Anonymity in the Bitcoin Peer-to-Peer Network

Shaileshh Bojja Venkatakrishnan, Giulia Fanti,
Andrew Miller, Pramod Viswanath

Why do People Use Cryptocurrencies?

Currency Stability

Investment

Technical Properties/ Ideology

“Untraceable Bitcoin”

Teenagers using untraceable currency Bitcoin to buy dangerous drugs online

Fears have been raised as children as young as 14 are getting parcels of legal highs delivered to their home

Mirror

This is false.

Bitcoin Reminder

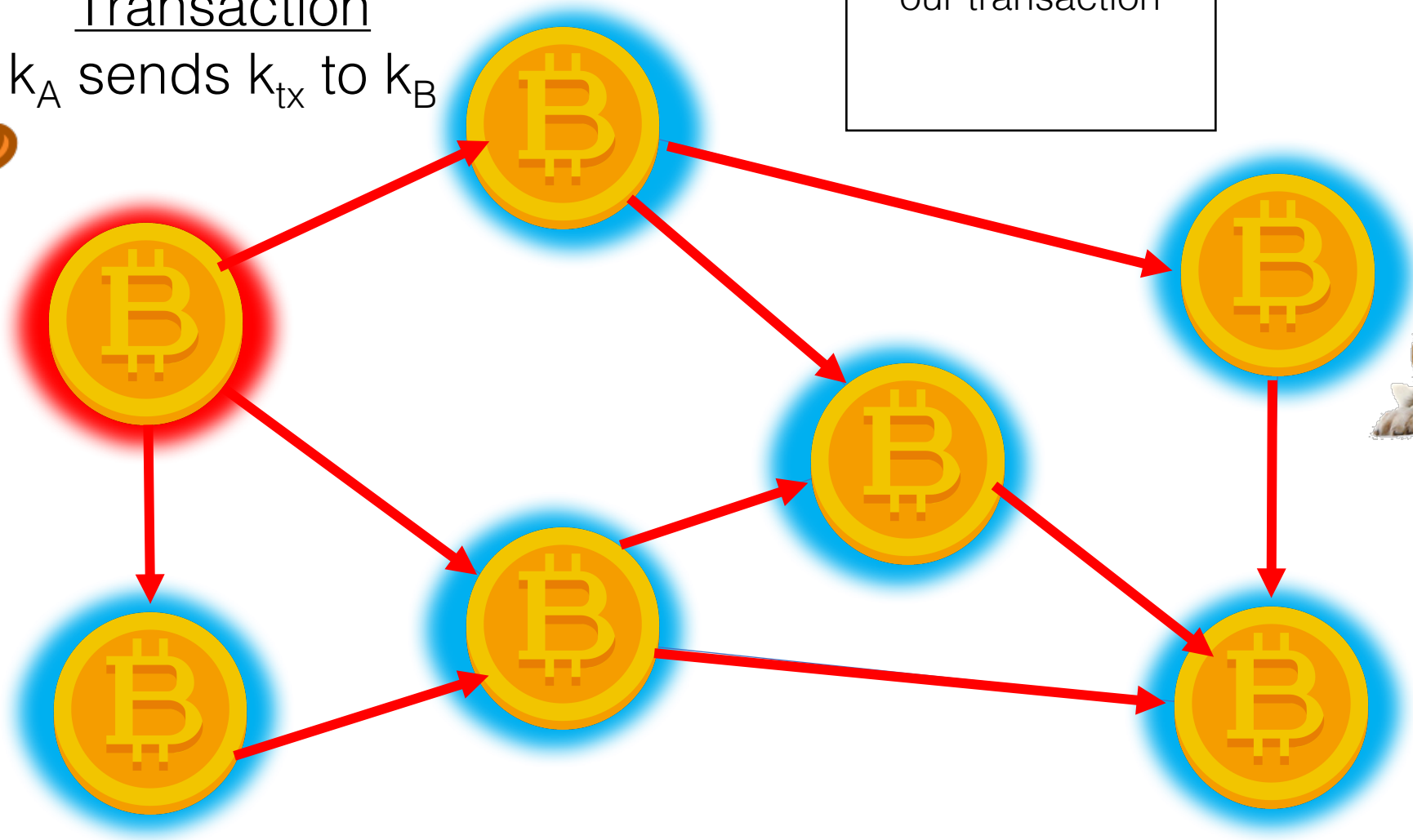
Transaction

k_A sends k_{tx} to k_B

Blockchain
sd93fjj2
pckrn29
…
our transaction

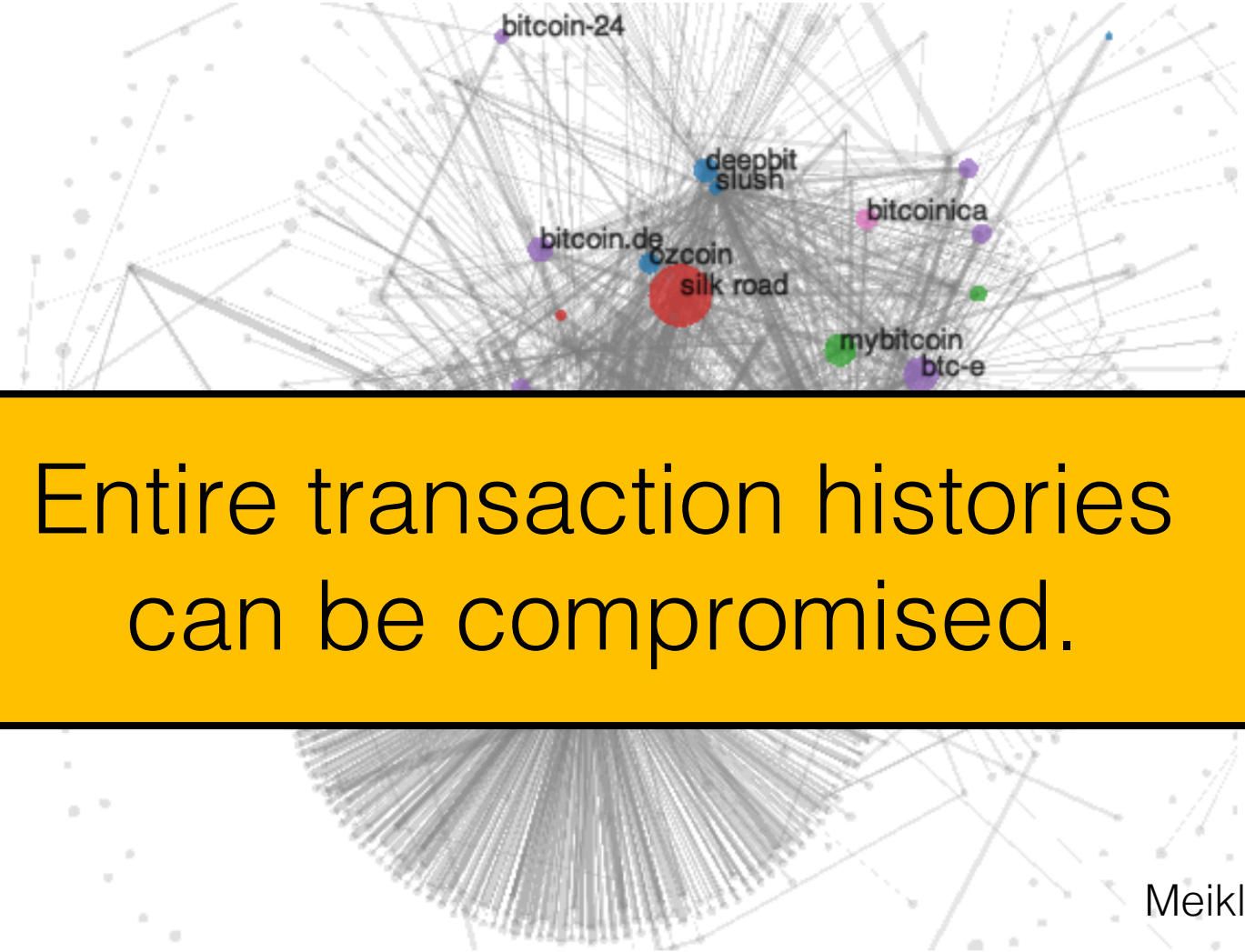
Alice
k_A

k_{tx}



Bob
k_B

How can users be deanonymized?



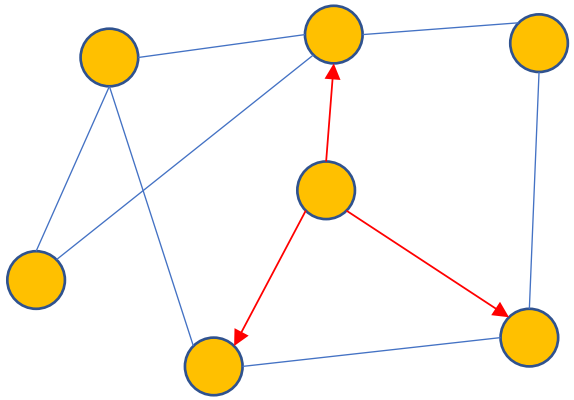
Entire transaction histories
can be compromised.

Meiklejohn et al., 2013

What about the peer-to-peer network?

Our Work

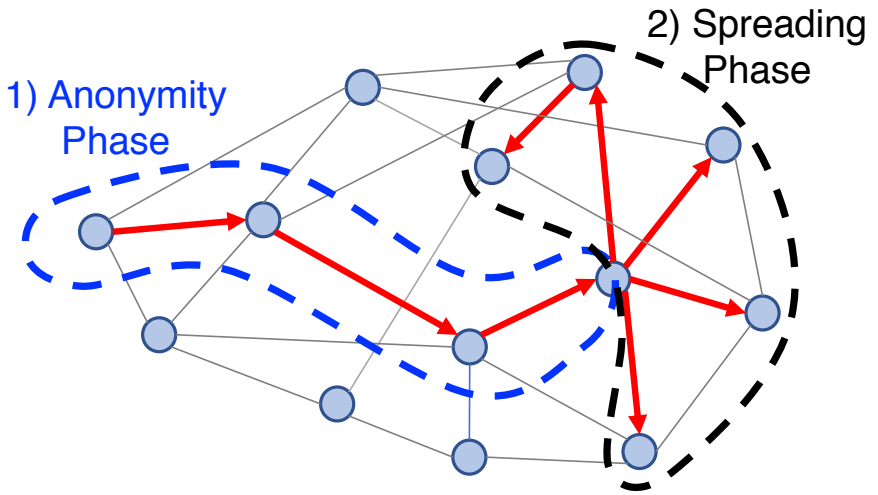
Analysis



Pr(detection)

