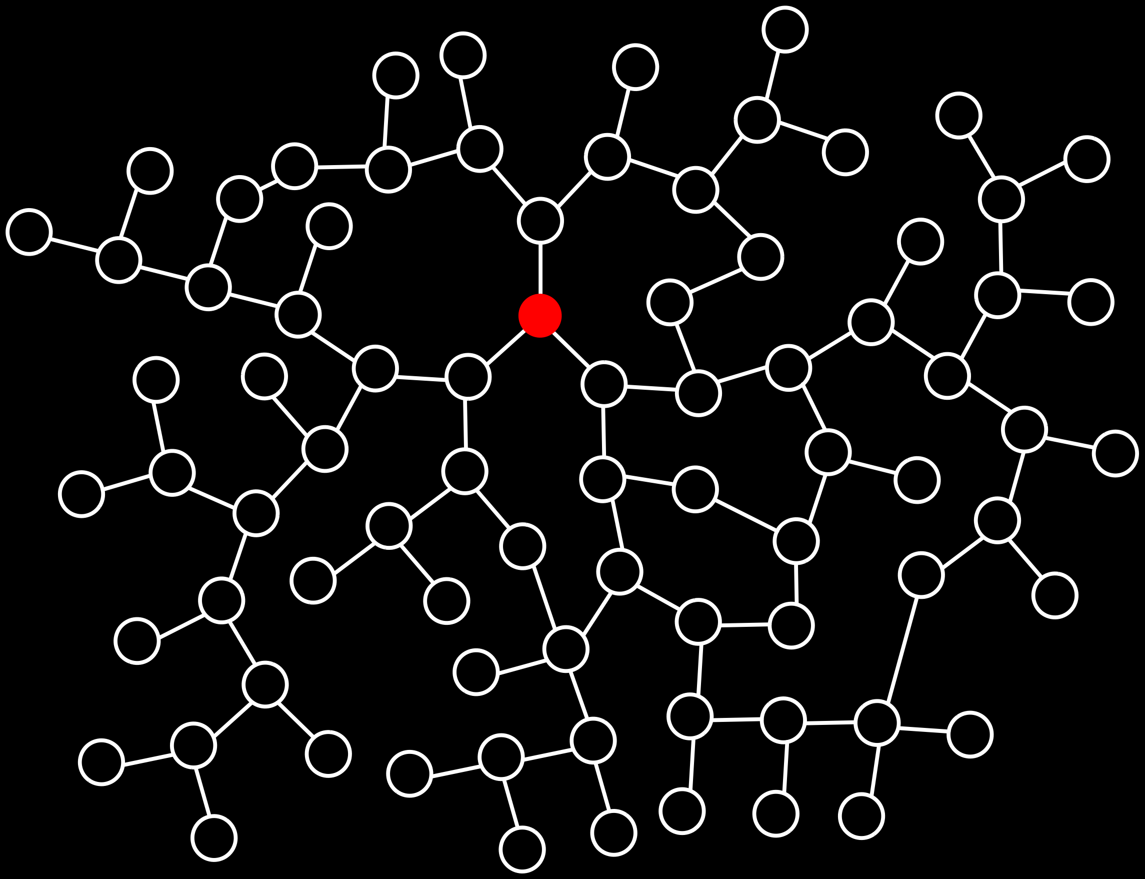


- N_T: **expected number** of nodes with the message at time T

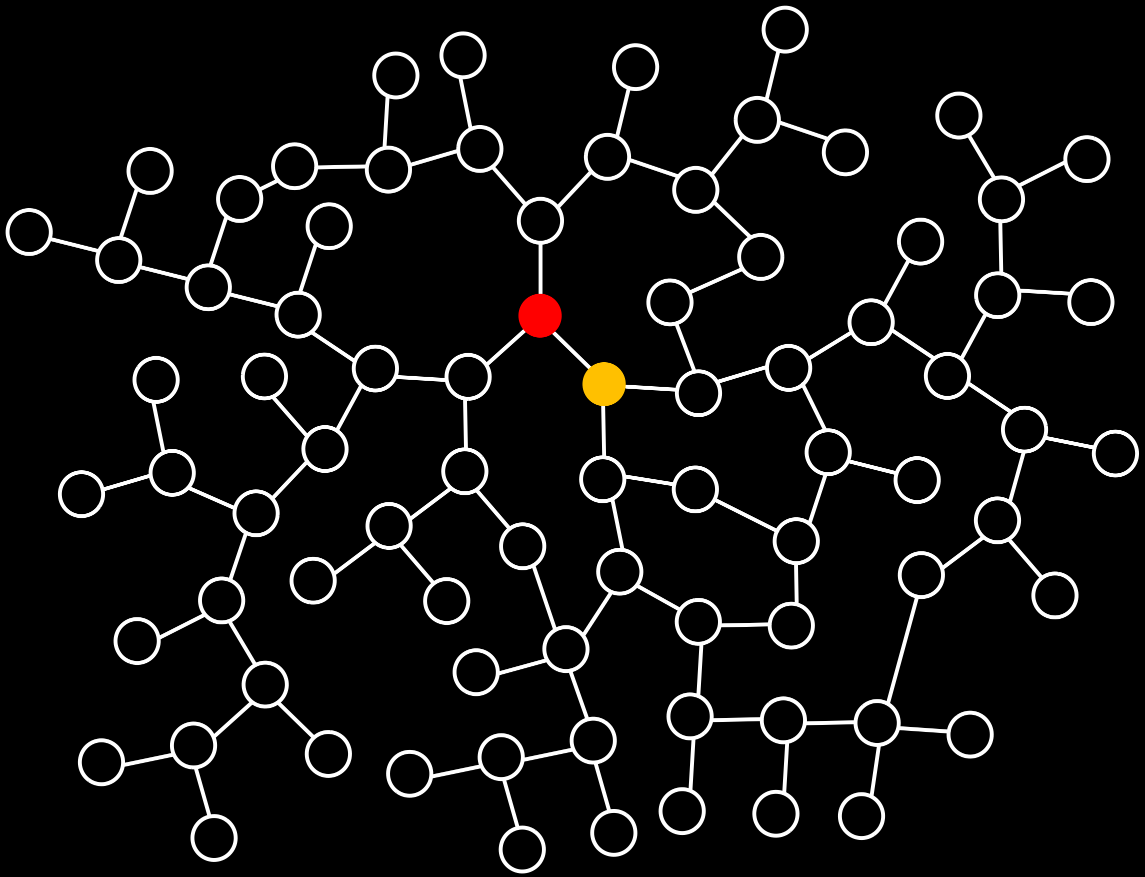
Main result: adaptive diffusion



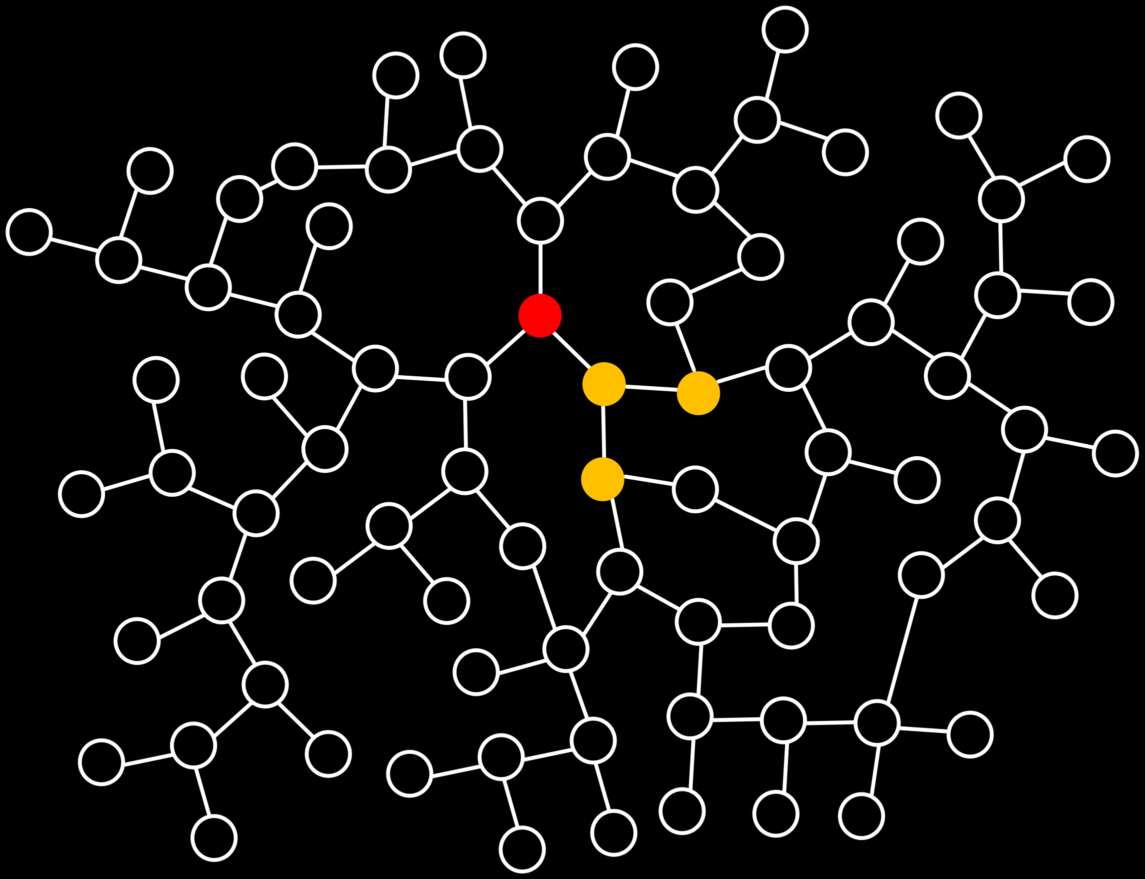
Main result: adaptive diffusion



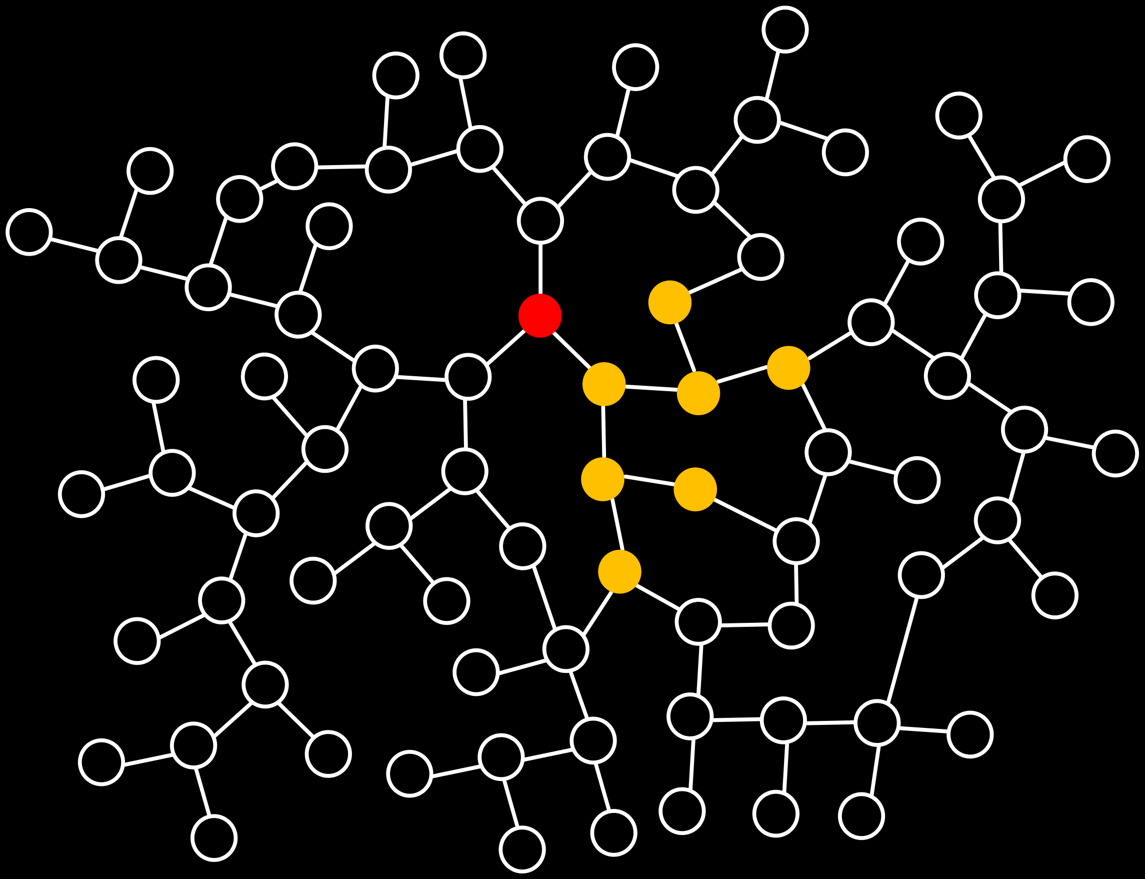
Main result: adaptive diffusion



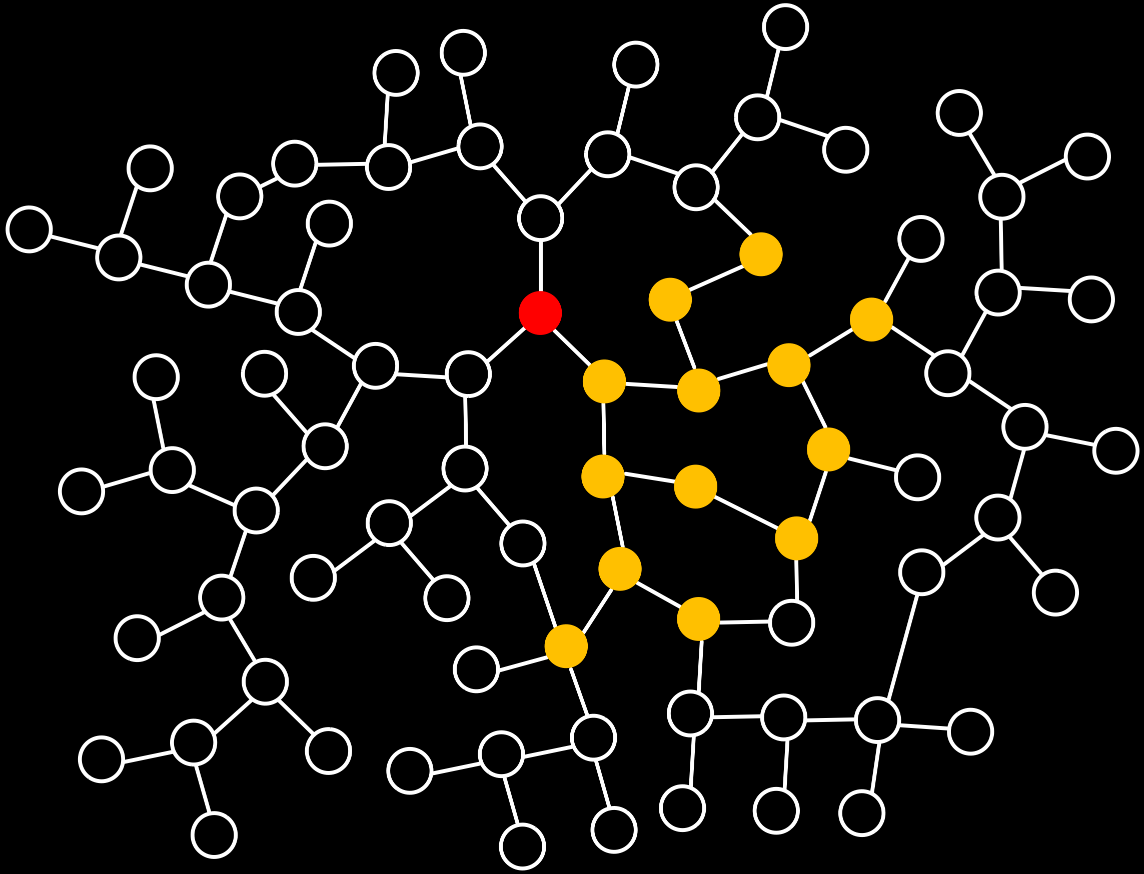
Main result: adaptive diffusion



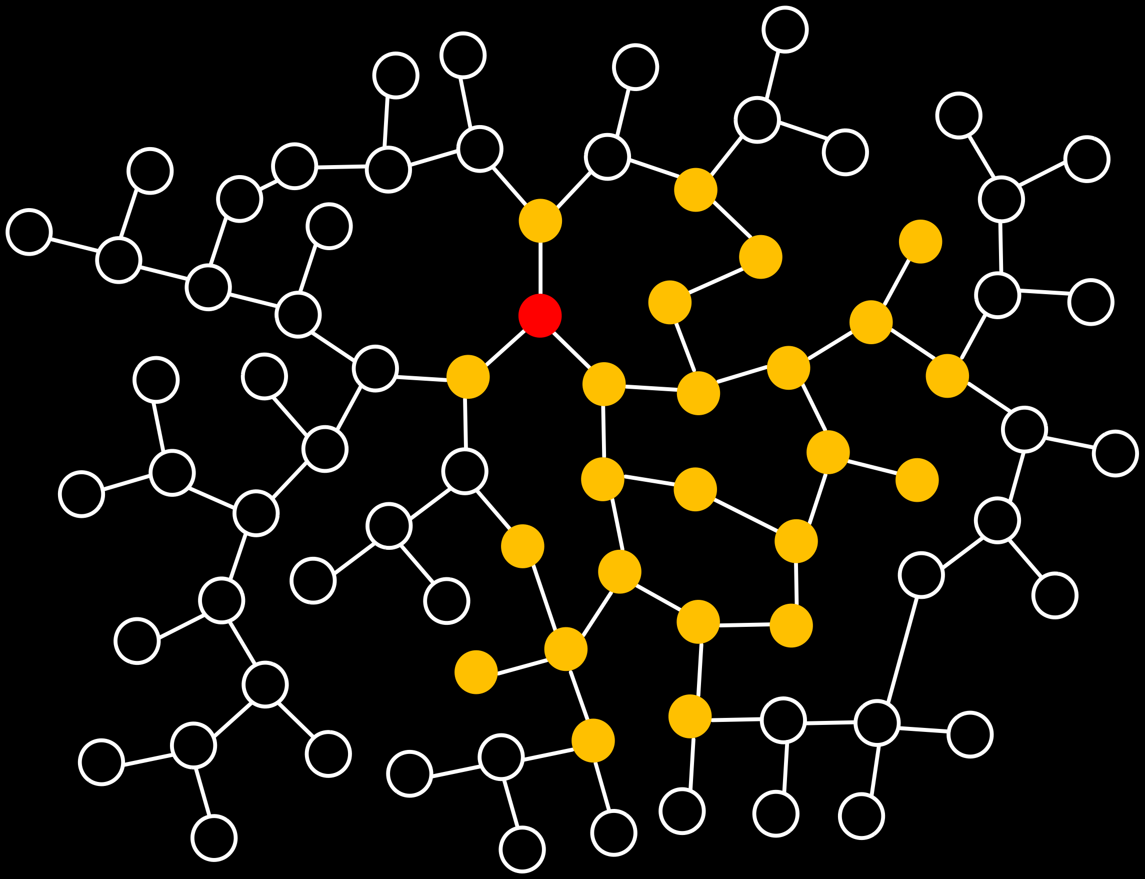
Main result: adaptive diffusion



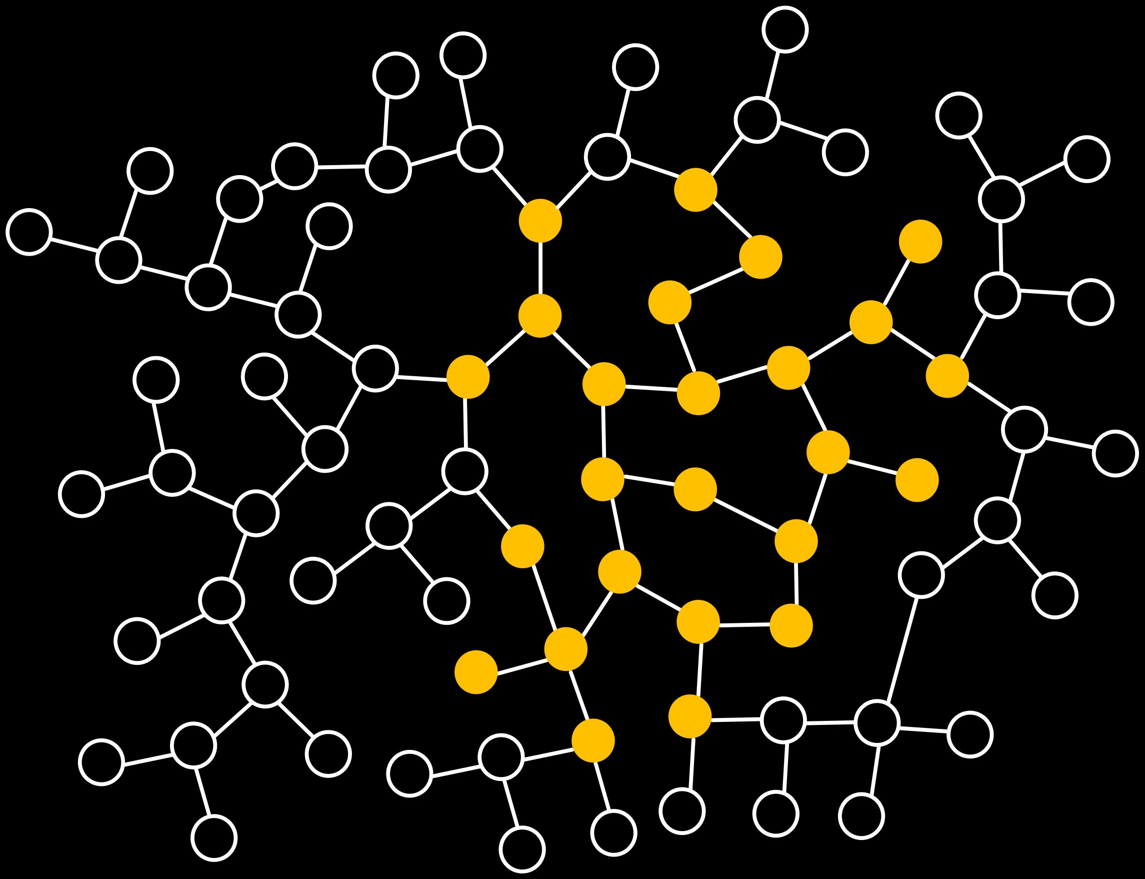
Main result: adaptive diffusion



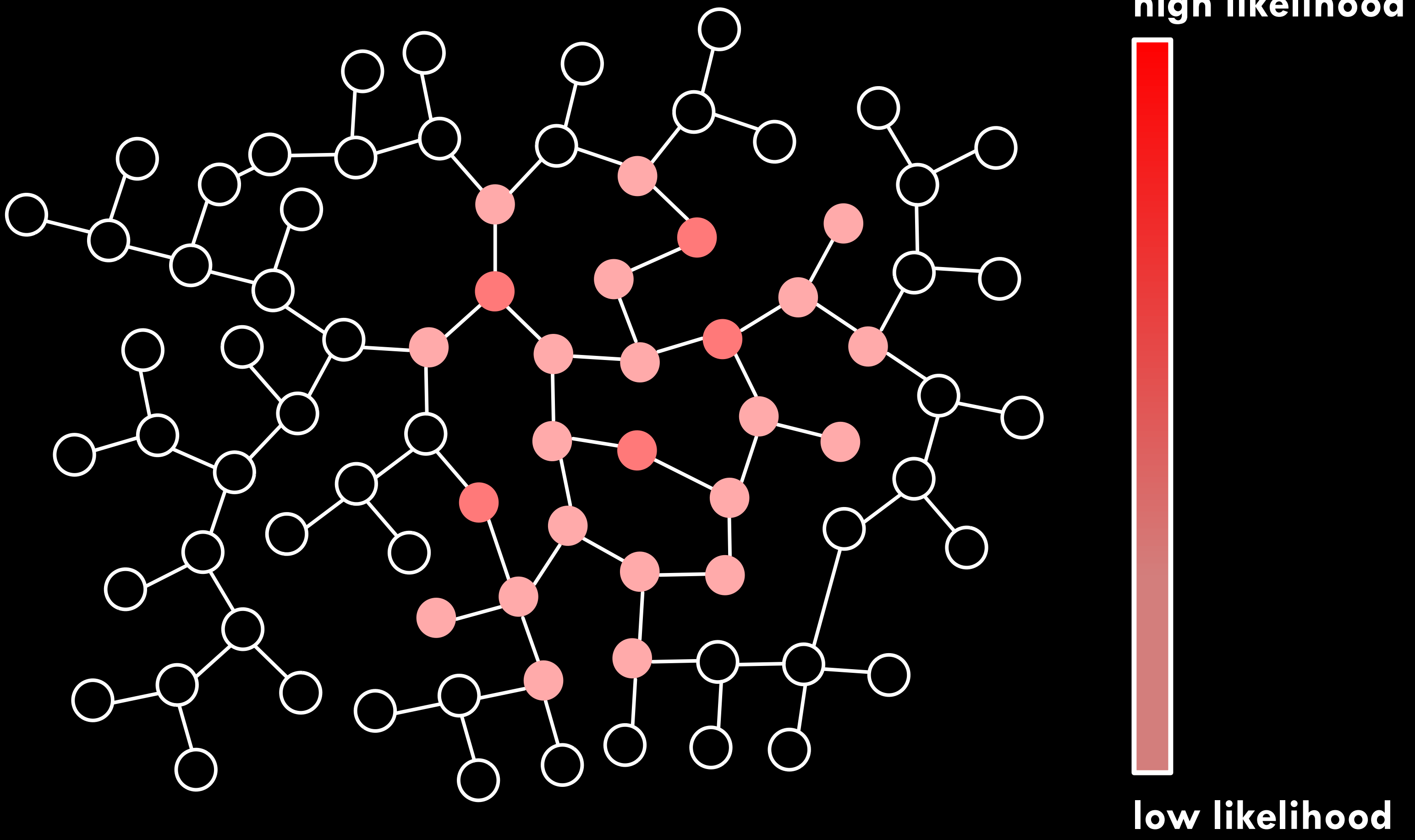
Main result: adaptive diffusion



Main result: adaptive diffusion

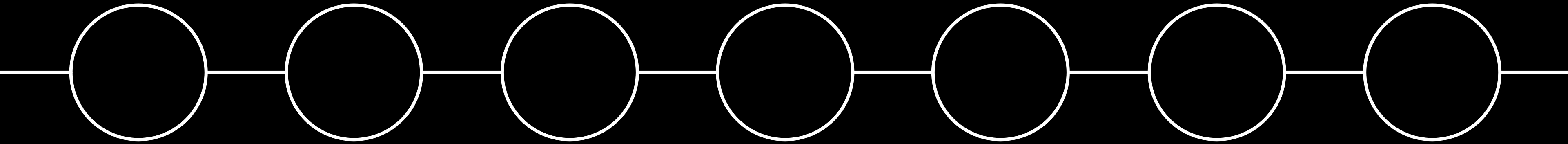


Main result: adaptive diffusion



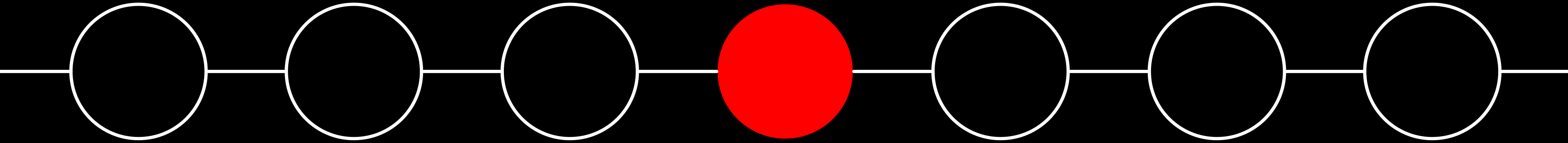
provides provable anonymity guarantees!