Under submission, 2017

Redesign

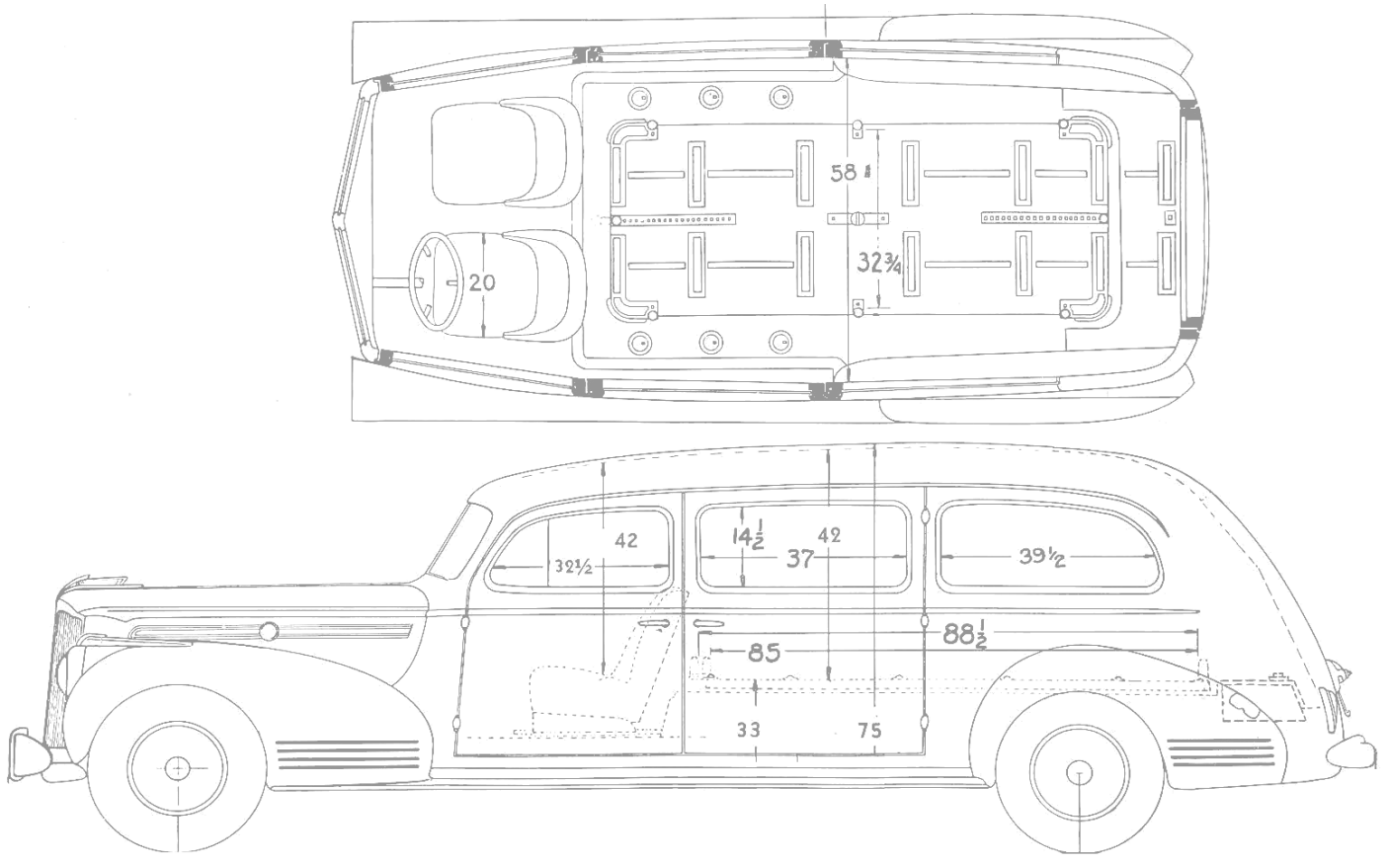


Dandelion

ACM Sigmetrics 2017

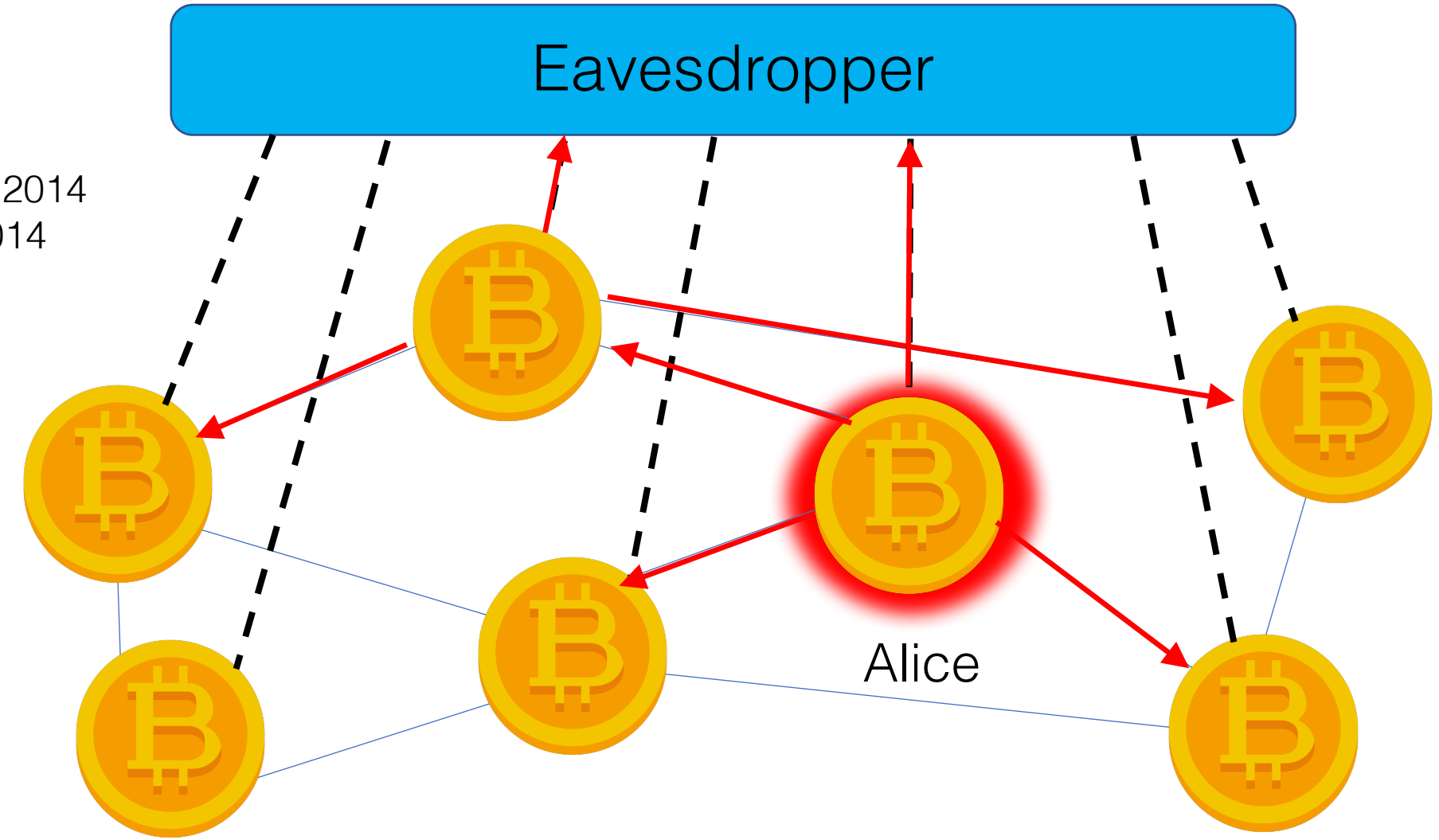
Model

Assumptions and Notation

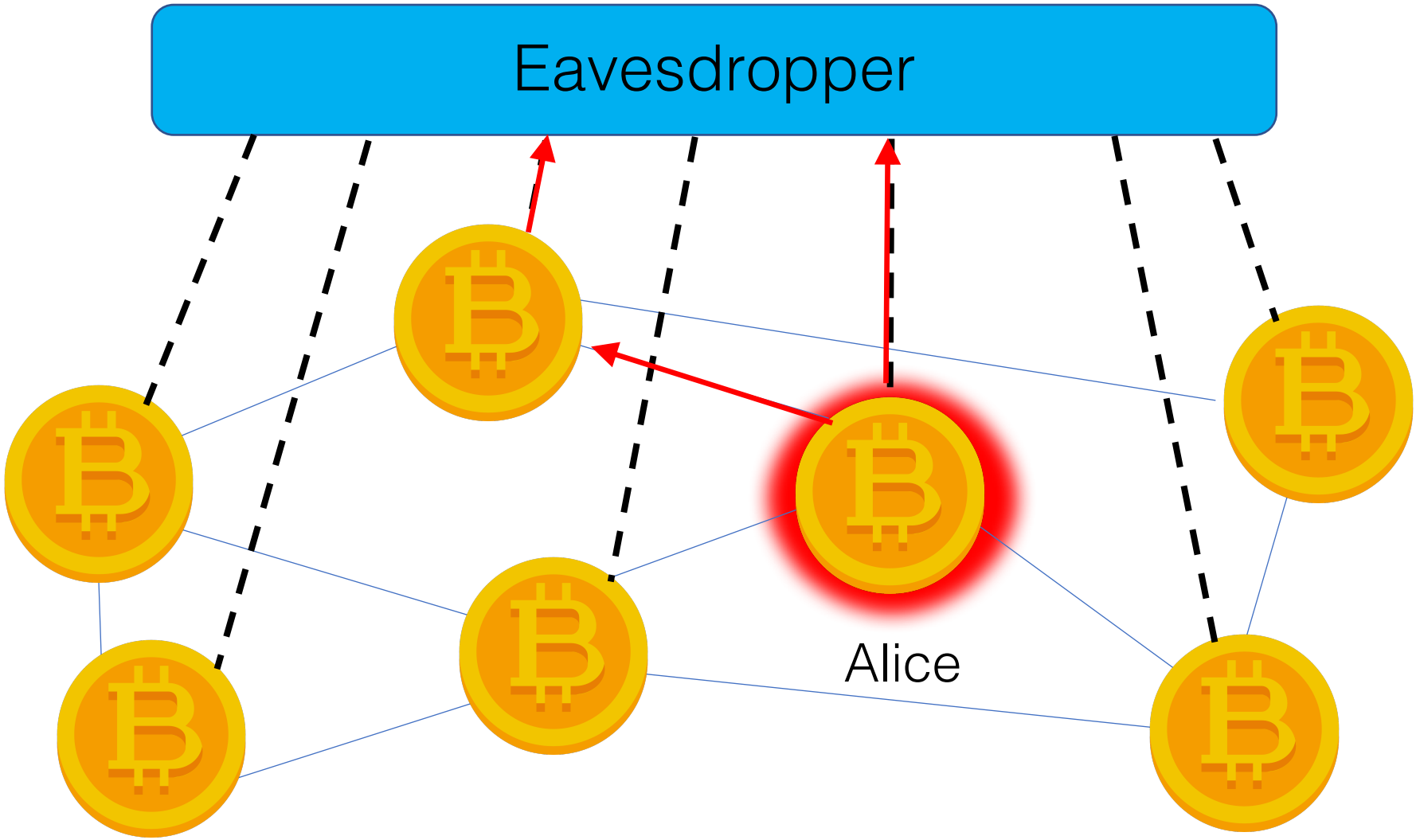


Attacks on the Network Layer

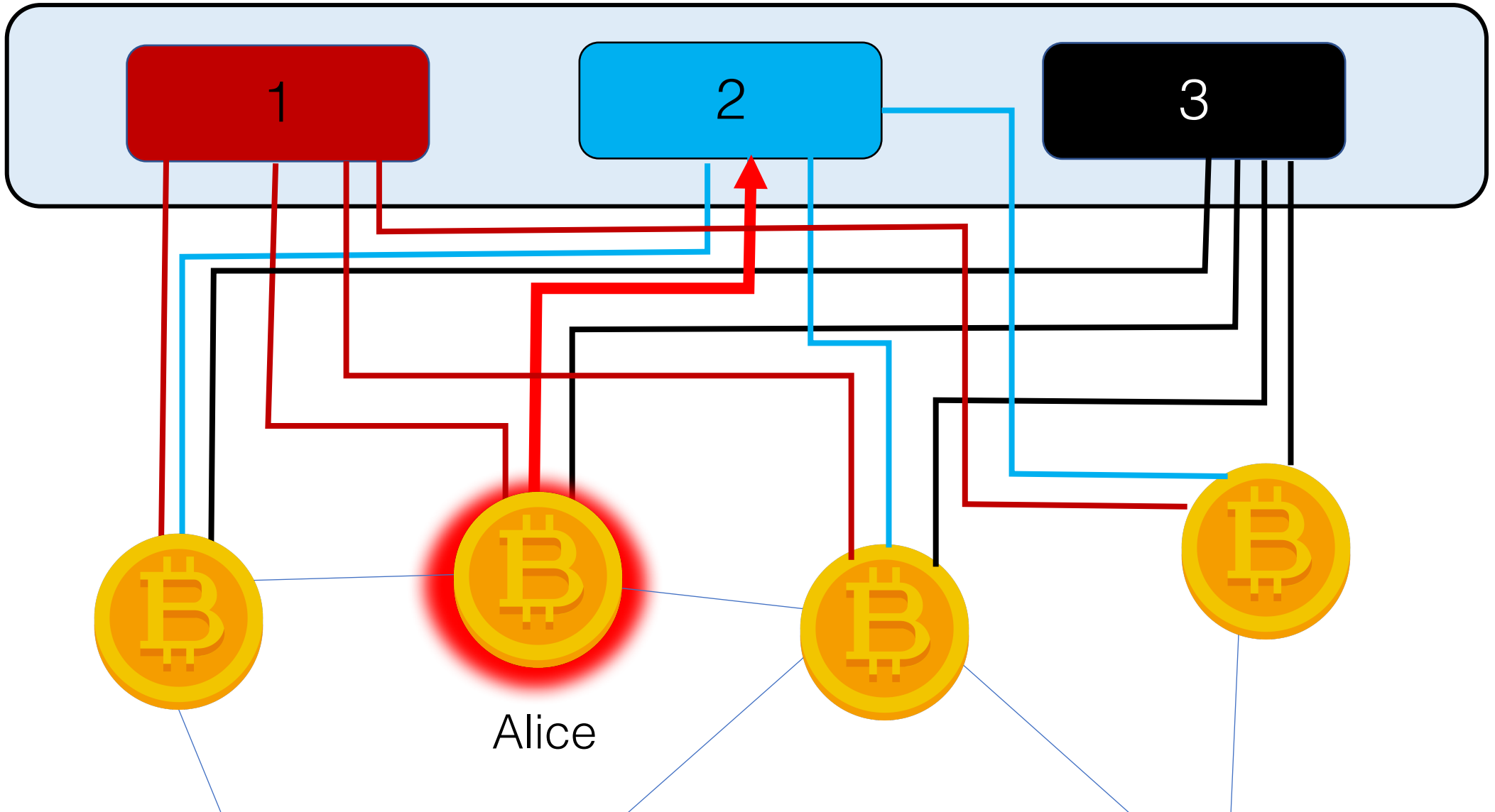
Biryukov et al., 2014
Koshy et al., 2014



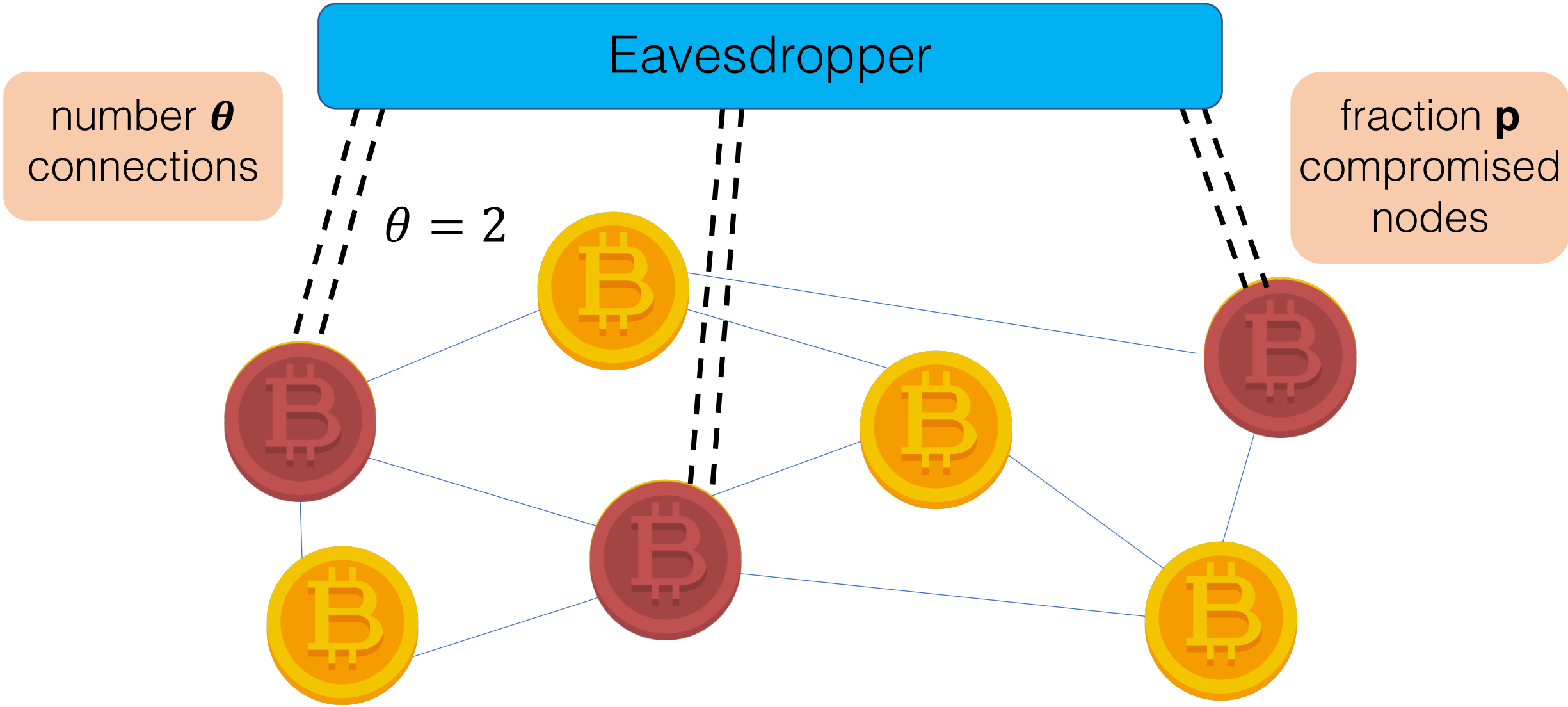
What can go wrong?

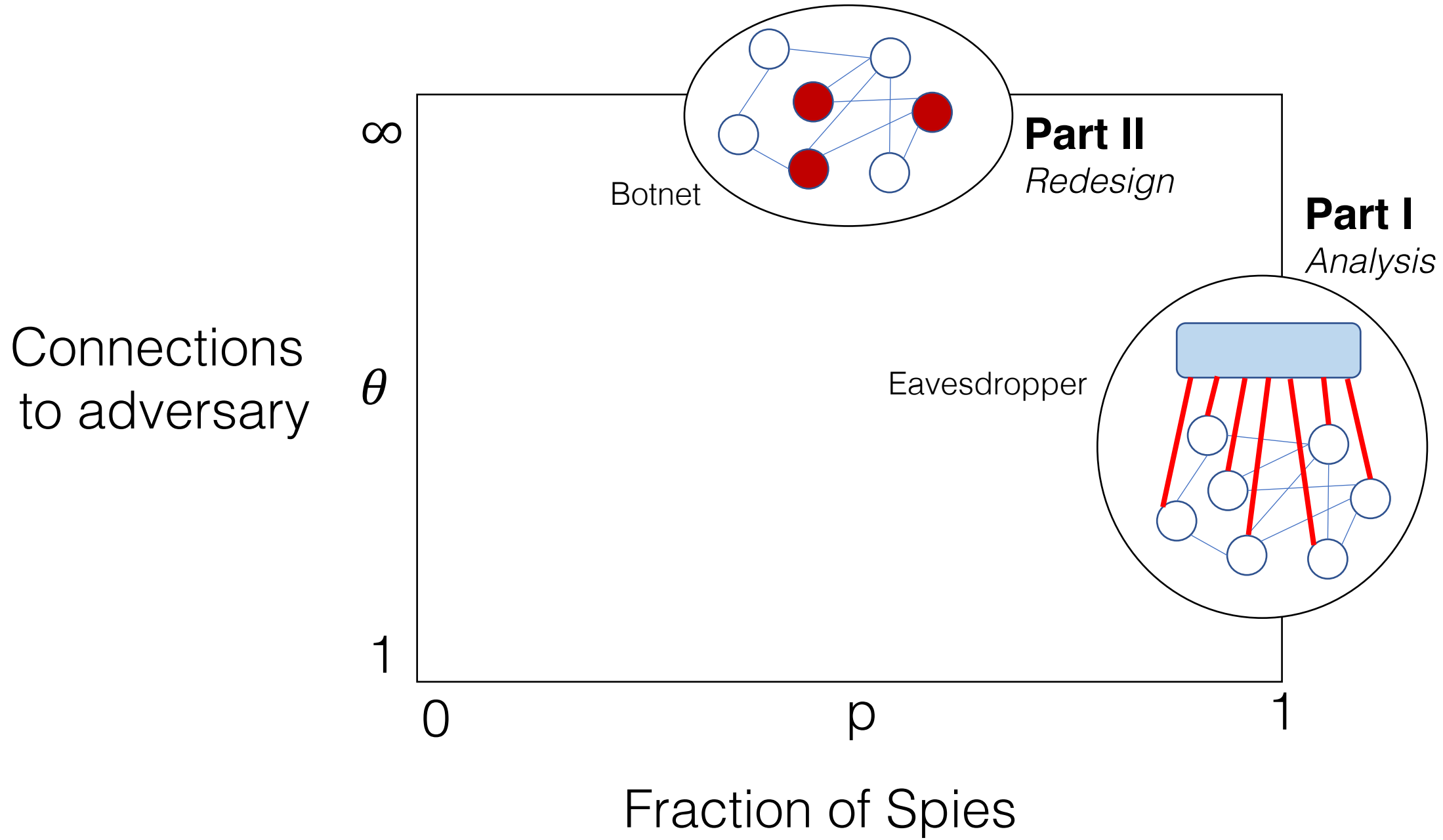


What the eavesdropper can do about it



Summary of adversarial model



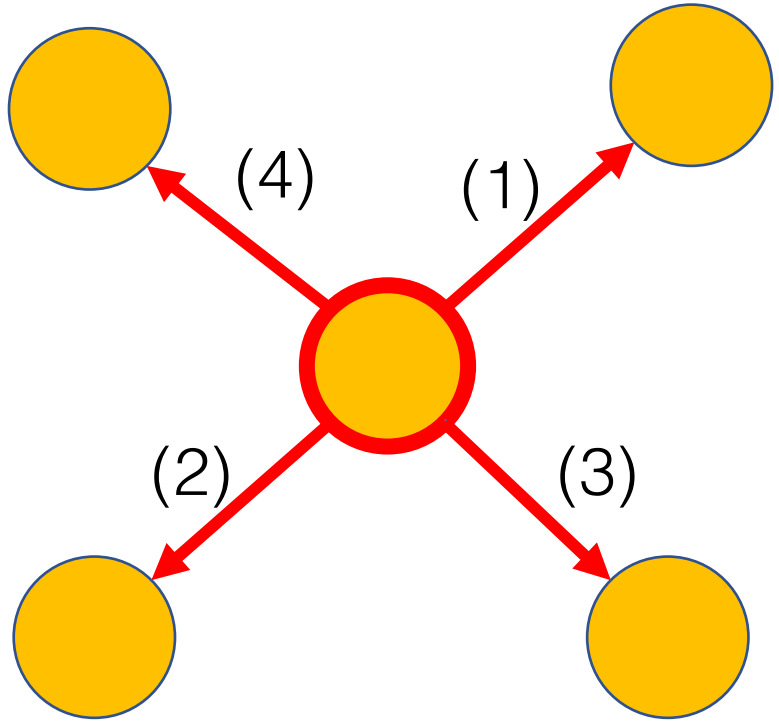

Connections to adversary
∞
θ
1
0
p
1
Fraction of Spies
Botnet
Part II
Redesign
Part I
Analysis
Eavesdropper

Analysis

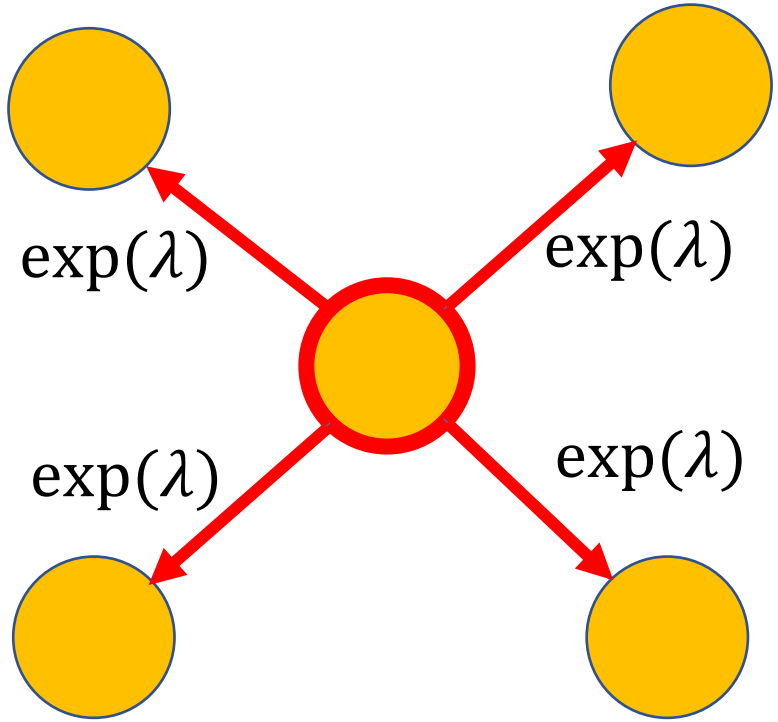
How bad is the problem?

Flooding Protocols

Trickle (pre-2015)



Diffusion (post-2015)

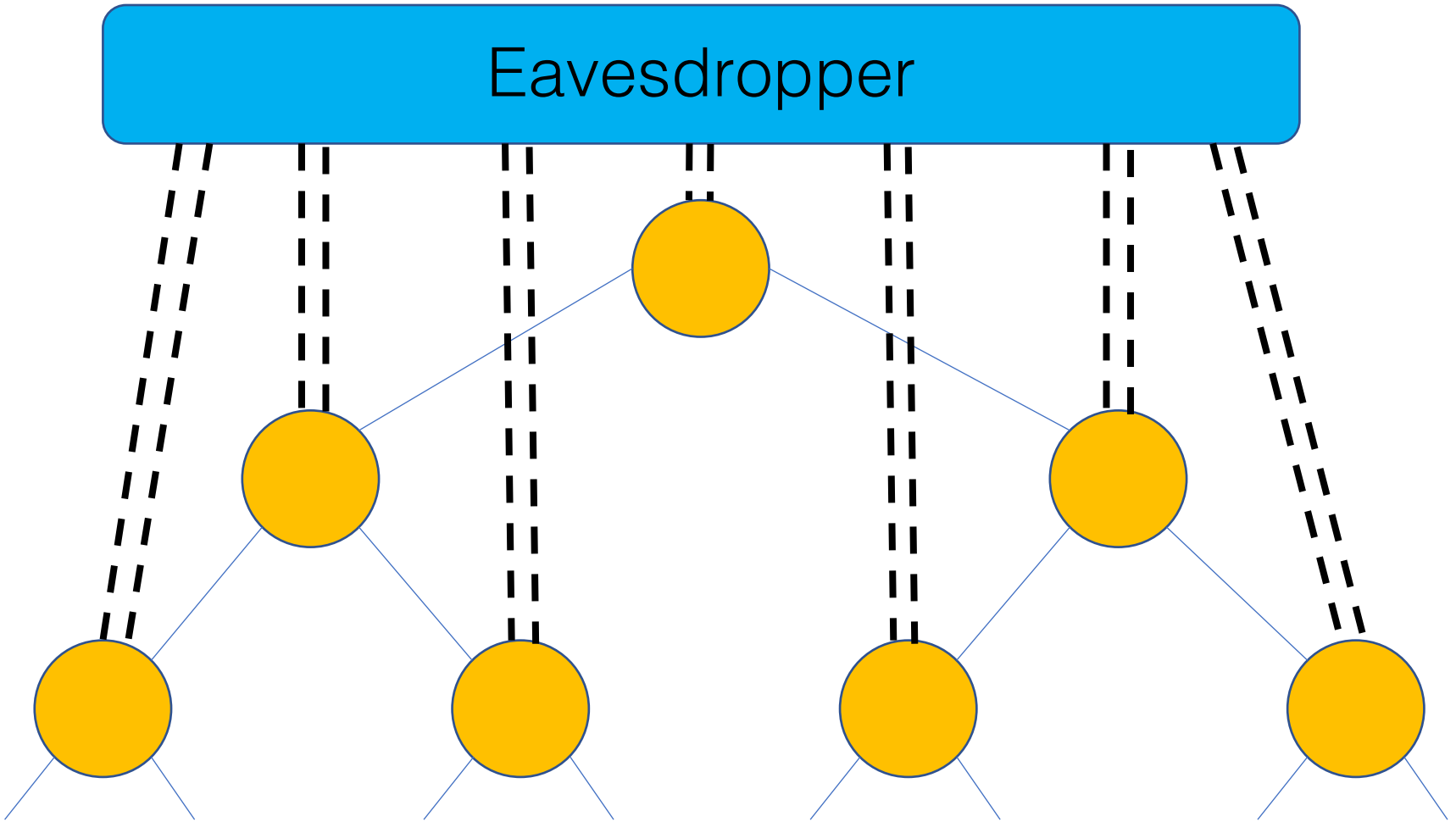


Does diffusion provide stronger anonymity than trickle spreading?

d-regular trees

Fraction of spies $p = 1$

Arbitrary number of connections θ



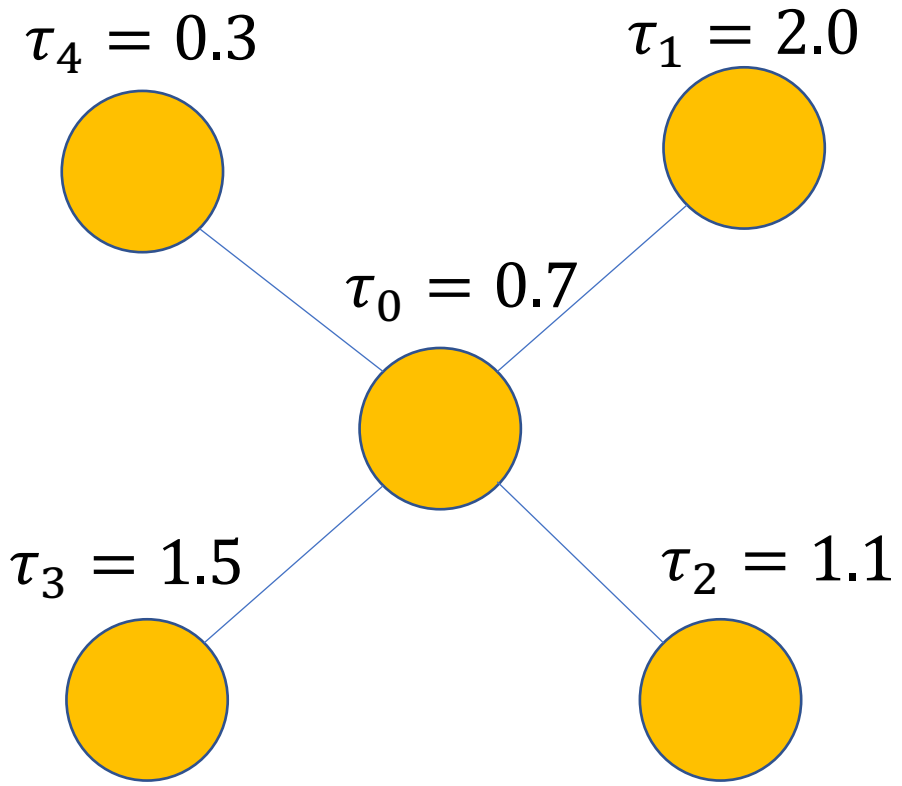
Anonymity Metric

$$P(\text{detection}|\boldsymbol{\tau}, G)$$

timestamps

graph

$$\boldsymbol{\tau} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \dots \\ \tau_n \end{bmatrix}$$

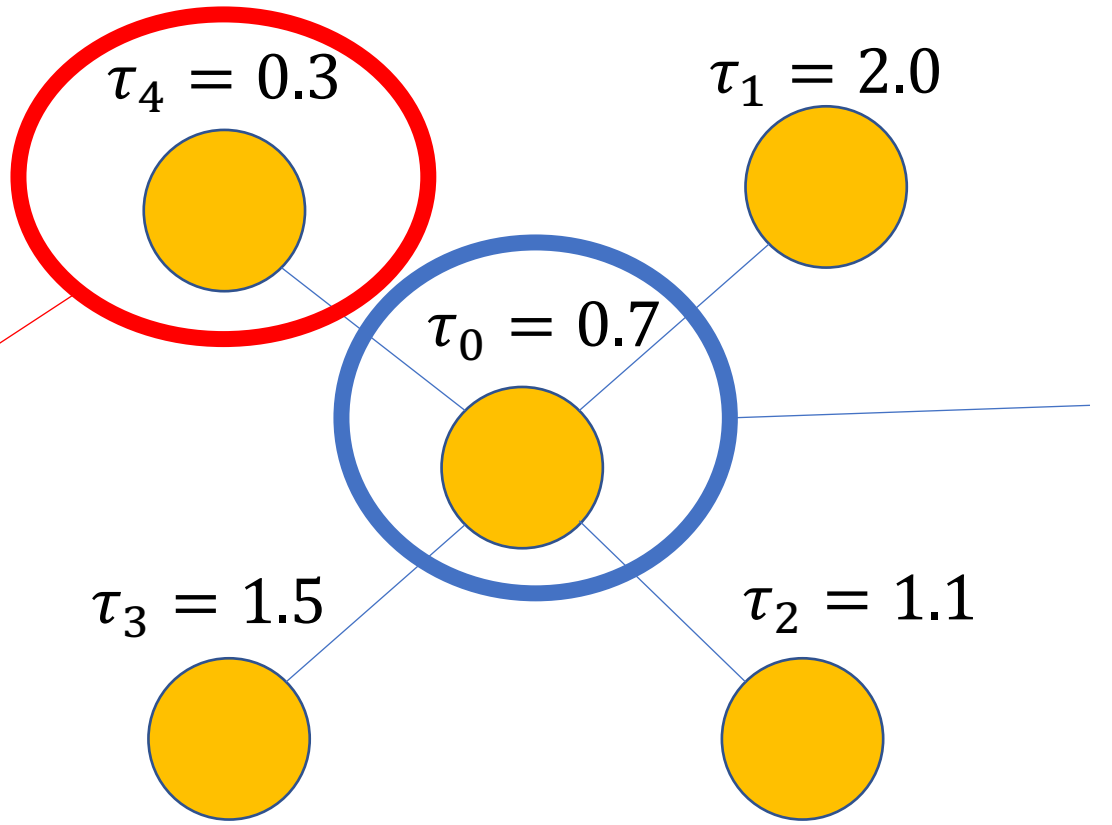


Estimators

$$P(\text{detection}|\boldsymbol{\tau}, G)$$

timestamps

graph



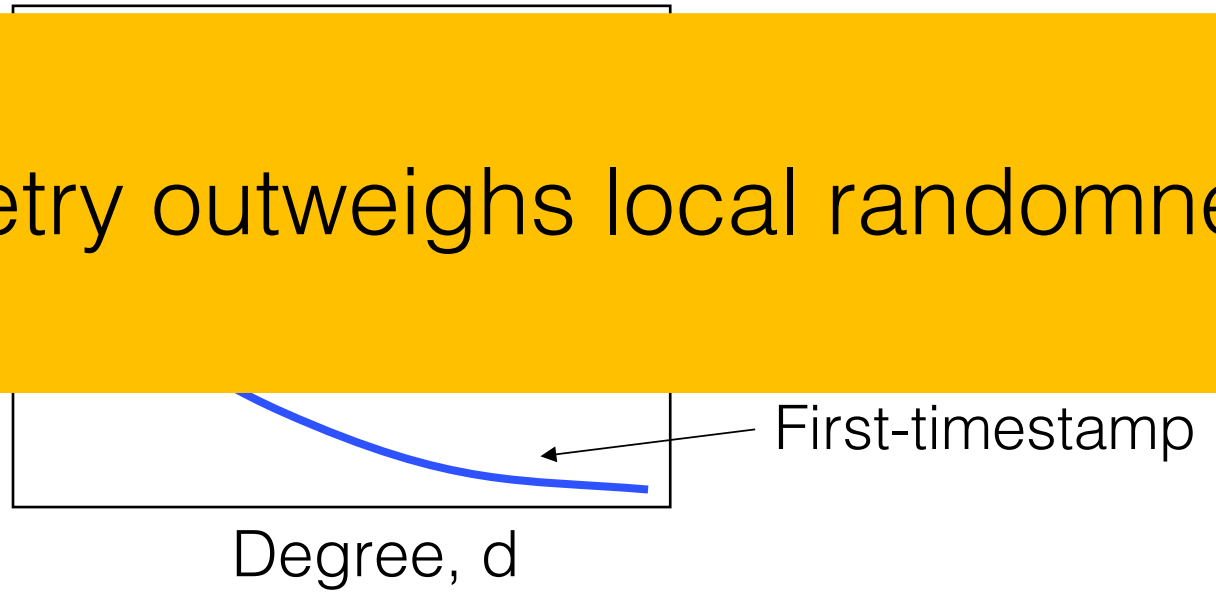
First-Spy

Maximum-Likelihood

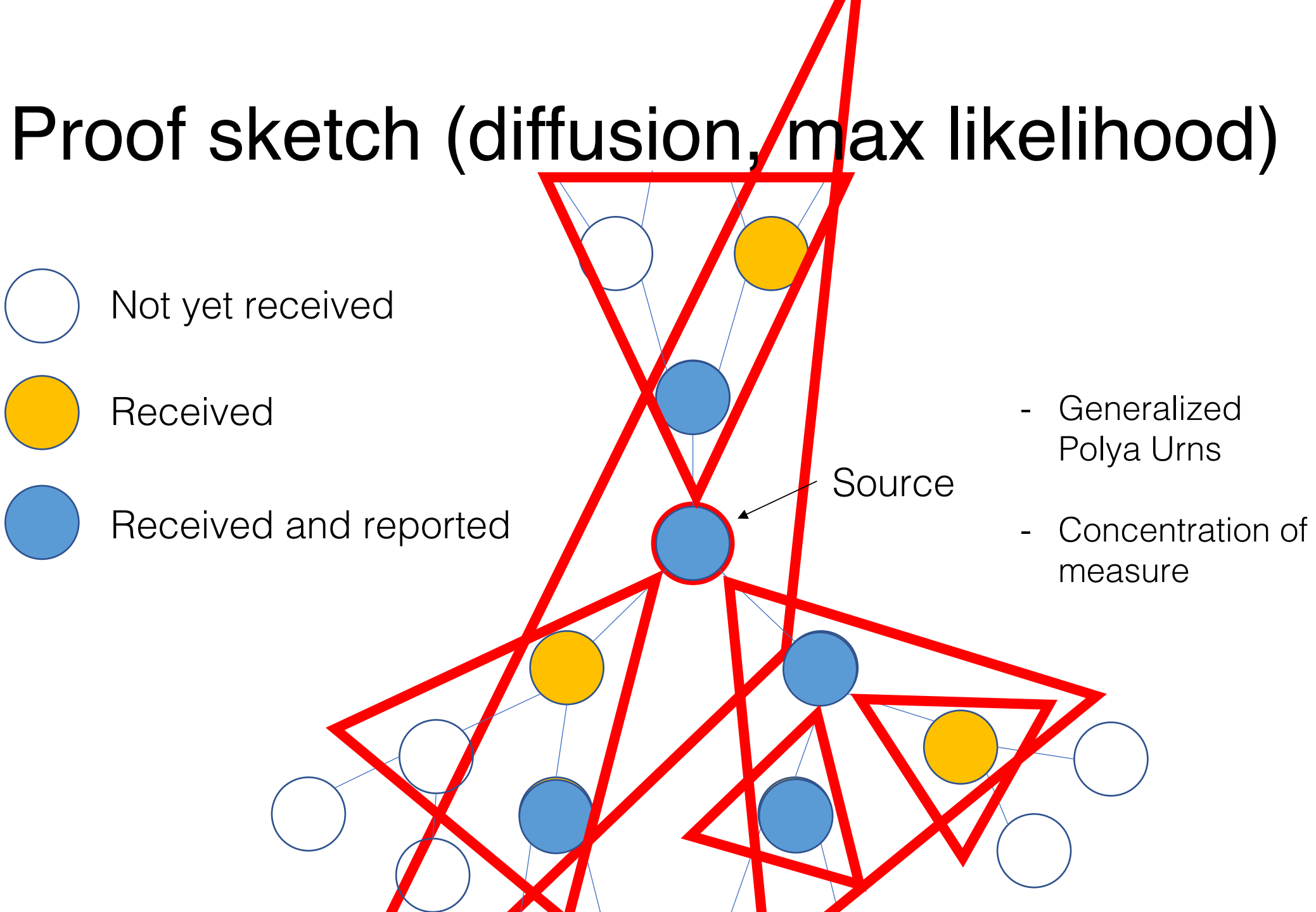
Results: d-Regular Trees

	Trickle	Diffusion
First-Timestamp	$o\left(\frac{\log d}{d}\right)$	$o\left(\frac{\log d}{d}\right)$
Maximum-Likelihood	$\Omega(1)$	$\Omega(1)$