Line graphs



- let’s start with line graphs

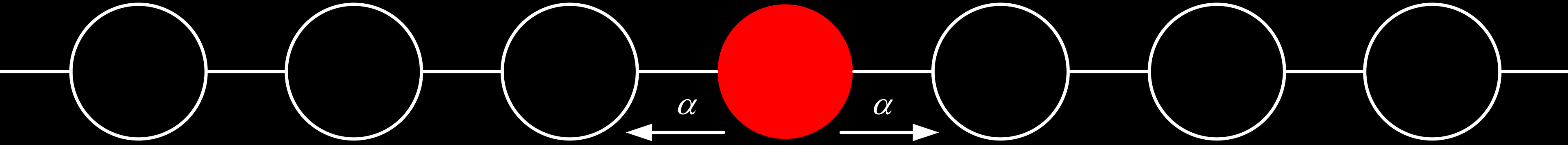
Line graphs: diffusion



$T = 0$

- the message author starts a rumor at $T = 0$

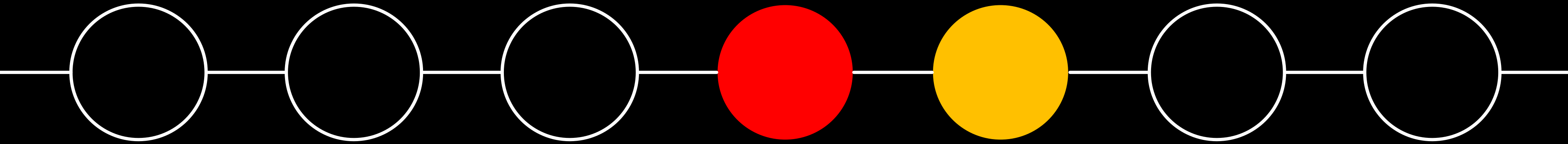
Line graphs: diffusion



$T = 1$

- with probability α, the left (right) node receives the message

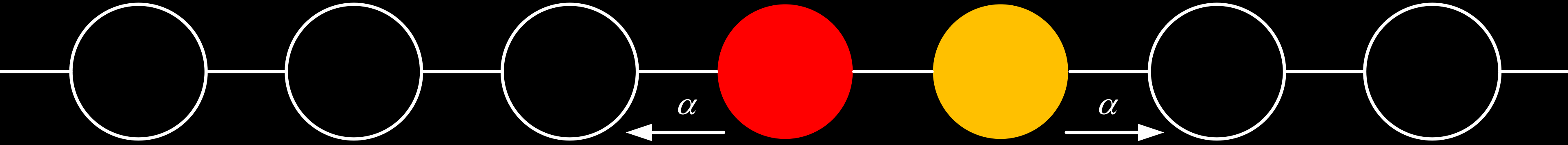
Line graphs: diffusion



$T = 1$

- the node to the right of the author receives the message

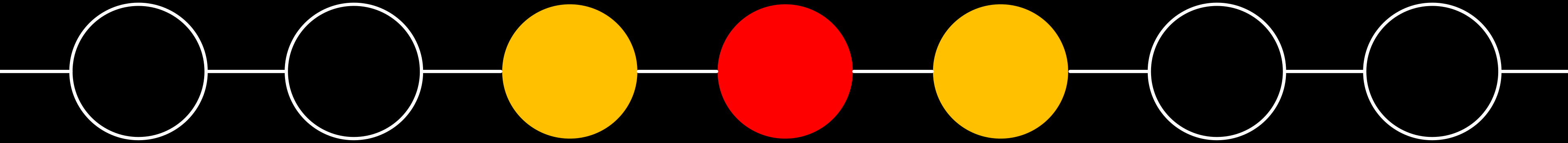
Line graphs: diffusion



$T = 2$

- the rumor propagates in **both directions** at the **same rate**

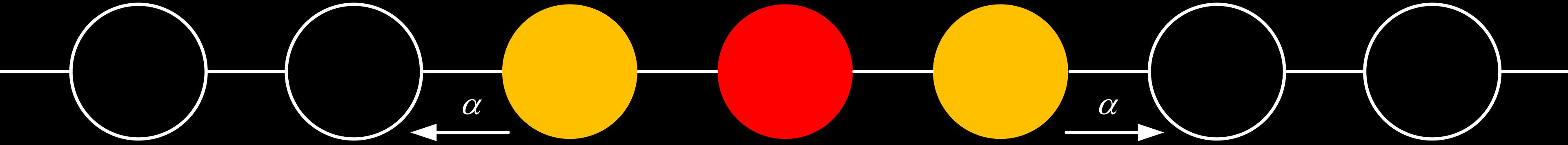
Line graphs: diffusion



$T = 2$

- the rumor propagates in **both directions** at the **same rate**

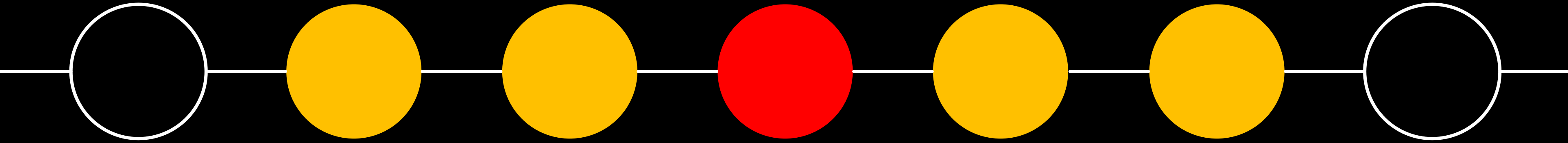
Line graphs: diffusion



$T = 3$

- α is **independent of time or hop distance** to message author

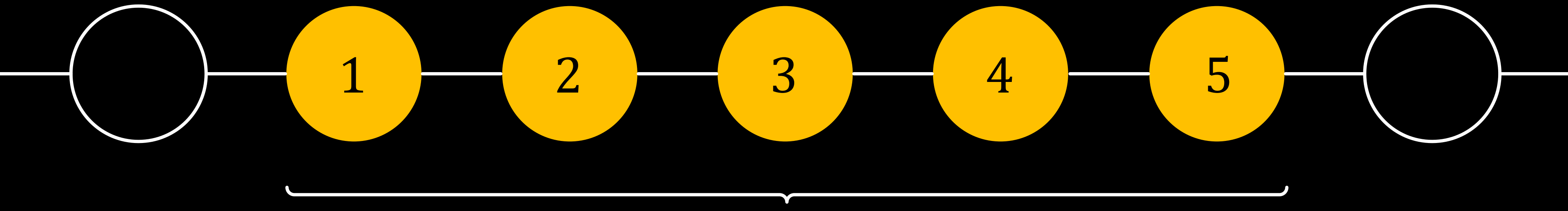
Line graphs: diffusion



$T = 3$

- diffusion on a line is equivalent to **two independent random walks**

Adversary's observation

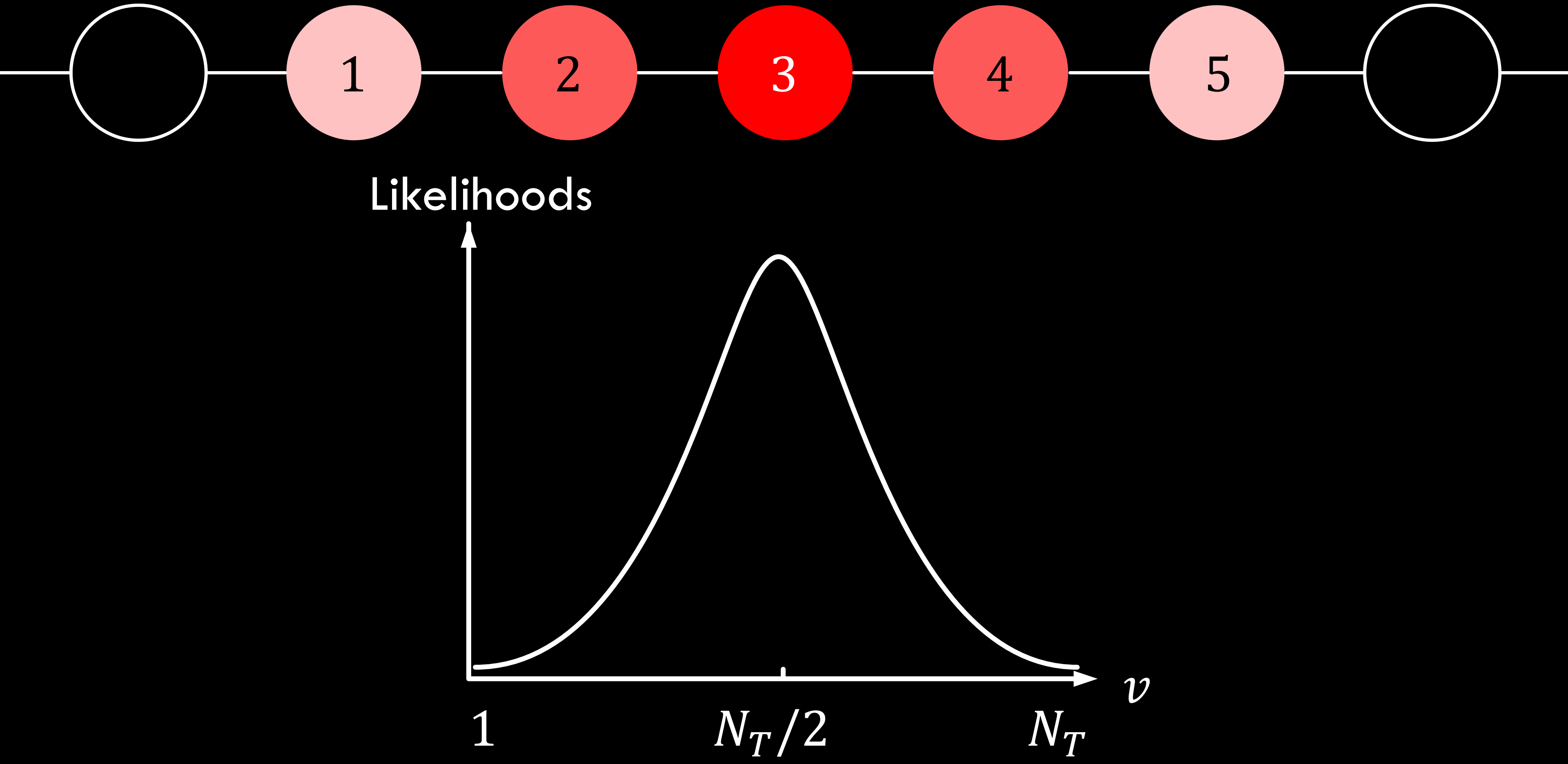


$N = 5$

nodes with the message

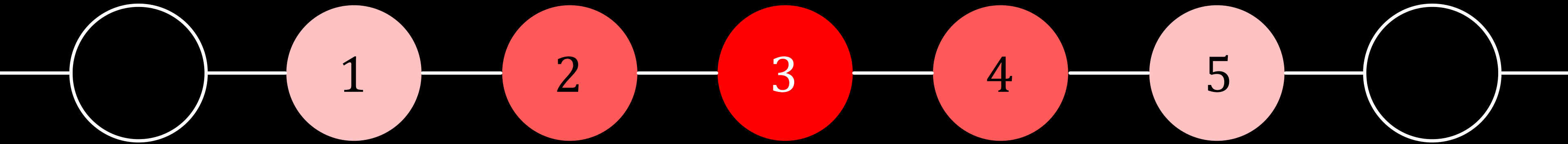
can the adversary locate the message author?