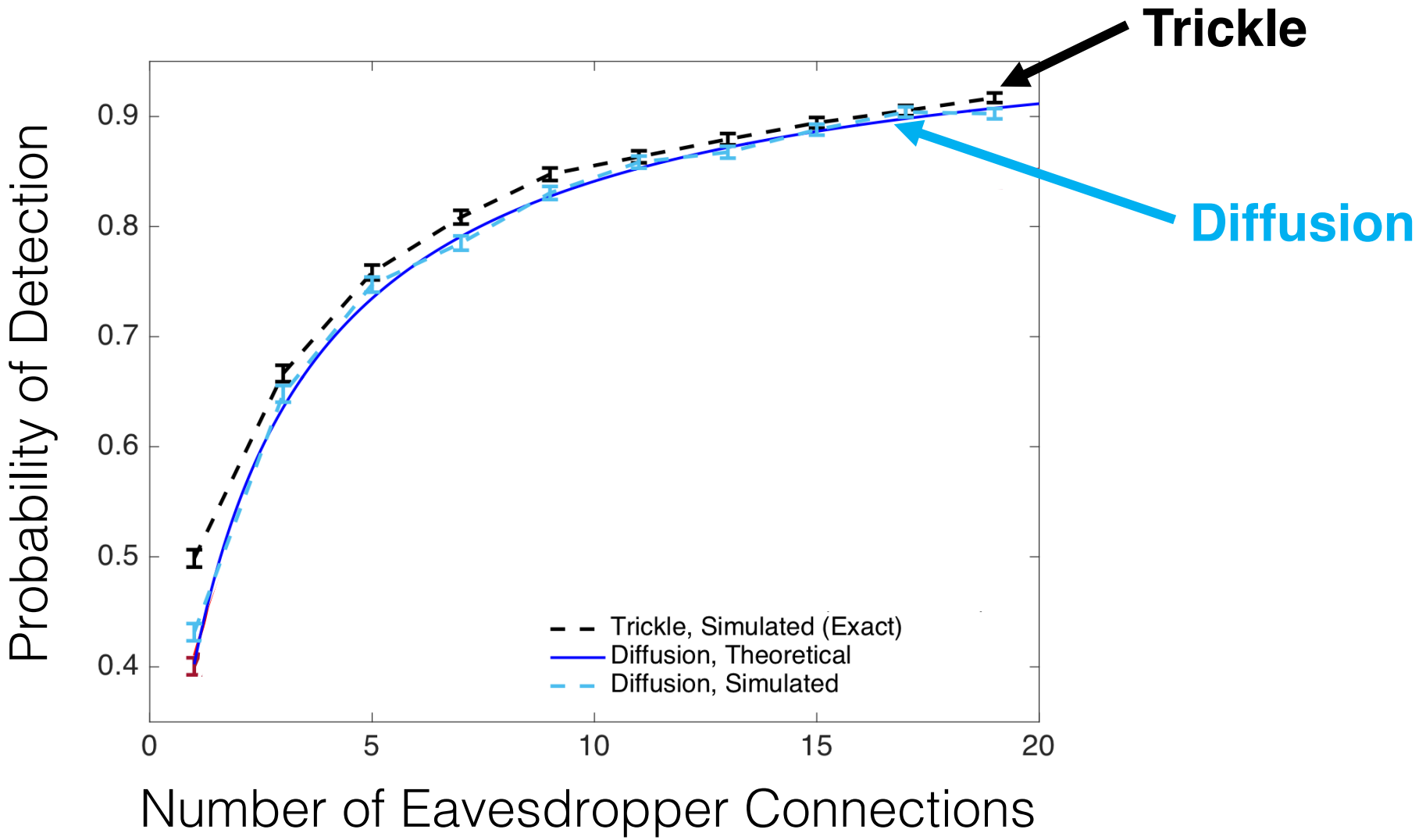
Intuition: Symmetry outweighs local randomness!



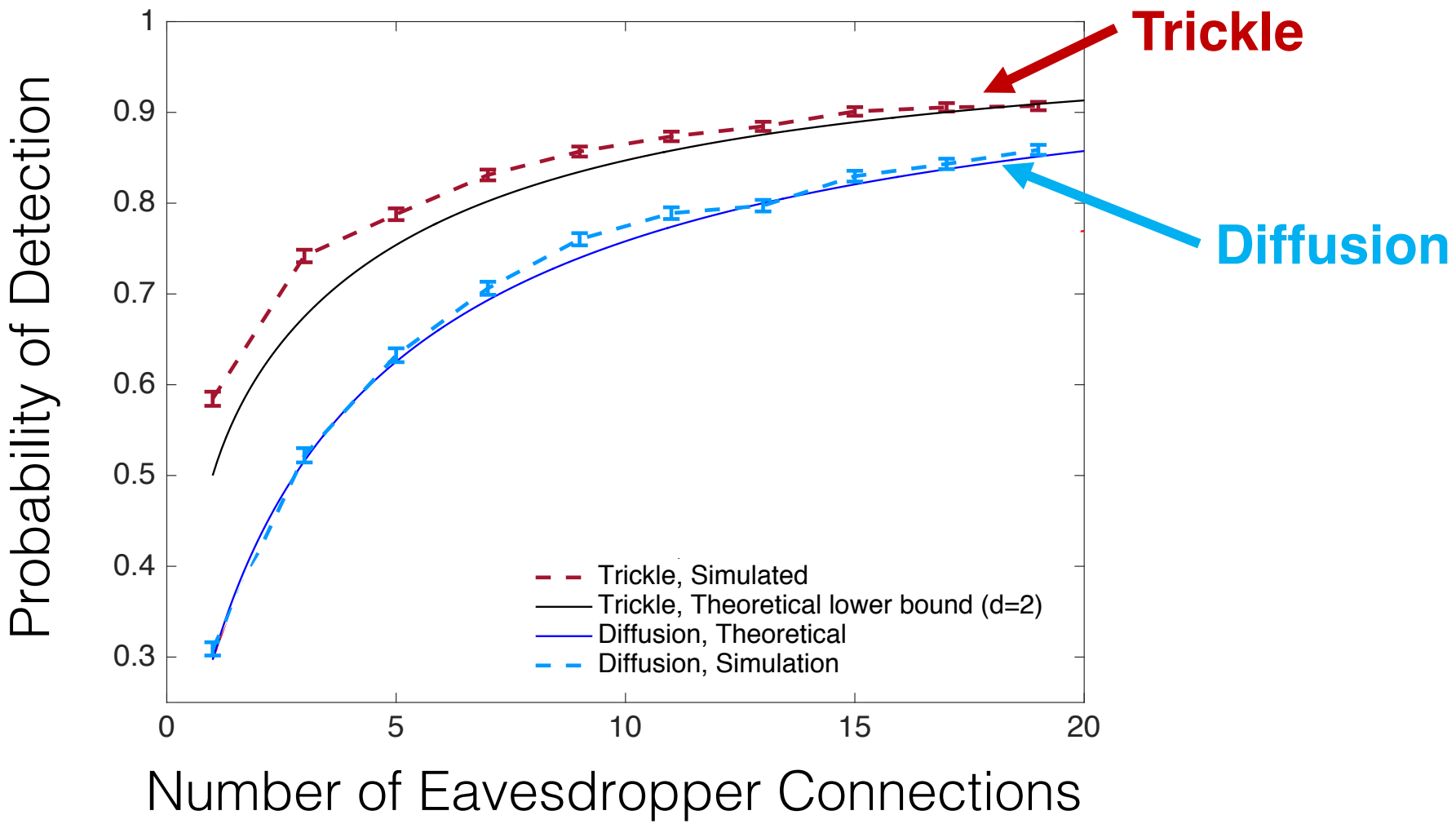
Proof sketch (diffusion, max likelihood)



Results: Trees



Results: Bitcoin Graph

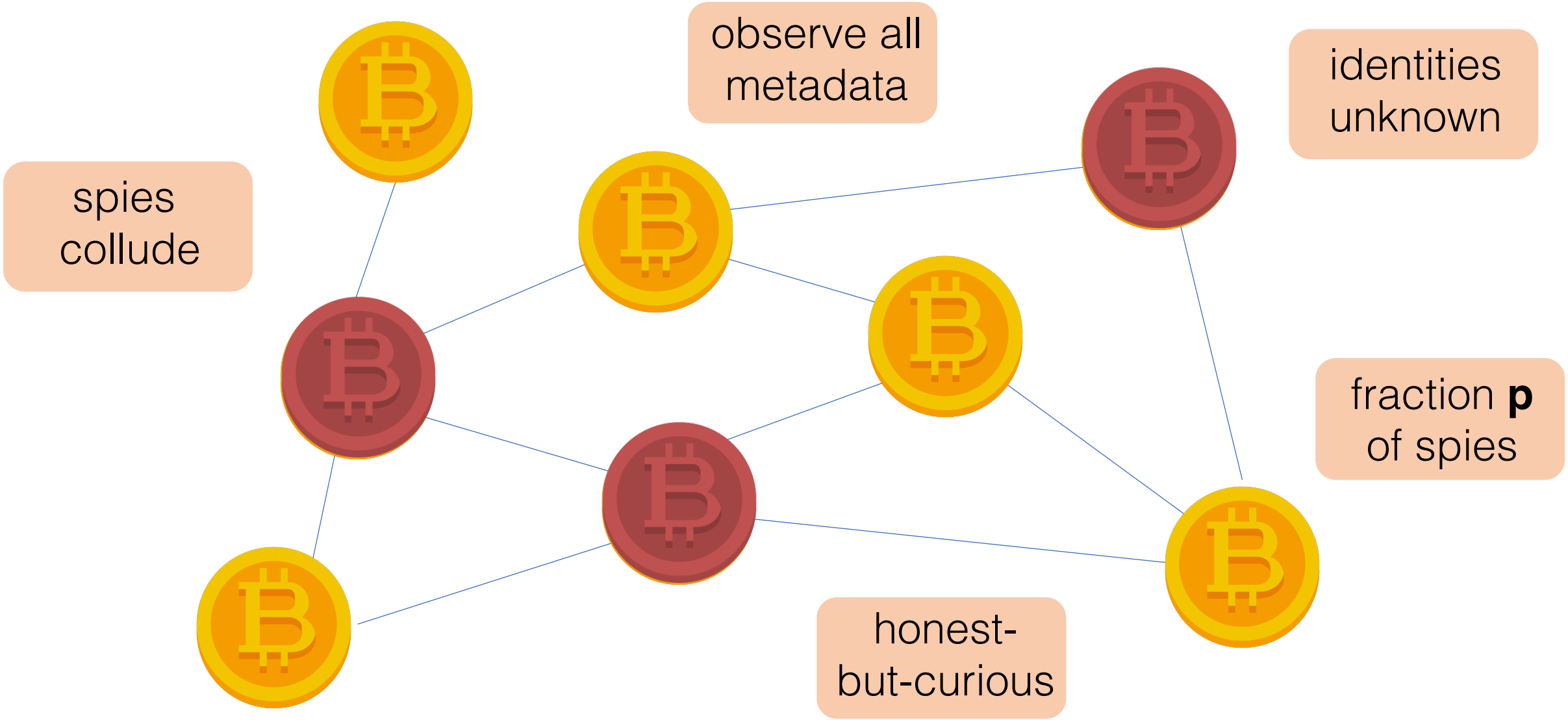


Diffusion does not have (significantly) better anonymity properties than trickle.

Redesign

Can we design a better network?

Botnet adversarial model



Metric for Anonymity

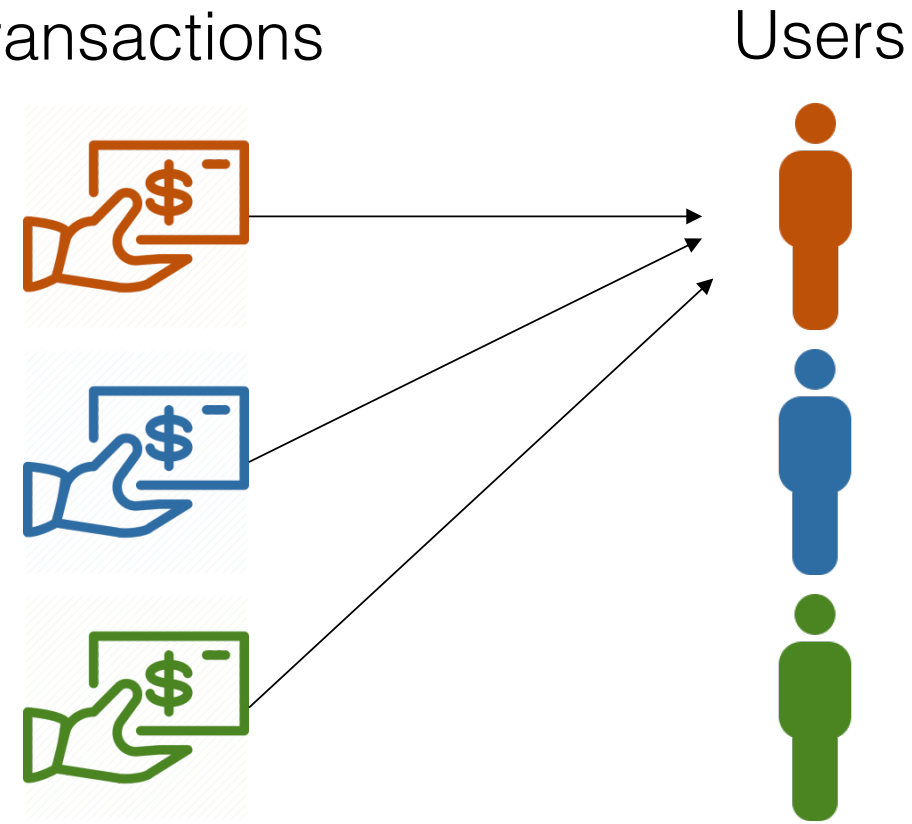
Recall

$$\frac{1}{n}\sum_{v} 1\{M(v\text{'s tx}) = v\}$$

Number honest users

User

Mapping



Precision

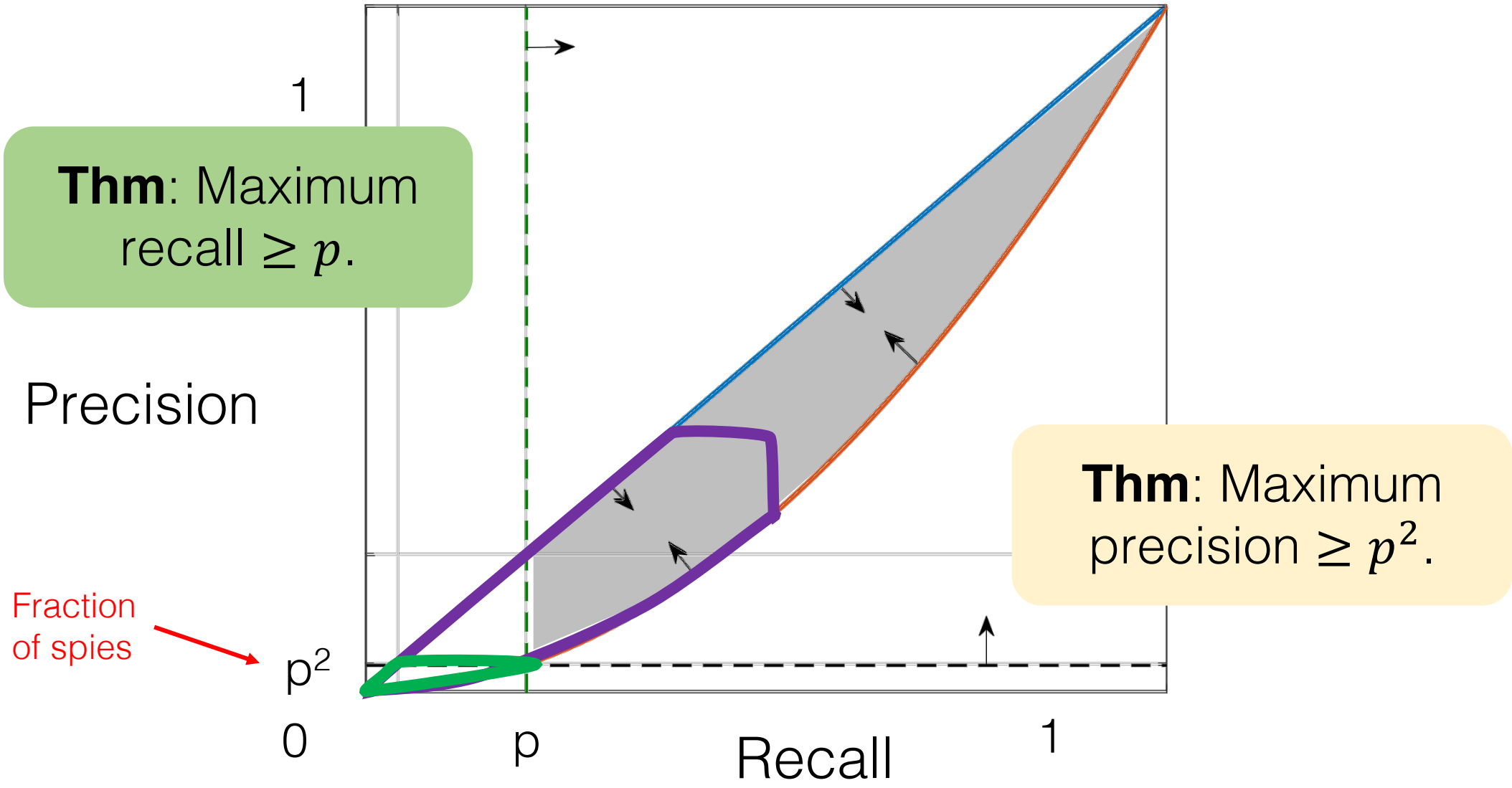
$$\frac{1}{n}\sum_{v} \frac{1\{M(v\text{'s tx}) = v\}}{\#\text{ tx mapped to v}}$$

$\mathbb{E}[\text{Recall}]$ = Probability of Detection

Goal:

Design a distributed flooding protocol that minimizes the maximum precision and recall achievable by a computationally-unbounded adversary.

Fundamental Limits

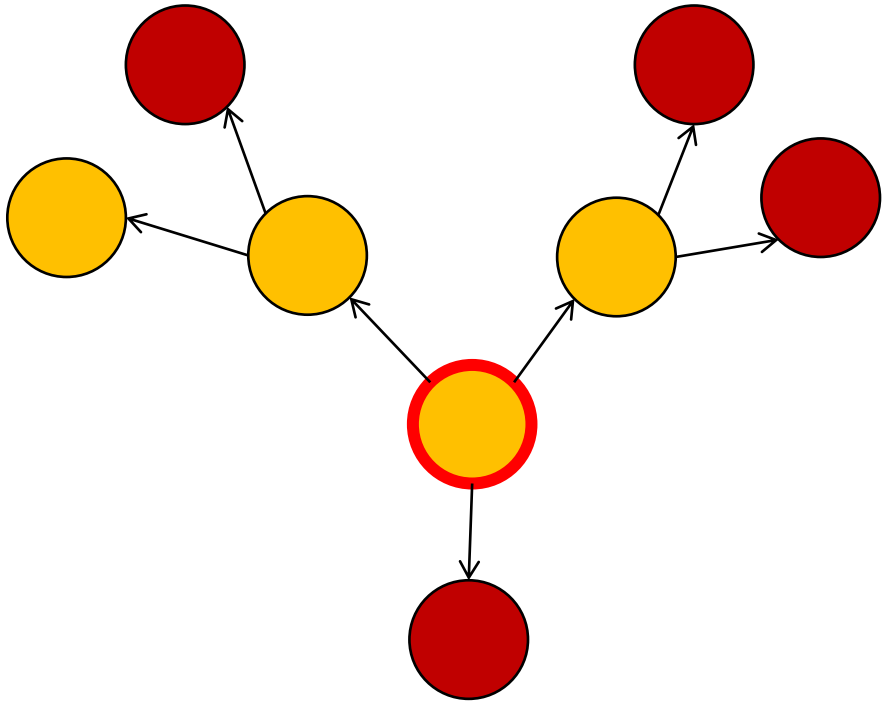


Thm: Maximum recall $\geq p$.

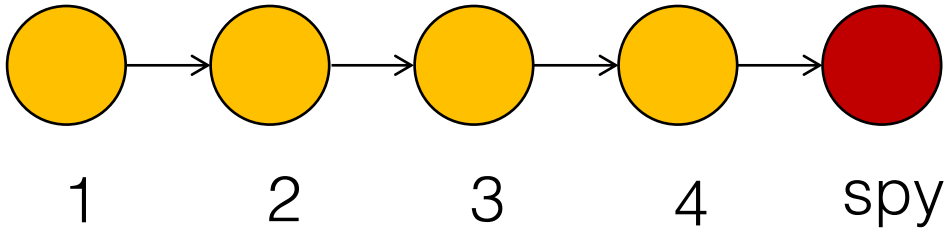
Thm: Maximum precision $\geq p^2$.

What are we looking for?

Asymmetry



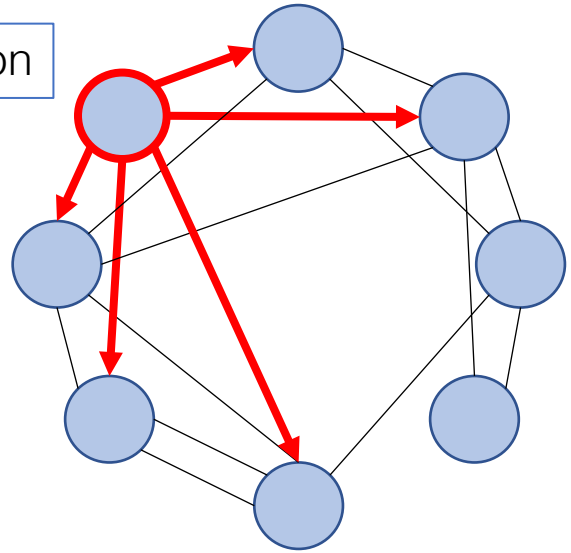
Mixing



What can we control?

Spreading Protocol

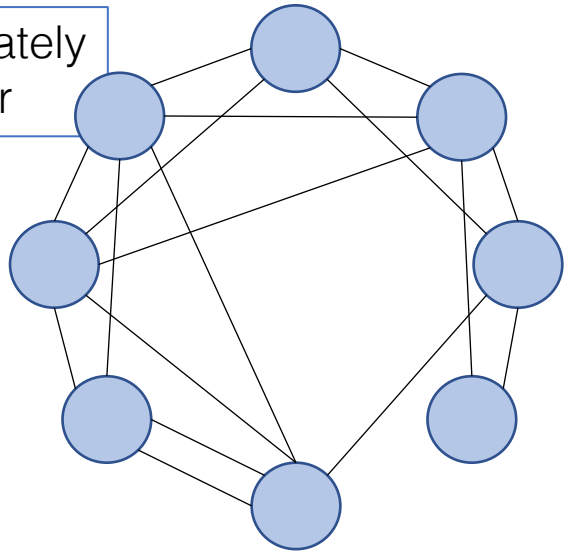
Diffusion



Given a graph, how do we spread content?

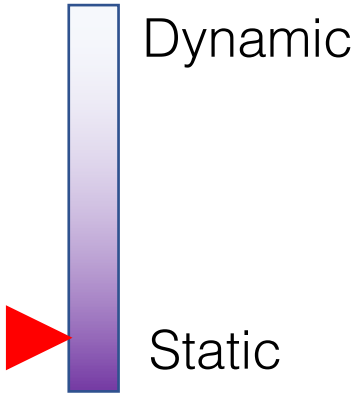
Topology

Approximately regular



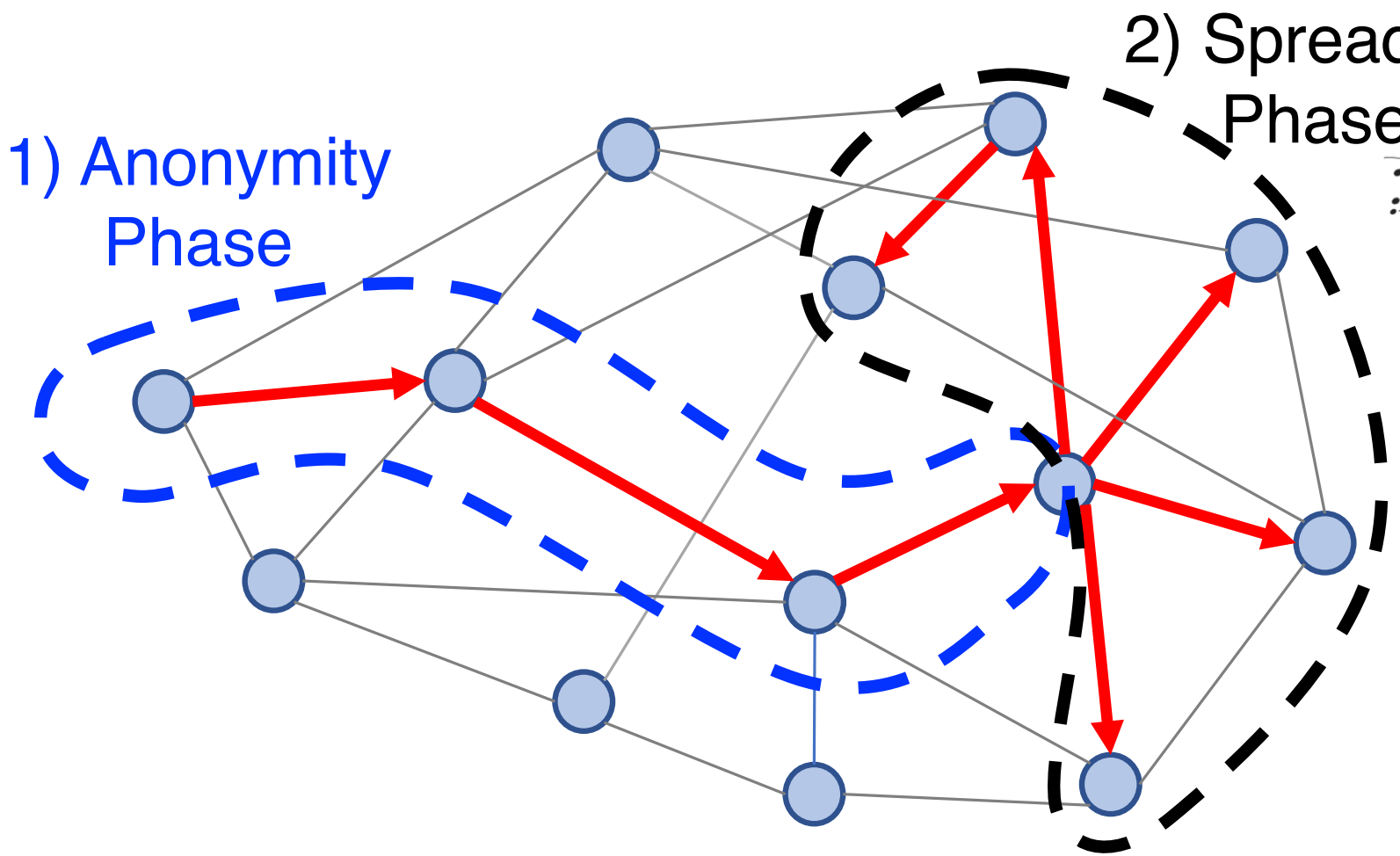
What is the underlying graph topology?

Dynamicity



How often does the graph change?

Spreading Protocol: Dandelion



Why Dandelion spreading?

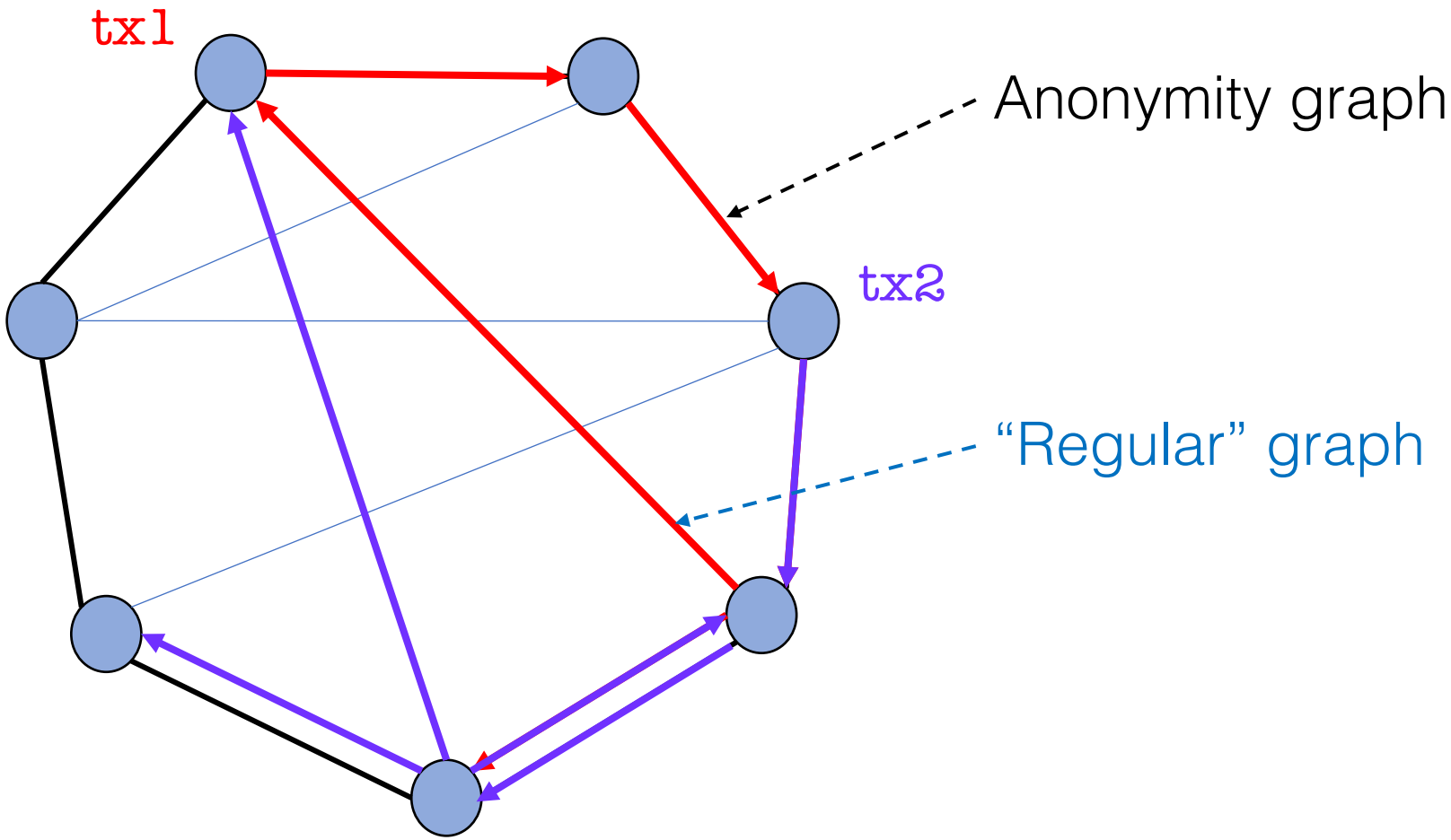
Theorem: Dandelion spreading has an optimally low maximum recall of $p + O\left(\frac{1}{n}\right)$.

lower bound = p

fraction of spies

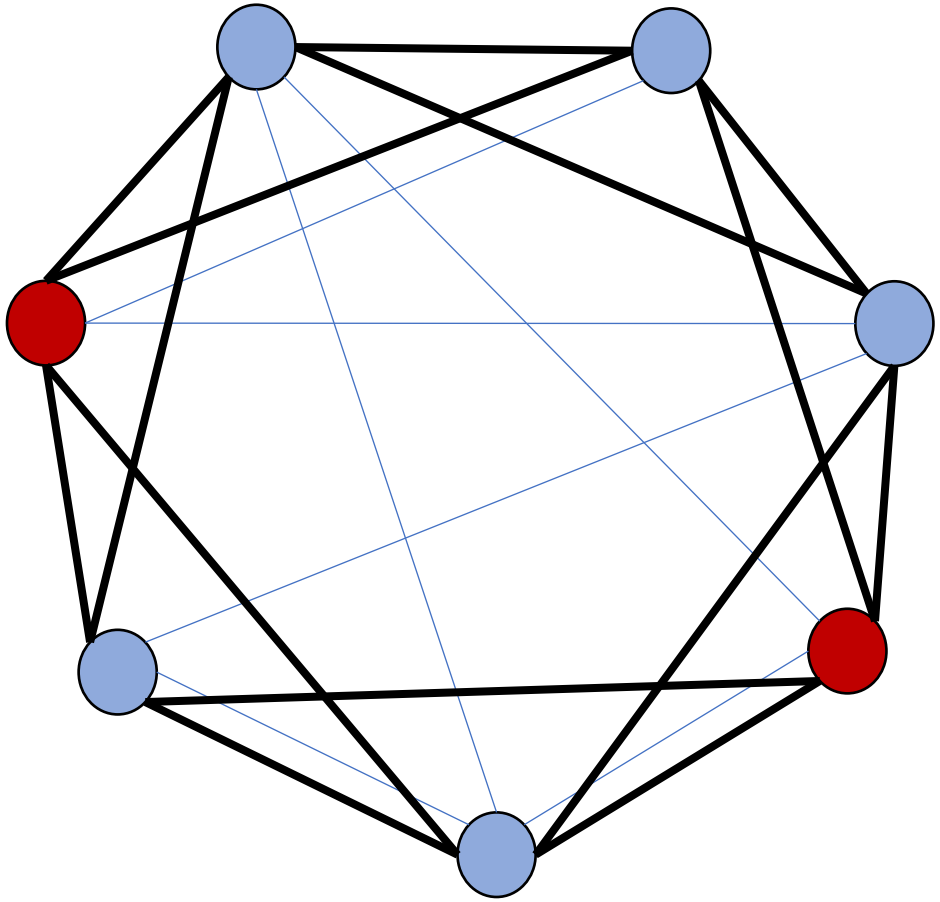
number of nodes

Graph Topology: Line



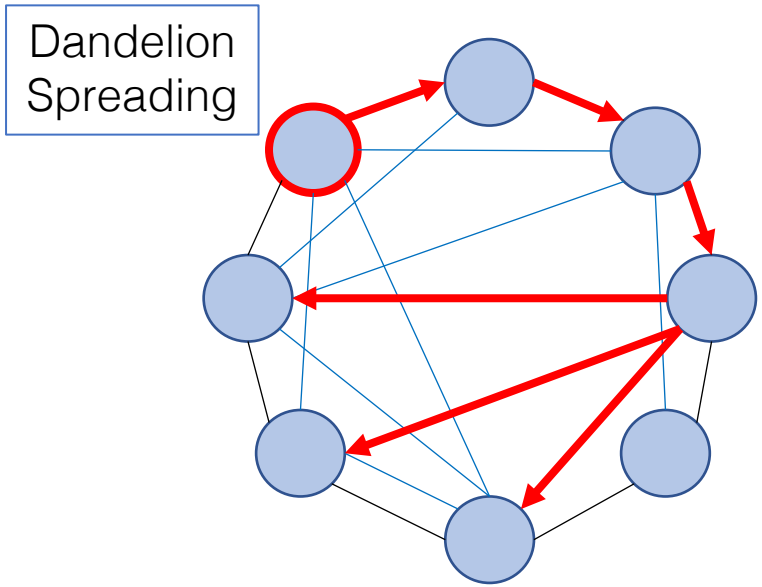
Dynamicity: High

Change the anonymity graph frequently.