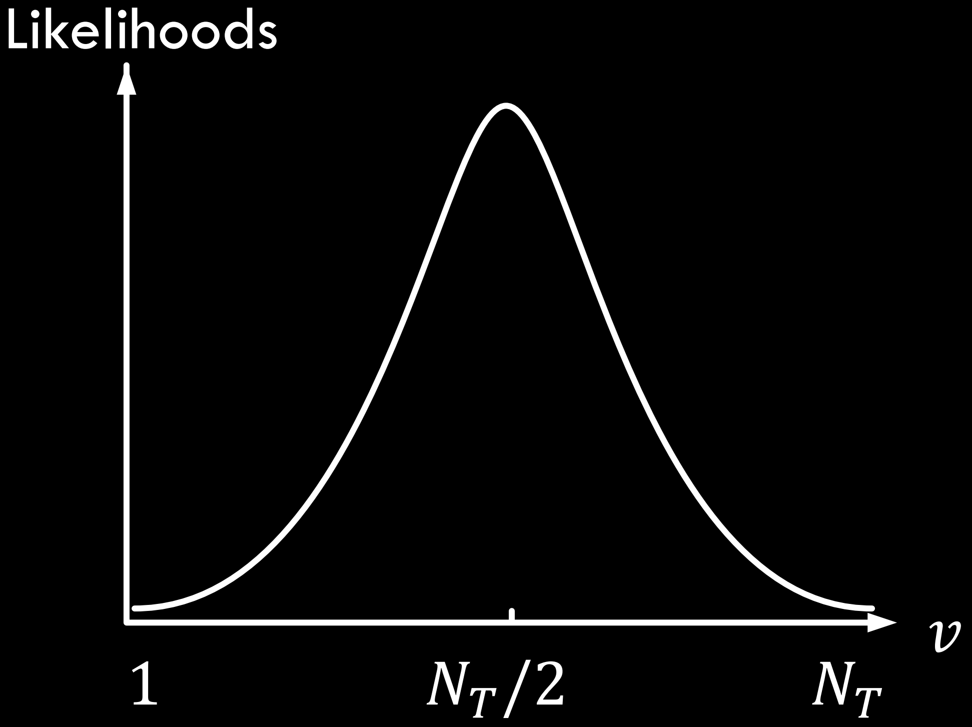
Maximum likelihood detection



- the **node in the middle** is the **mostly likely author**

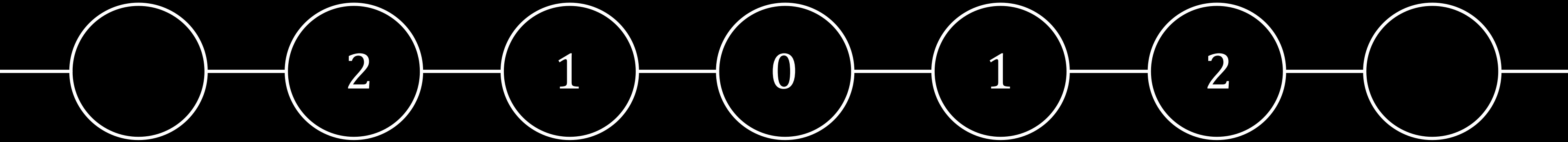
Maximum likelihood detection





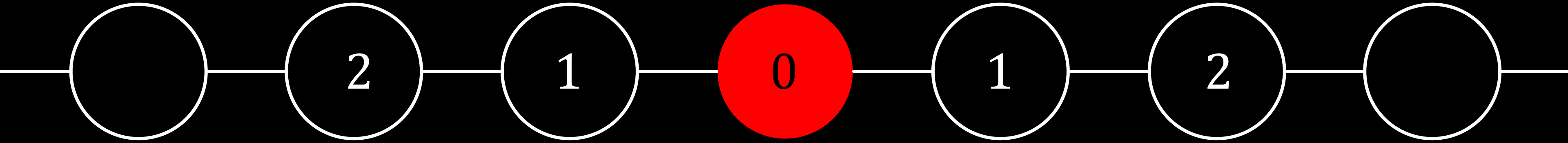
Probability of detection $\approx \frac{1}{\sqrt{N_T}}$