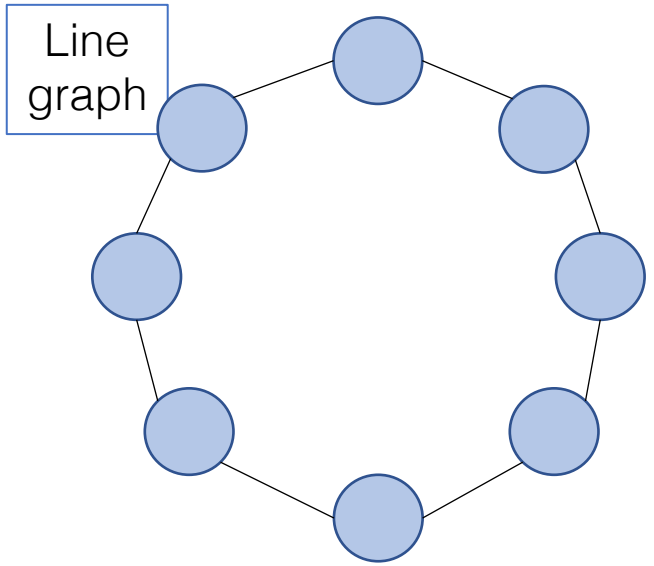
DANDELION Network Policy

Spreading Protocol



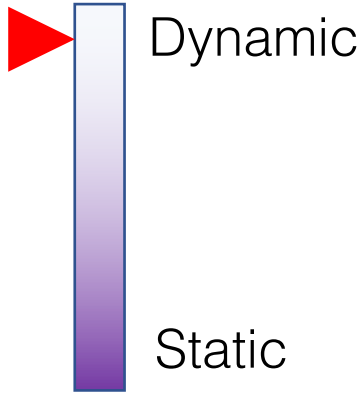
Given a graph, how do we spread content?

Topology



What is the anonymity graph topology?

Dynamicity



How often does the graph change?

lower bound = p^2

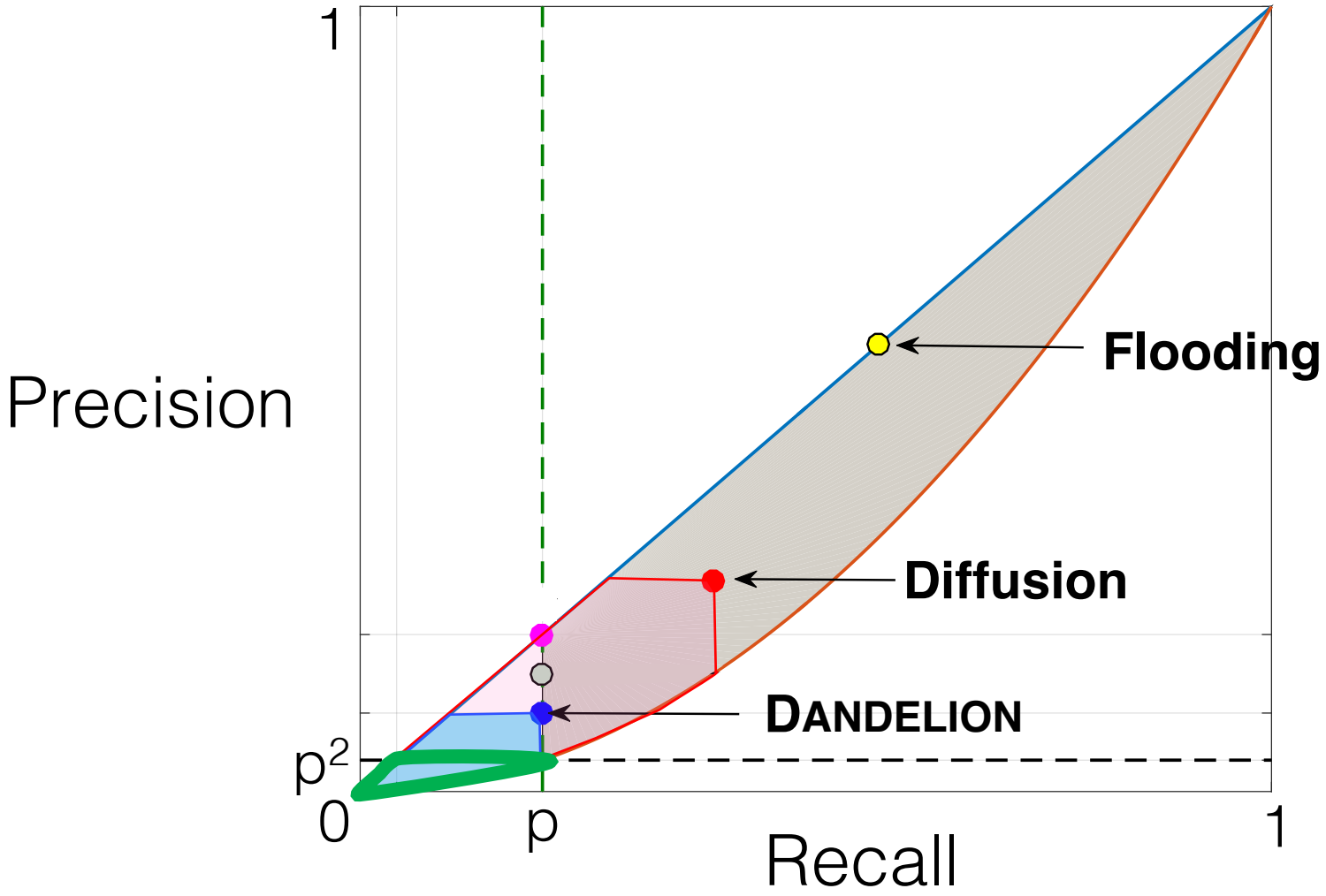
Theorem: DANDELION has a nearly-optimal maximum precision of $\frac{2p^2}{1-p}\log\left(\frac{2}{p}\right) + O\left(\frac{1}{n}\right)$.*

fraction of spies

number of nodes

*For $p < \frac{1}{3}$

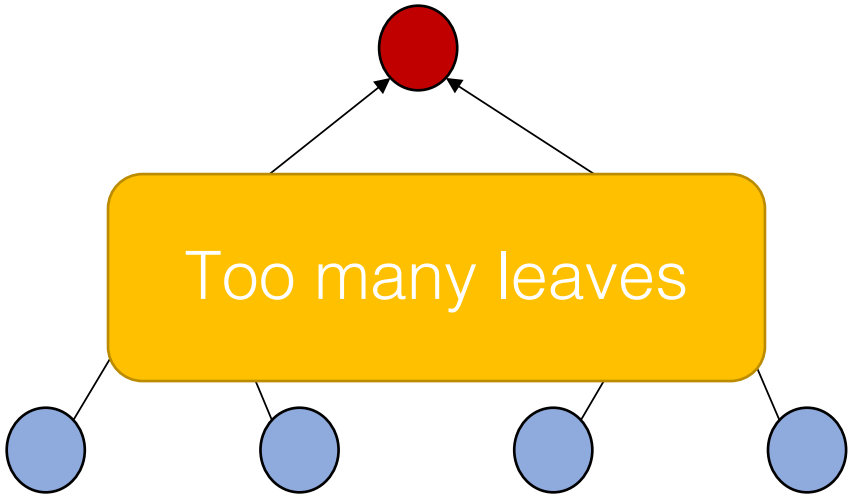
Performance: Achievable Region



Why does DANDELION work?

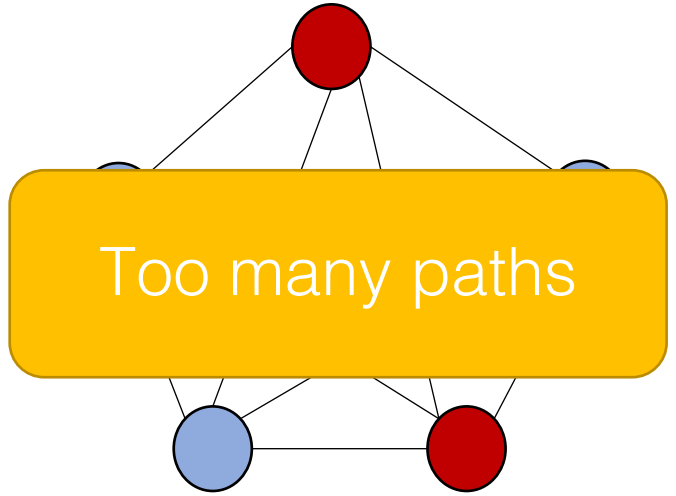
Strong mixing properties.

Tree



Precision: $O(p)$

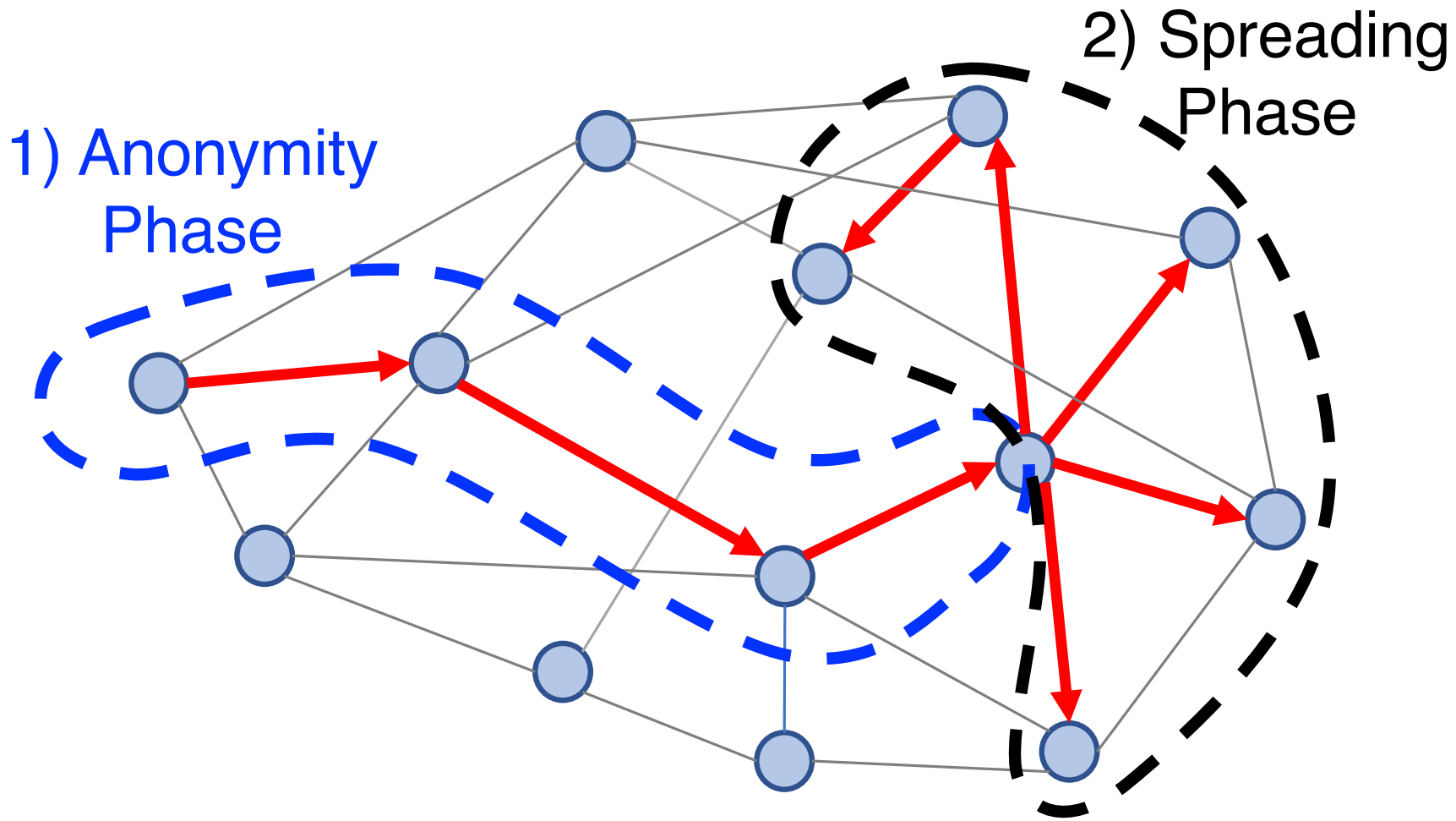
Complete graph



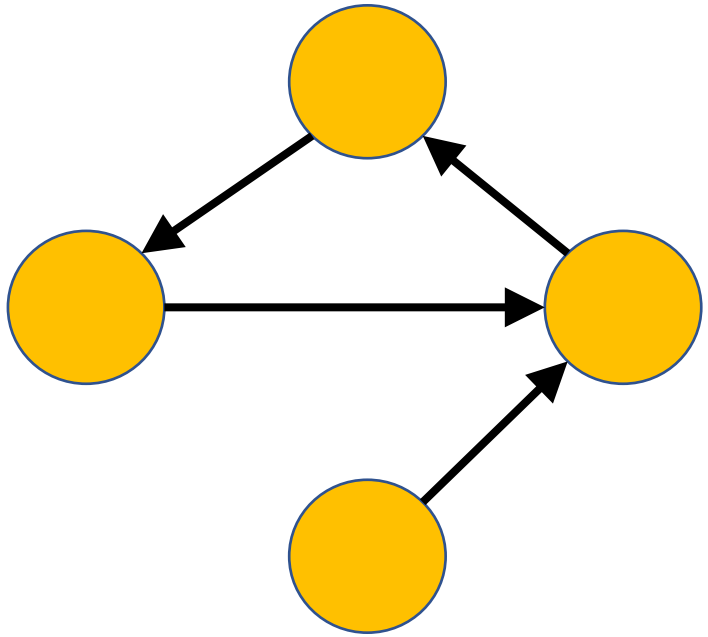
Precision: $\frac{p}{1-p}(1-e^{p-1})$

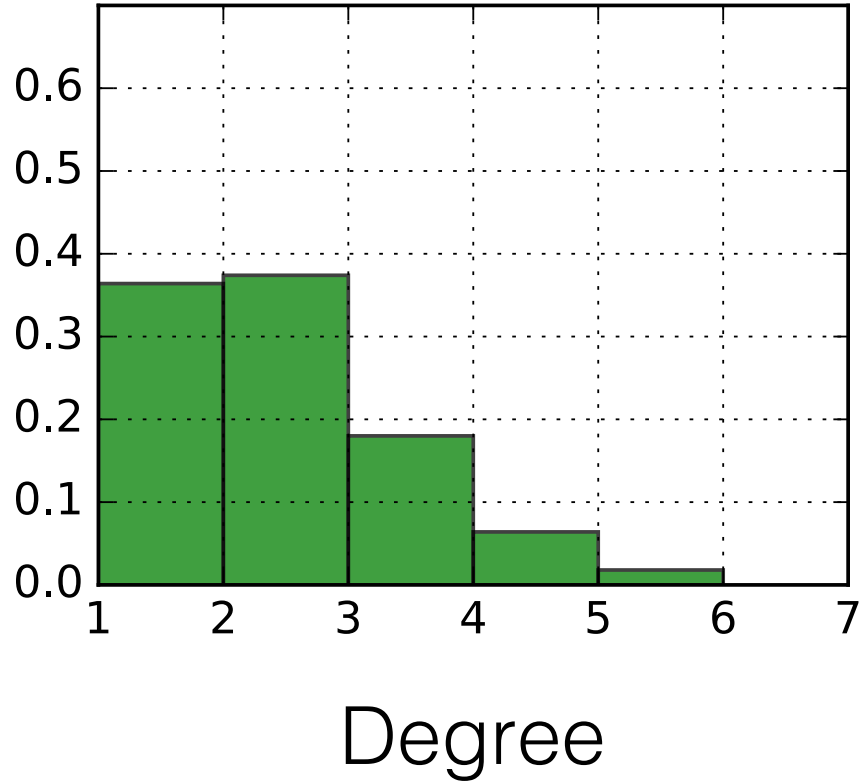
How practical is this?