Line graphs: adaptive diffusion



- consider a line graph

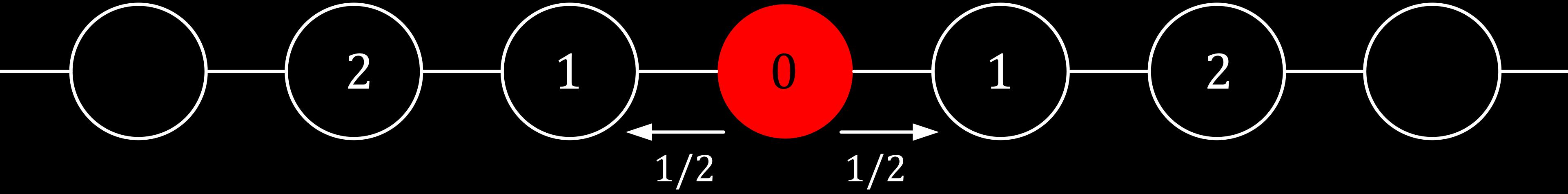
Line graphs: adaptive diffusion



$T = 0$

- node 0 starts a rumor at $T = 0$

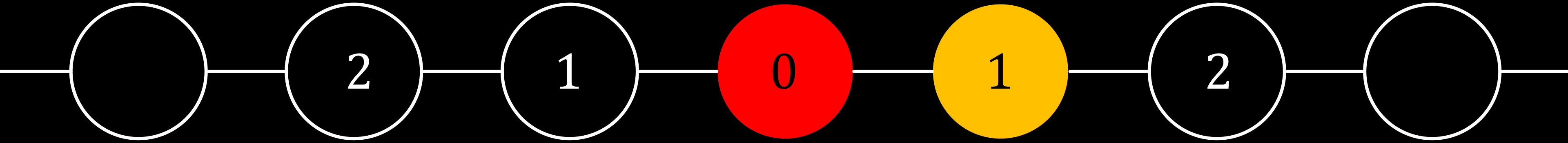
Line graphs: adaptive diffusion



$T = 1$

- with probability $1/2$, the left (right) node receives the message

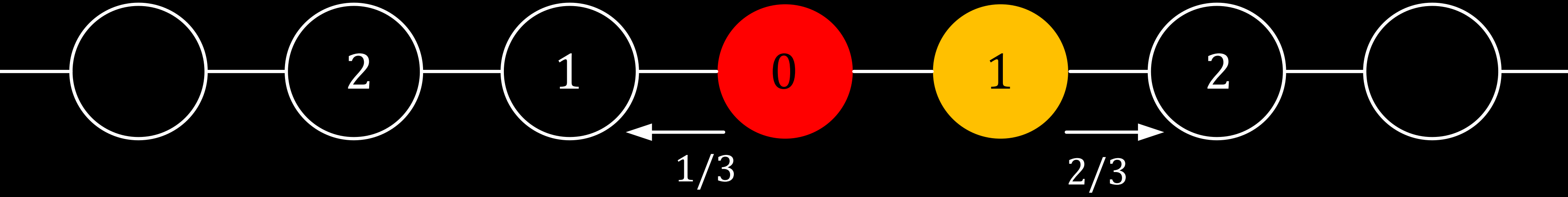
Line graphs: adaptive diffusion



$T = 1$

- right node 1 receives the message

Line graphs: adaptive diffusion



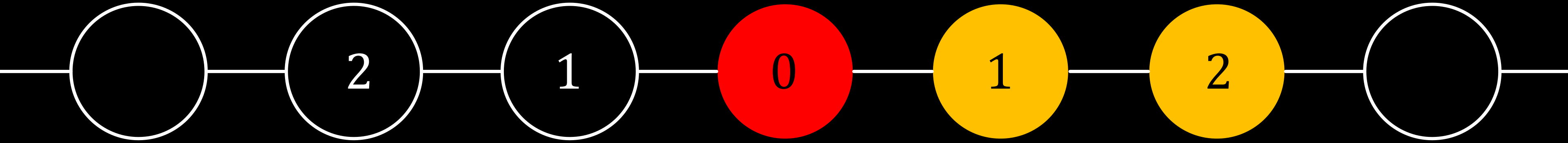
$T = 2$

- probability of passing message: $\alpha = \frac{h+1}{T+1}$

hop distance to message author

elapsed time

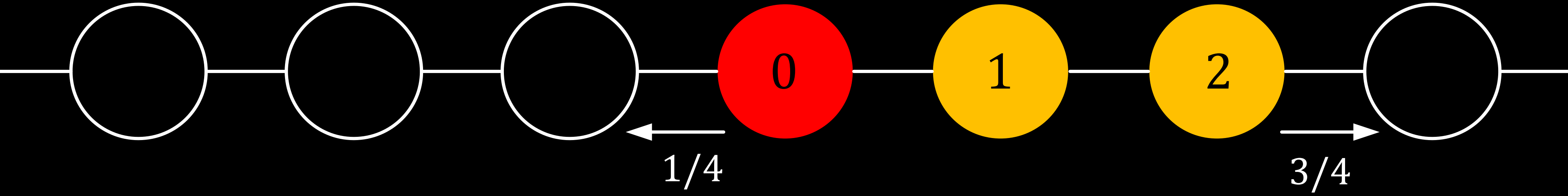
Line graphs: adaptive diffusion



$T = 2$