Dandelion spreading

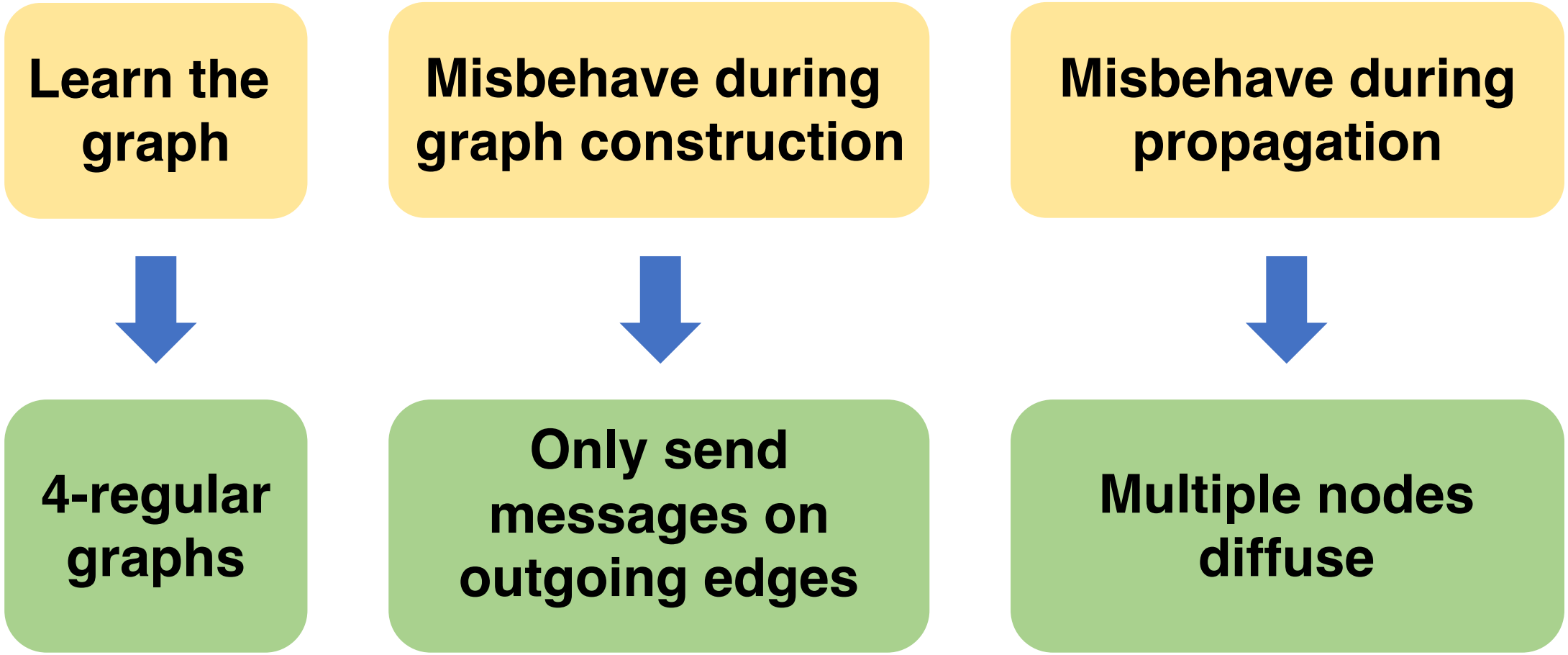


Anonymity graph construction

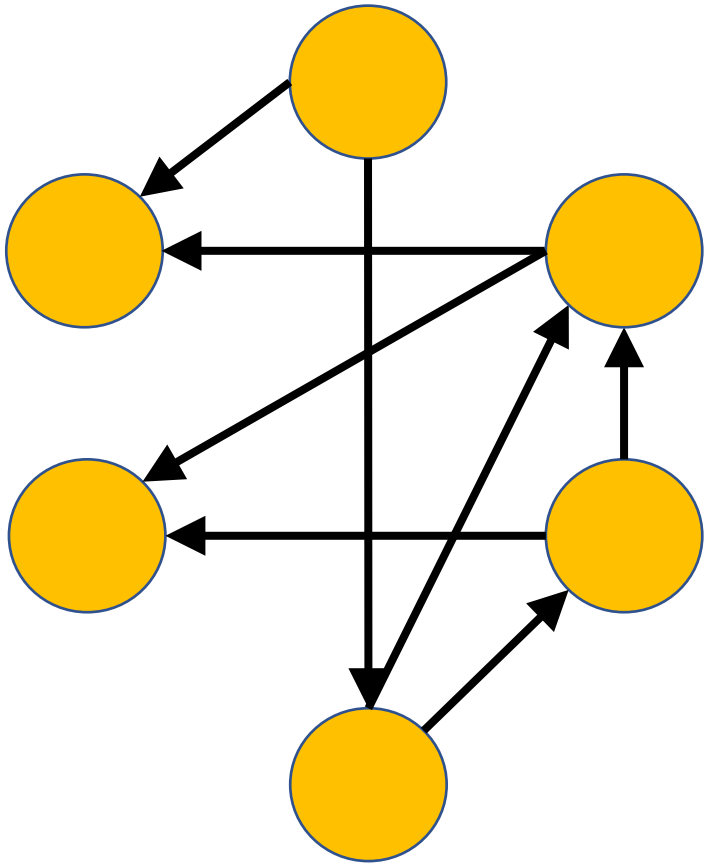




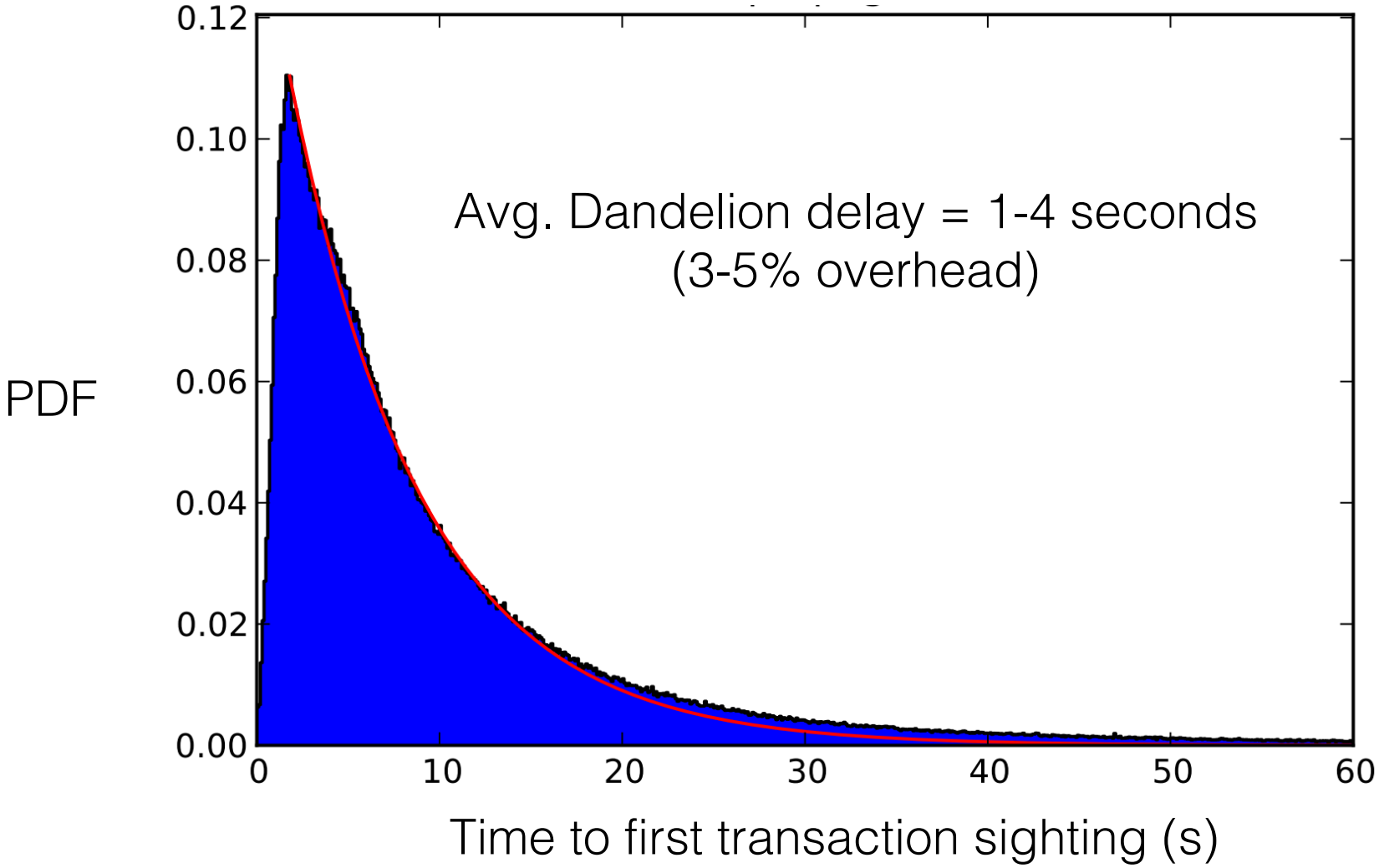
Dealing with stronger adversaries



Anonymity graph construction

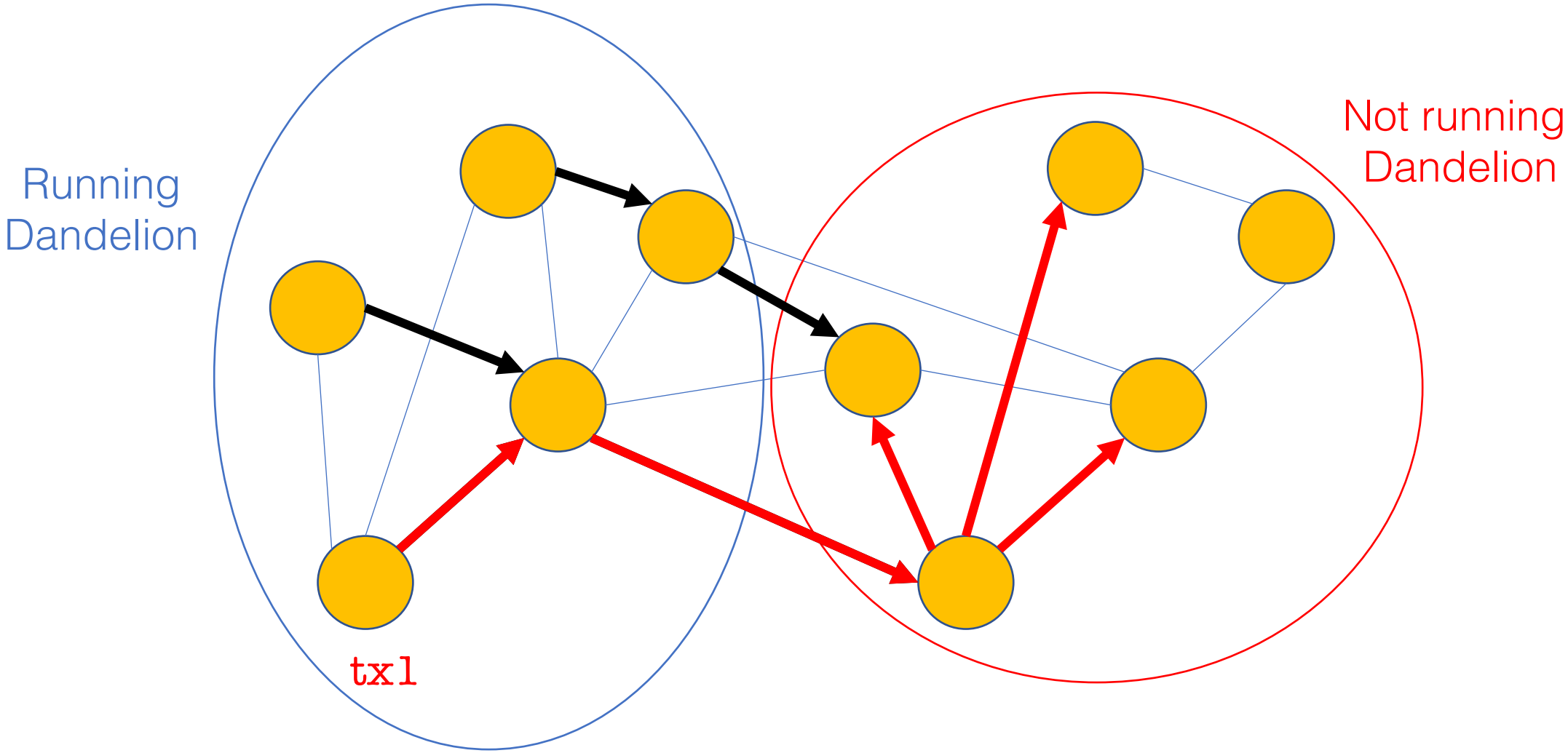


Latency Overhead: Estimate



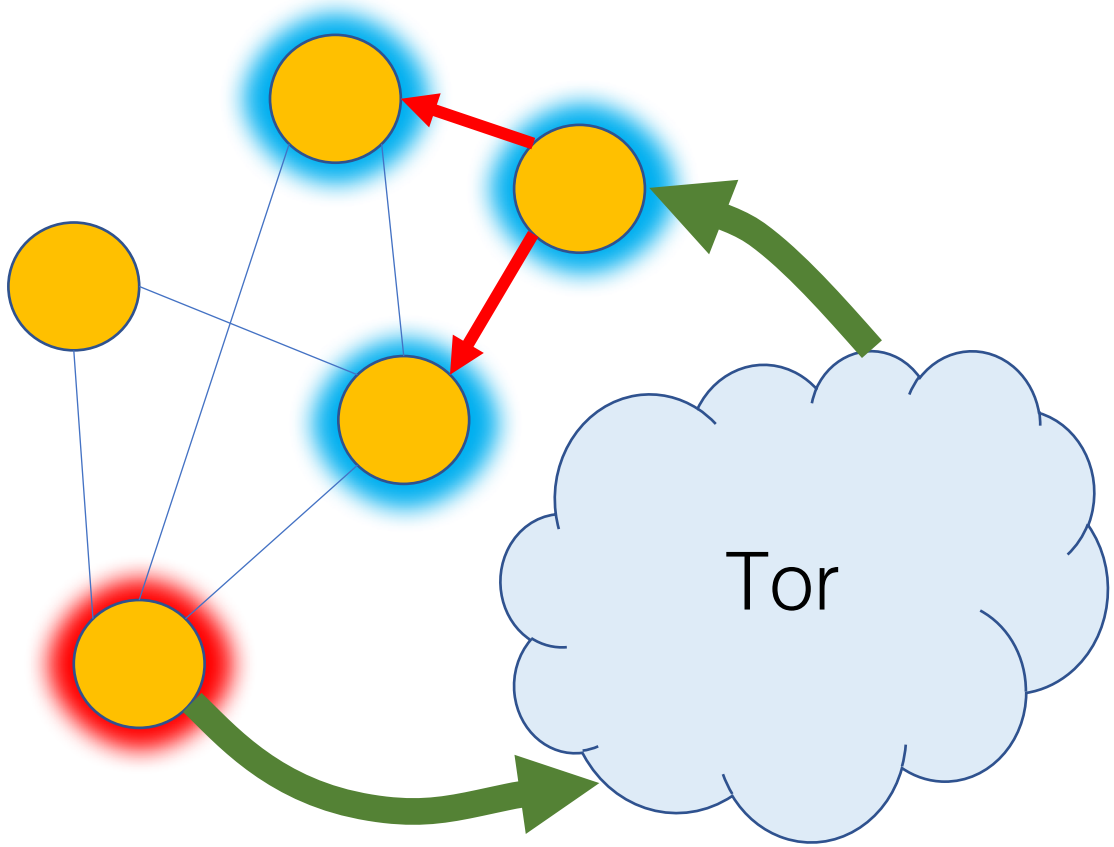
Information Propagation in the Bitcoin Network, Decker and Wattenhofer, 2013