- right node 2 receives the message

Line graphs: adaptive diffusion



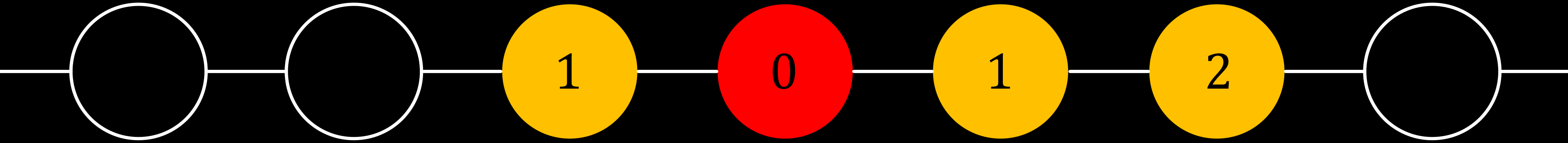
$T = 3$

hop distance to message author

- probability of passing message: $\alpha = \frac{h+1}{T+1}$

elapsed time

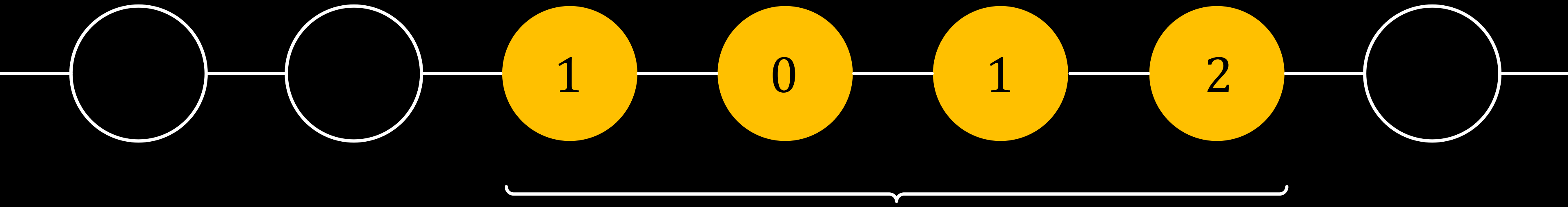
Line graphs: adaptive diffusion



$T = 3$

- left node 1 receives the message

Adversary's observation

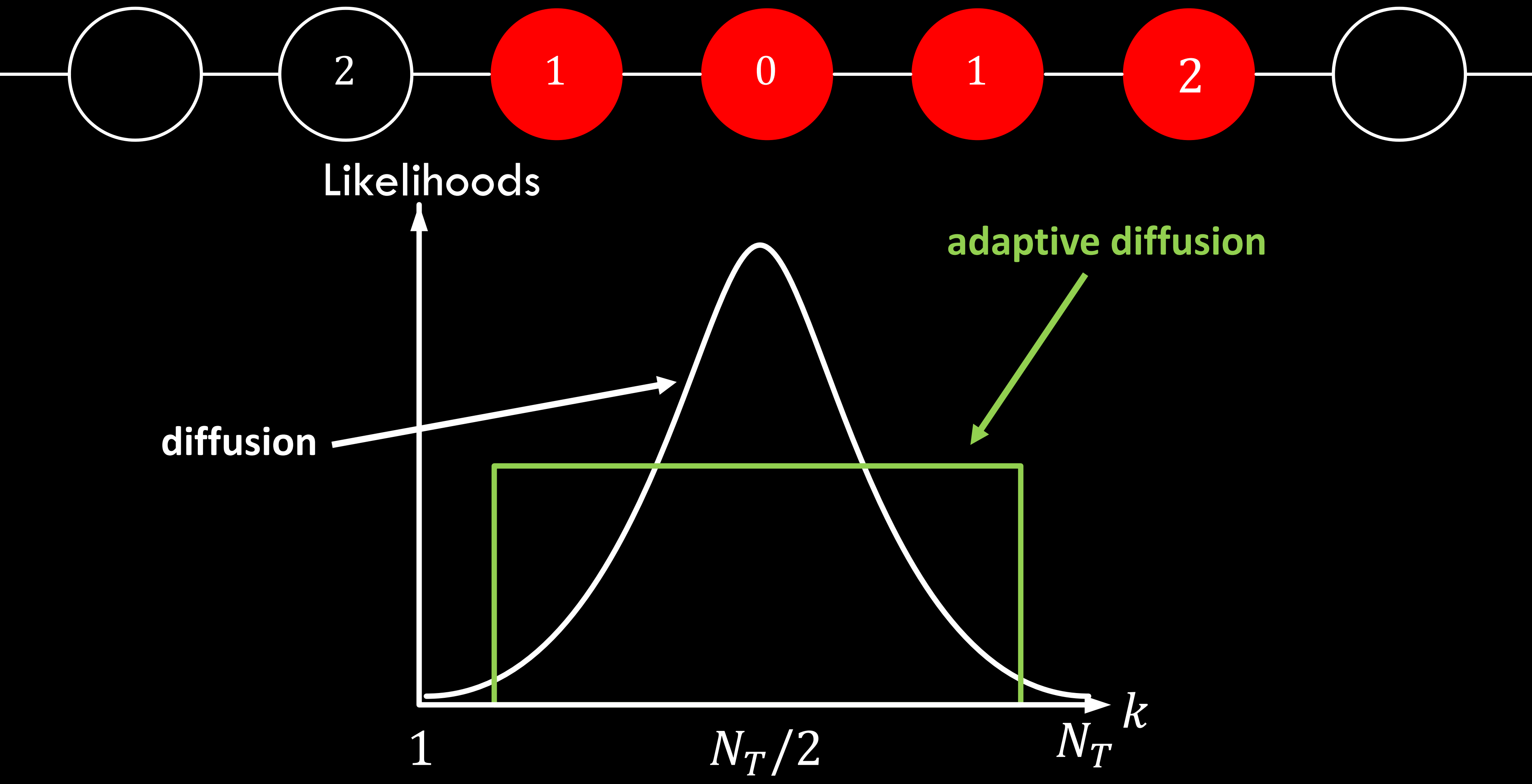


$N_T = 4$

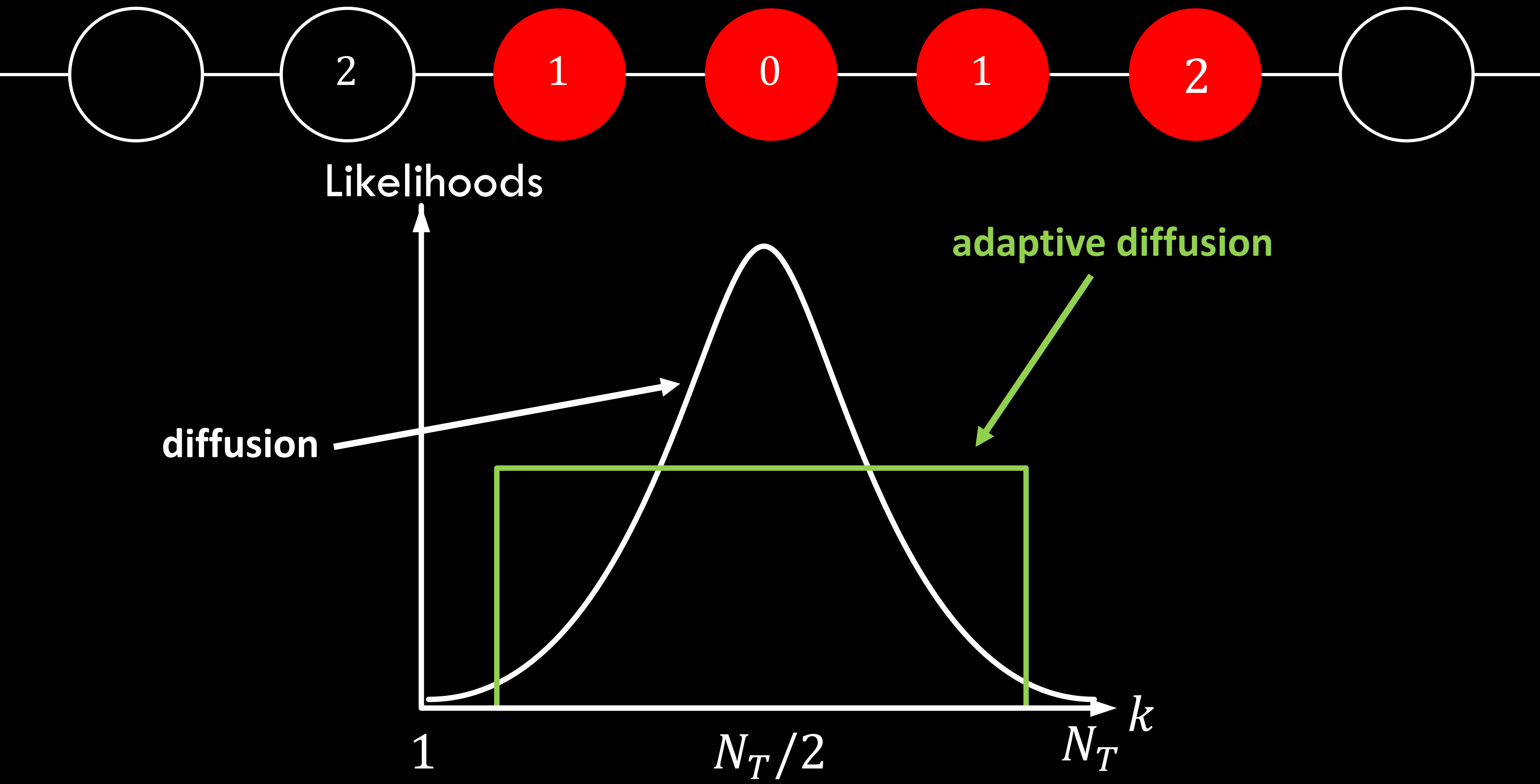
nodes with the message

can the adversary locate the message author?

Maximum likelihood detection

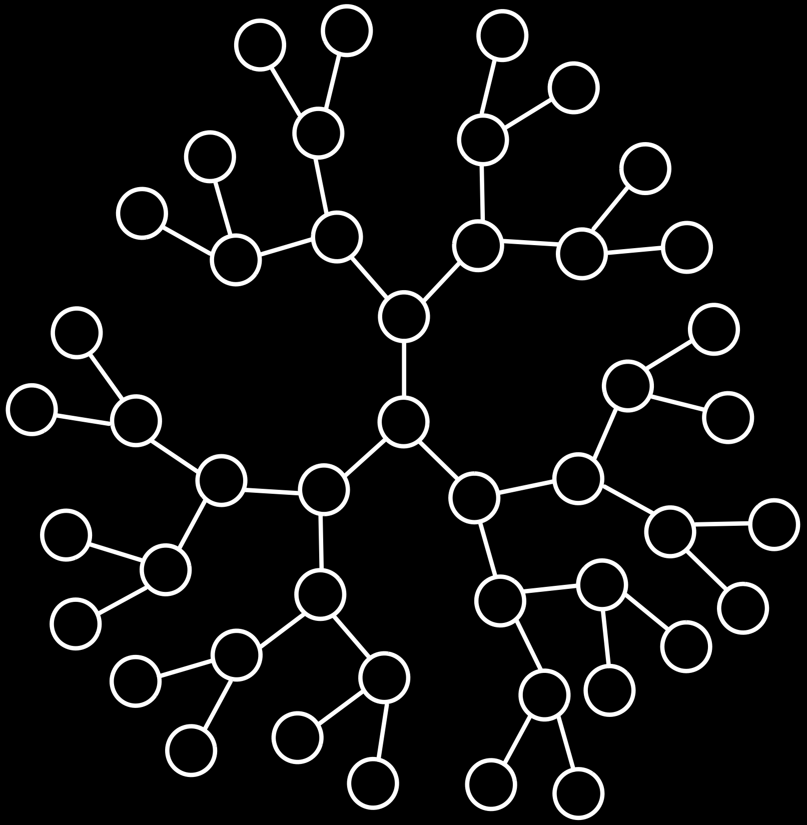


Maximum likelihood detection



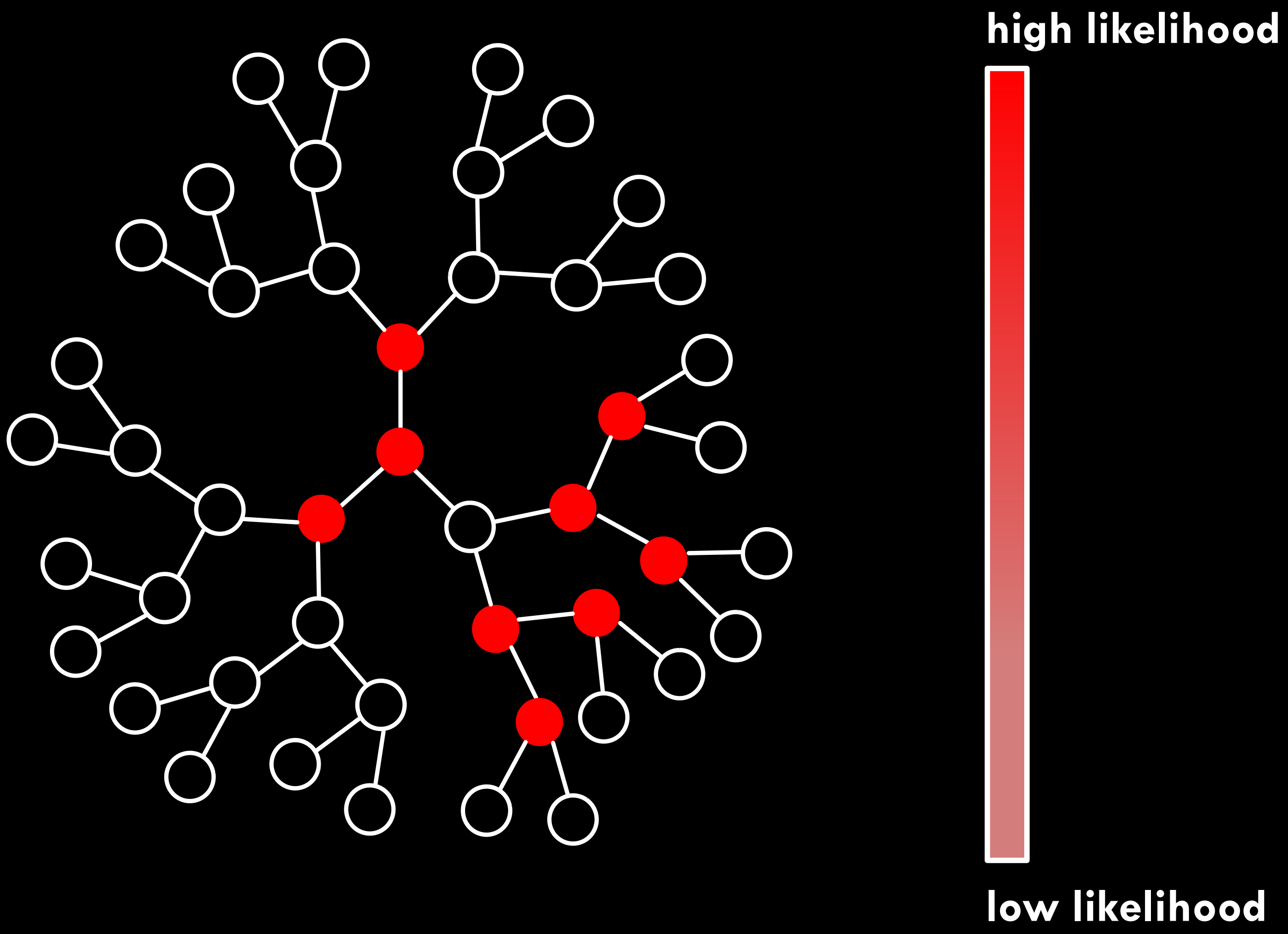
Probability of detection $\approx \frac{1}{N_T}$

d-regular trees



adaptive diffusion for regular trees?

Maximum likelihood detection



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

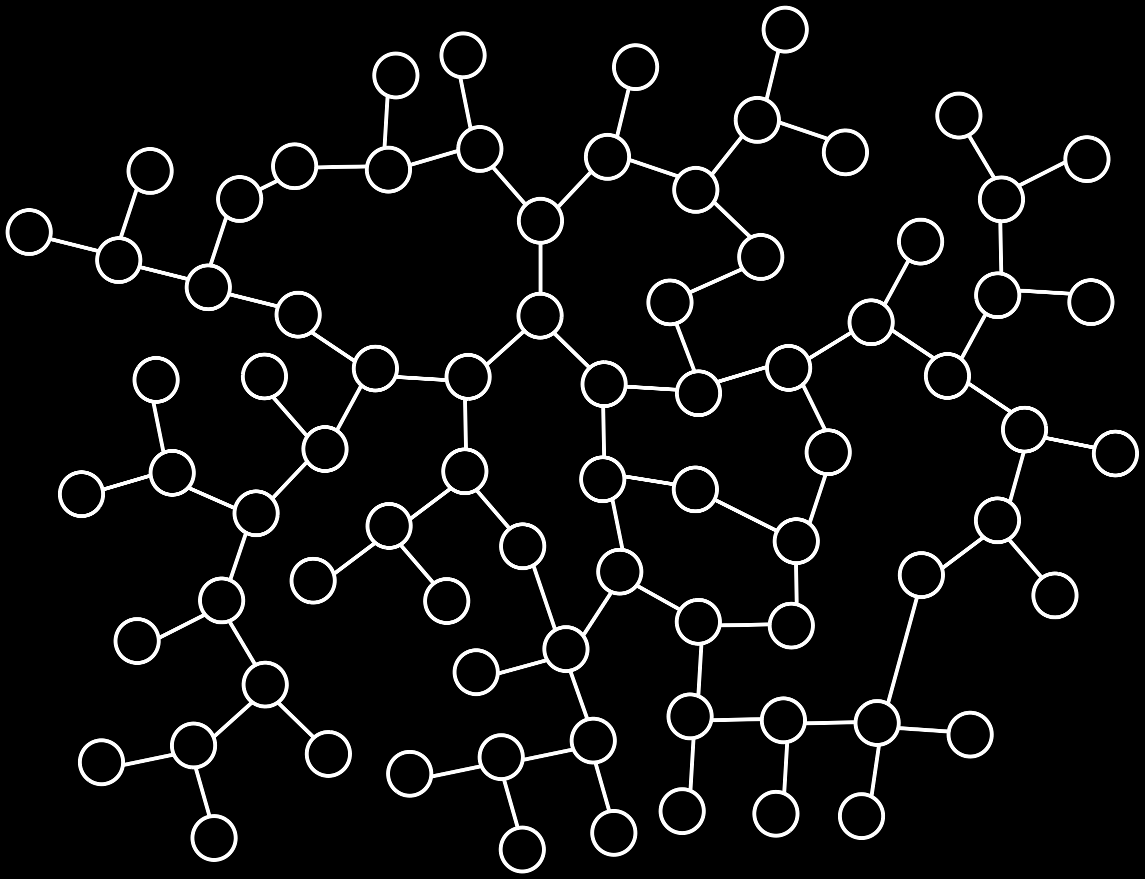
Main Theorem

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*

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

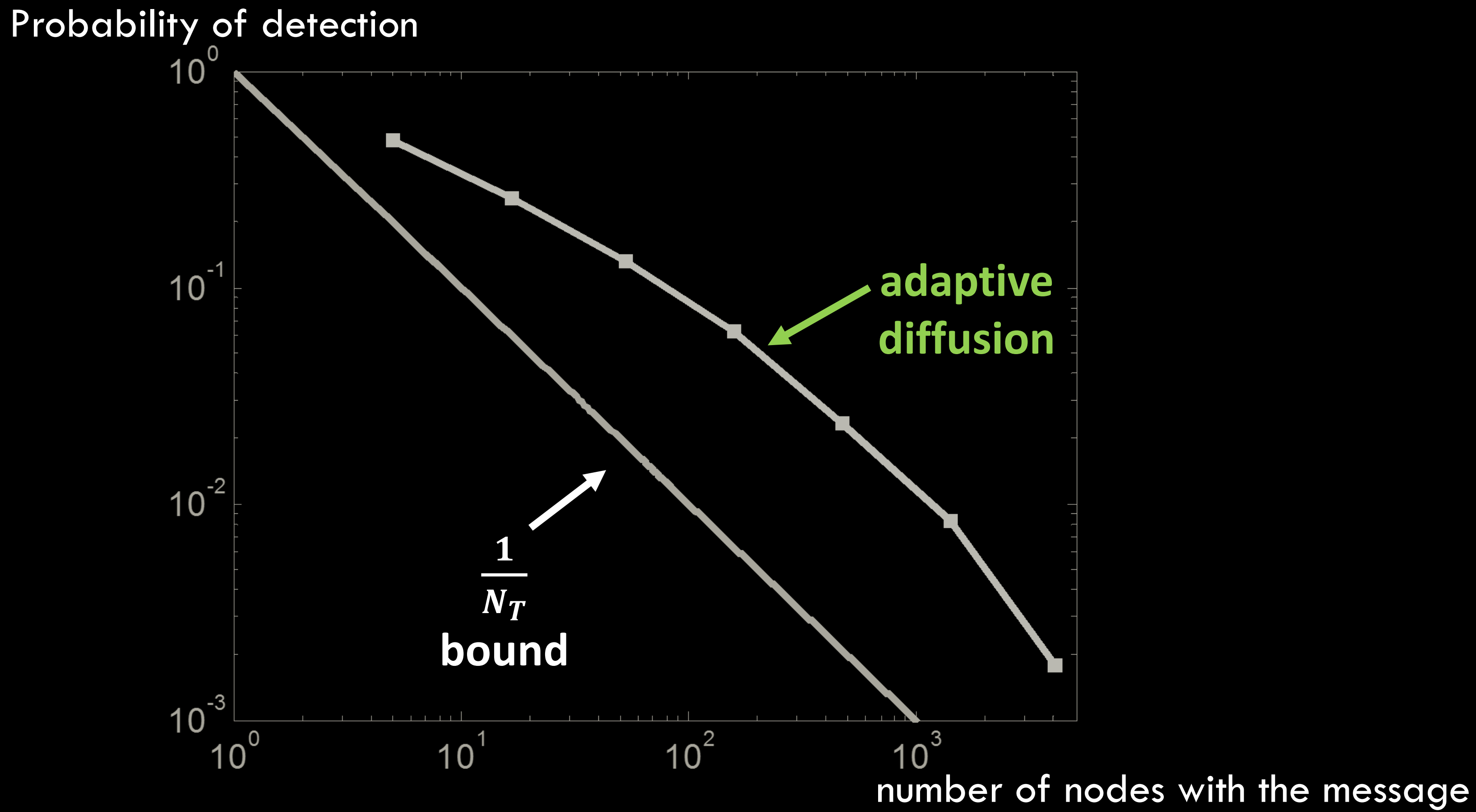
3. *The expected distance between the estimated and true source is at least* $\frac{T}{2}$.

General graphs



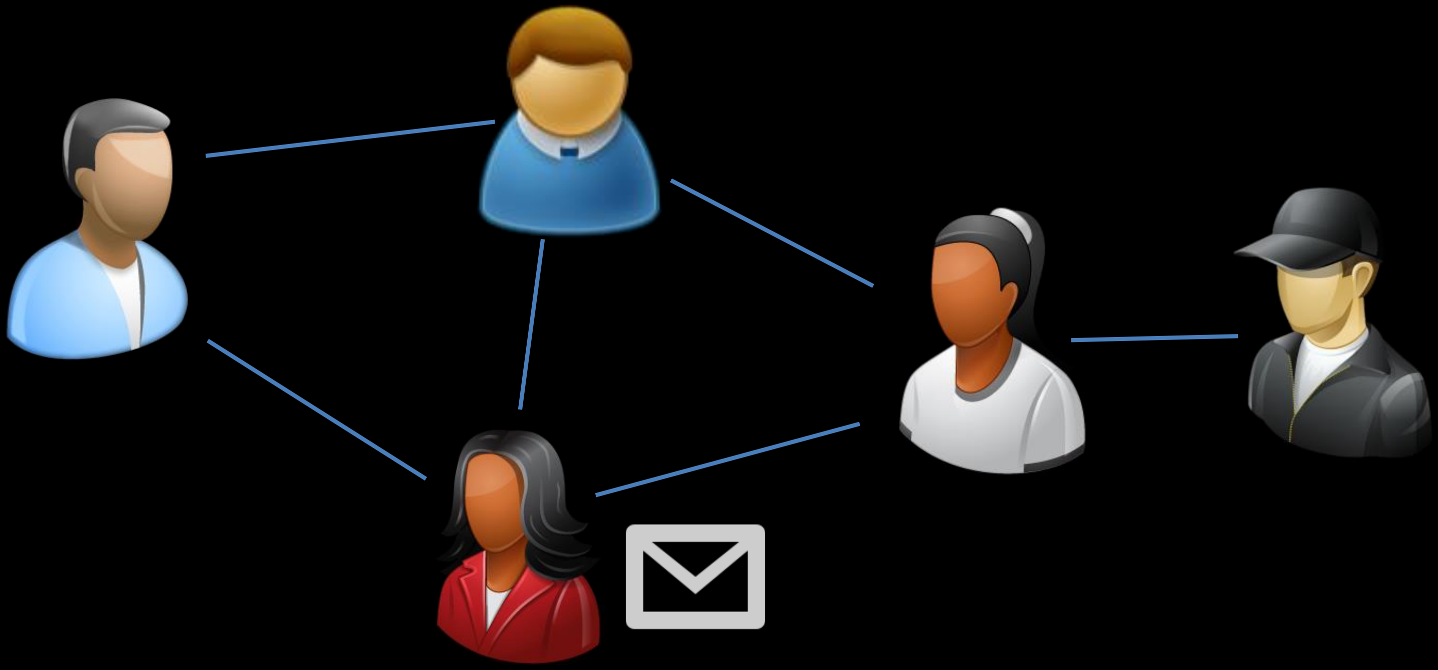
adaptive diffusion for general graphs?

Simulation results: Facebook graph



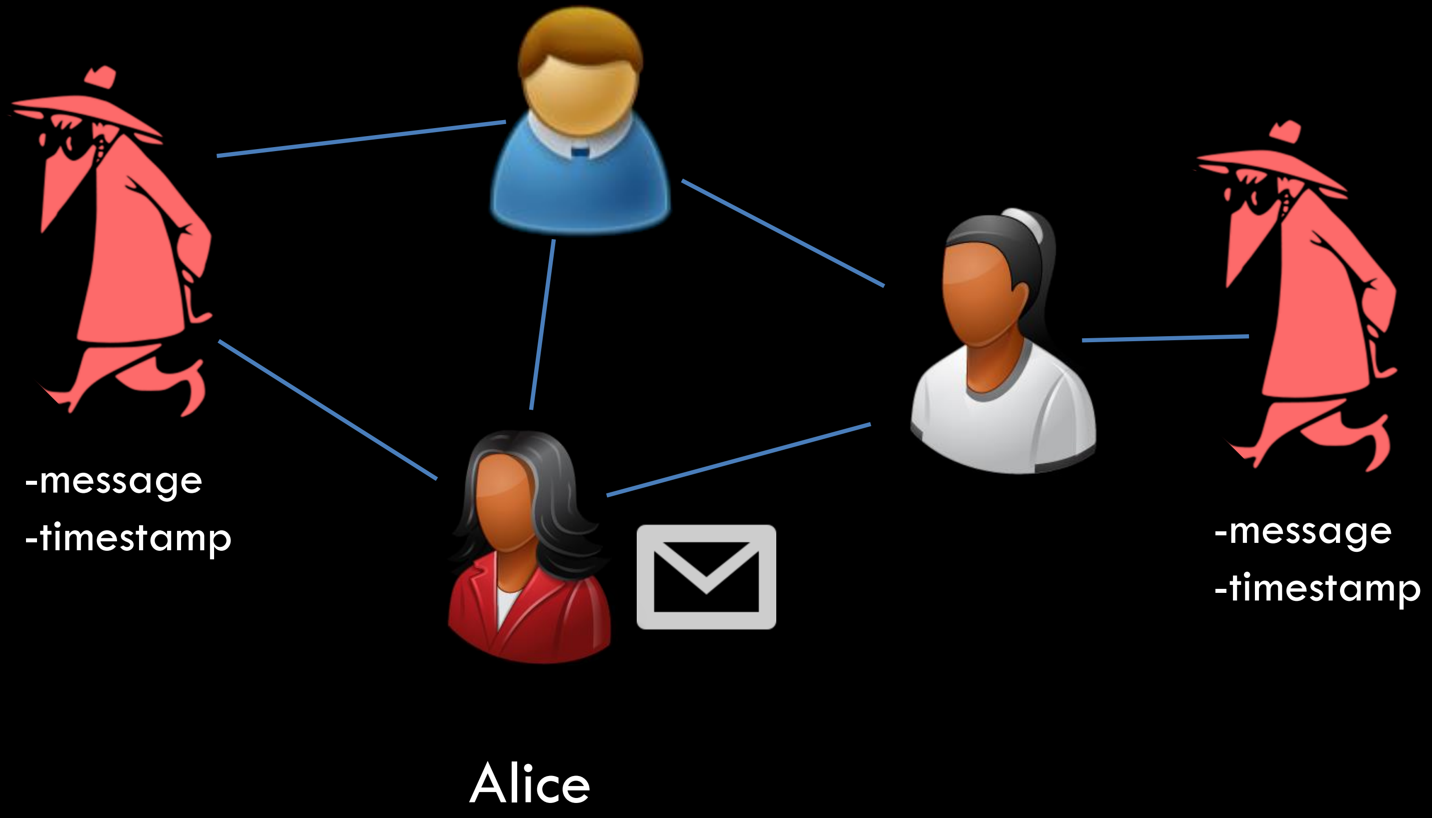
- likelihoods can be **approximated** numerically

Adversary with timing



Alice

Adversary with timing

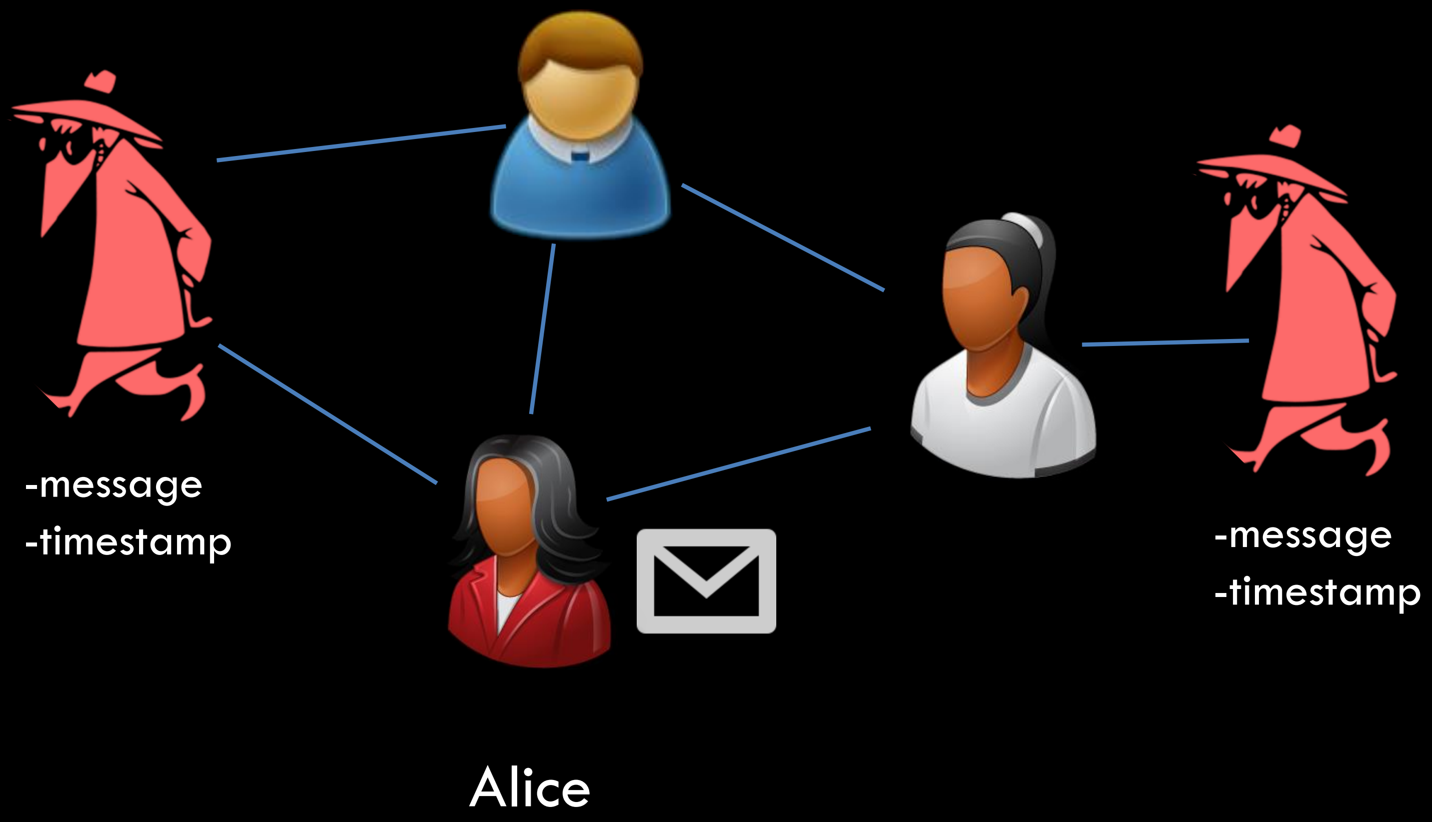


-message
-timestamp

-message
-timestamp

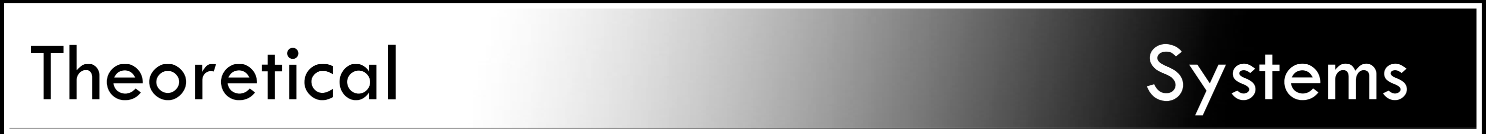
Alice

Adversary with timing


-message
-timestamp
-message
-timestamp
Alice

adaptive diffusion is order "optimal" for trees!

Extensions and related work



Theoretical

Systems

- Adversaries with timing information
- Peer-to-peer dynamic networks
- Hiding relays
- Multiple message sources

- Cyber-bullying detection
- Anonymous video sharing
- Message caching
- Bootstrapping contacts