Deployment considerations

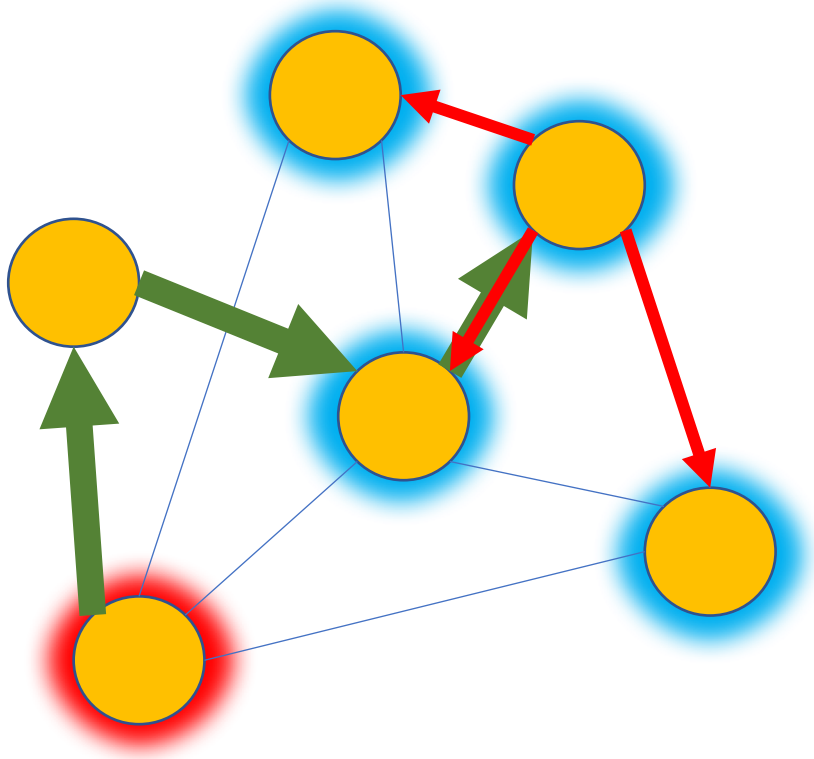


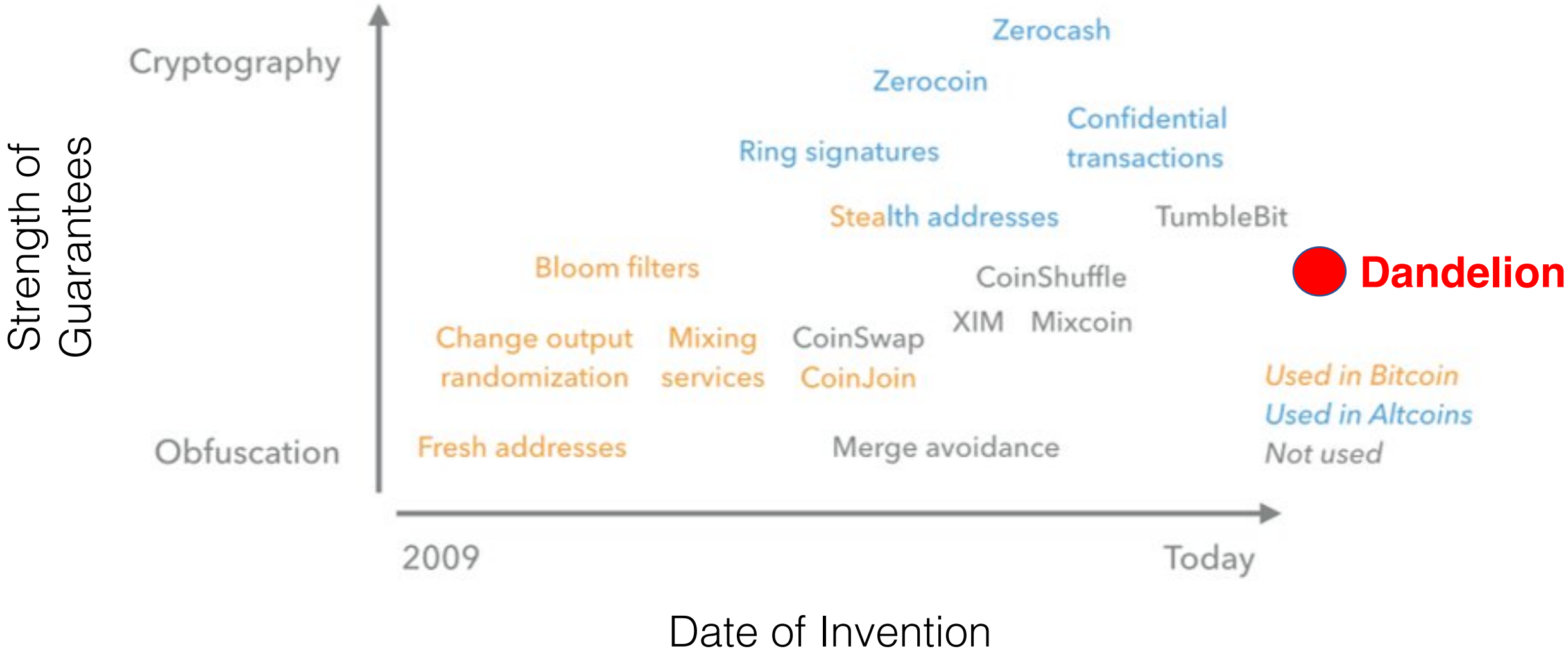
Why not alternative solutions?

Connect through Tor



I2P Integration (e.g. Monero)





Narayanan and Möser, 2017

Take-Home Messages

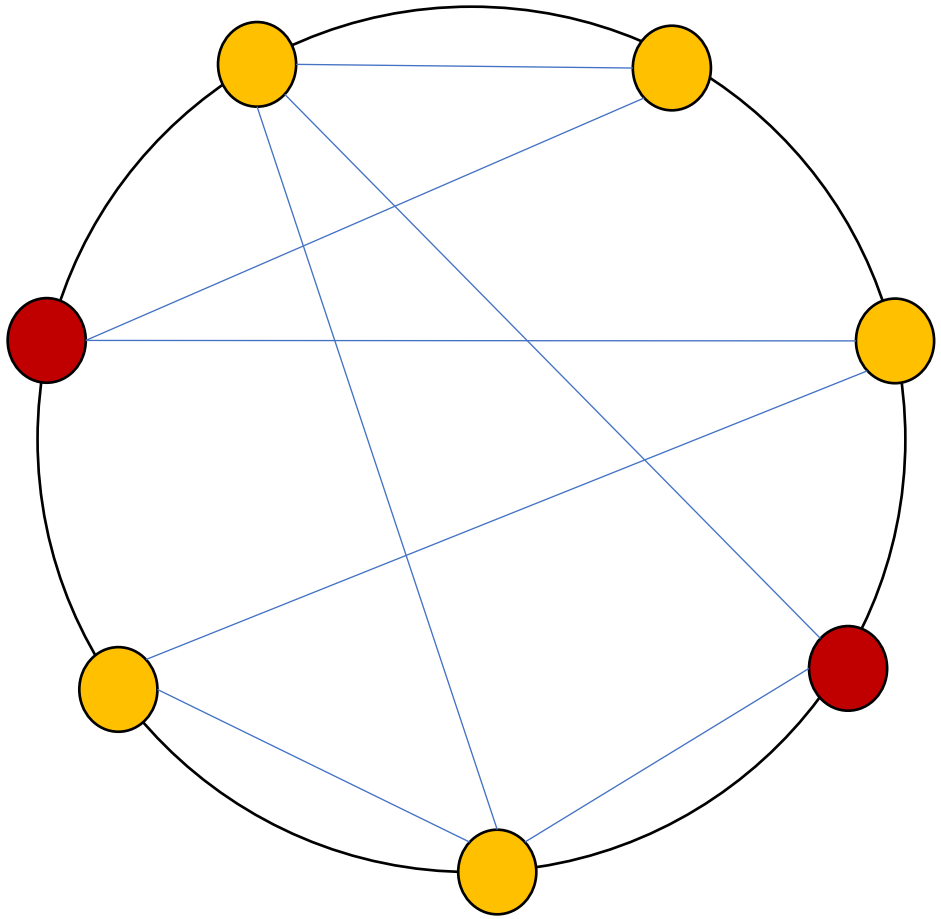
1) Bitcoin's P2P network has poor anonymity.

2) Moving from trickle to diffusion did not help.

3) DANDELION may be a lightweight solution for certain classes of adversaries.

https://github.com/gfanti/bitcoin

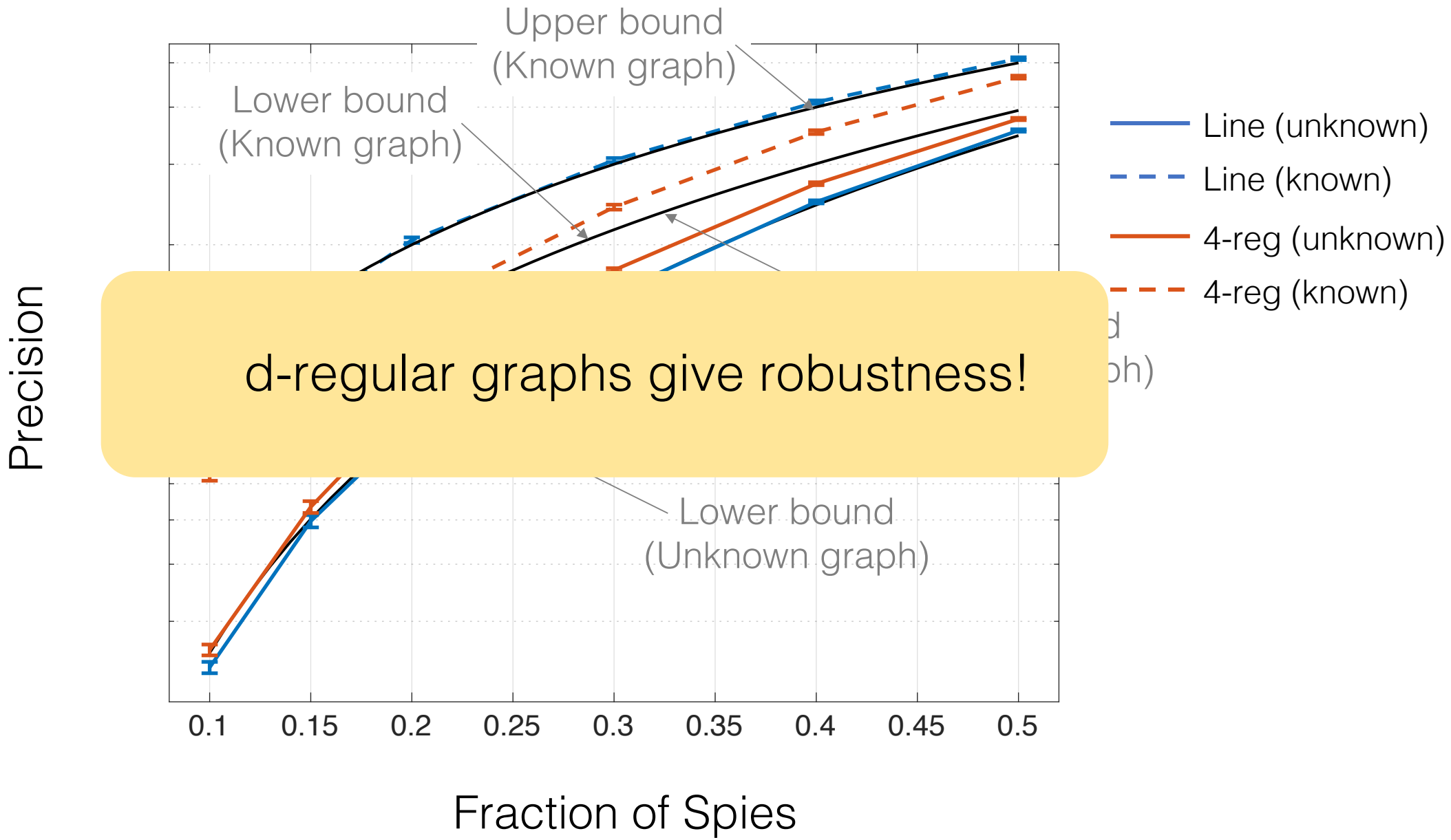
DANDELION vs. Tor, Crowds, etc.

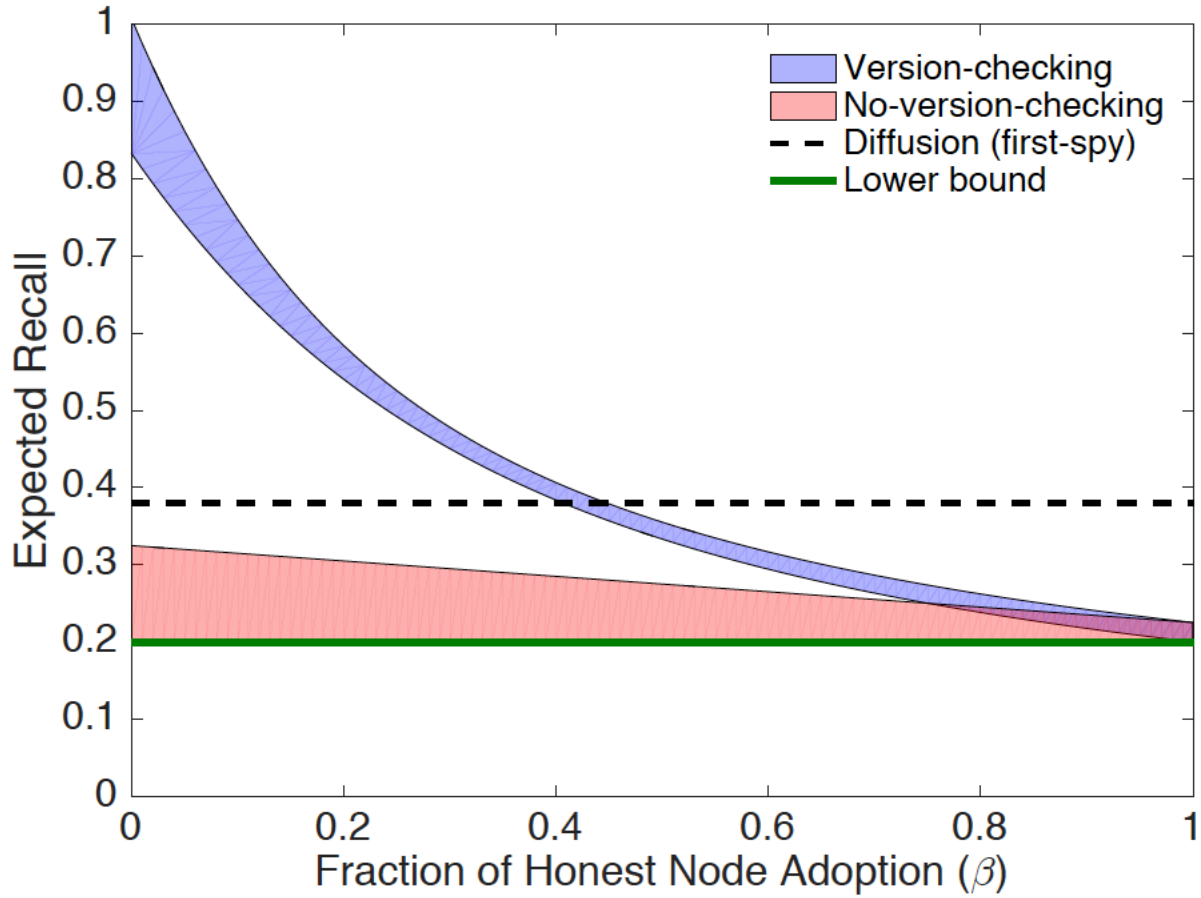


1) Messages propagate over the **same** cycle graph

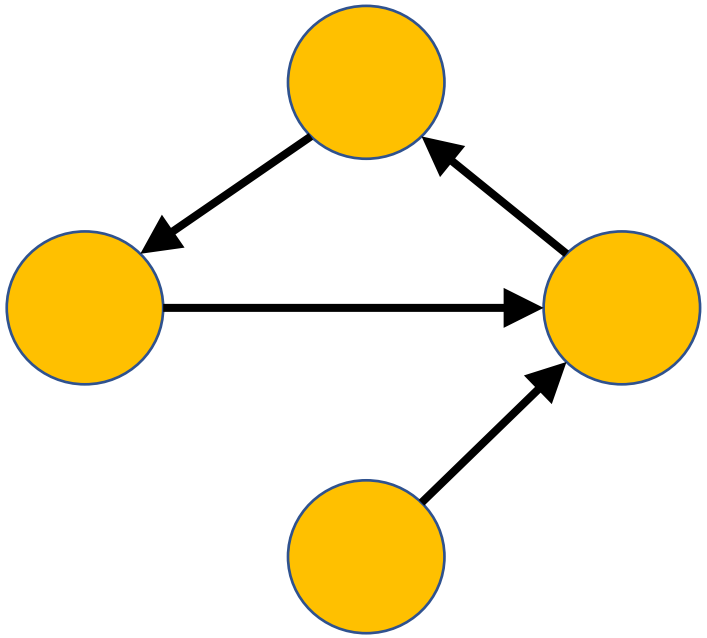
2) Anonymity graph changes dynamically.

3) No encryption required.

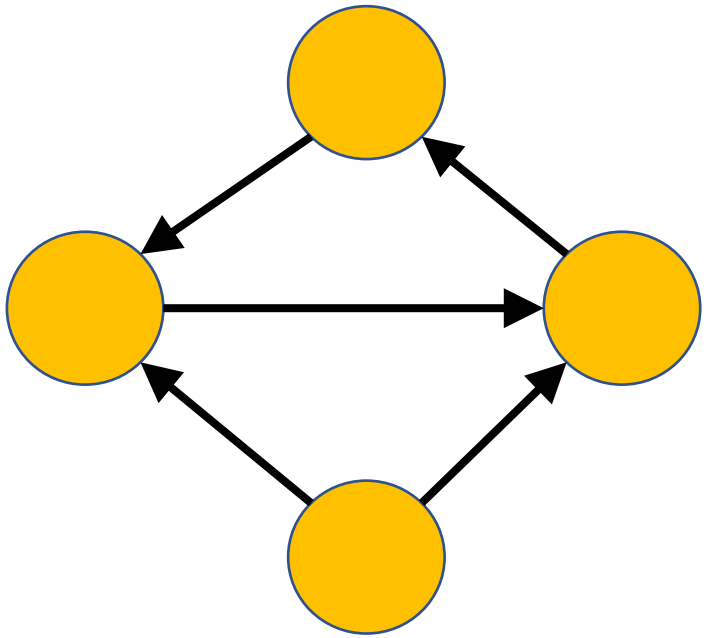

Upper bound
(Known graph)
Lower bound
(Known graph)
Line (unknown)
Line (known)
4-reg (unknown)
4-reg (known)
d-regular graphs give robustness!
Lower bound
(Unknown graph)
Precision
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
Fraction of Spies


Version-checking
No-version-checking
Diffusion (first-spy)
Lower bound
Expected Recall
Fraction of Honest Node Adoption (β)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.2
0.4
0.6
0.8
1

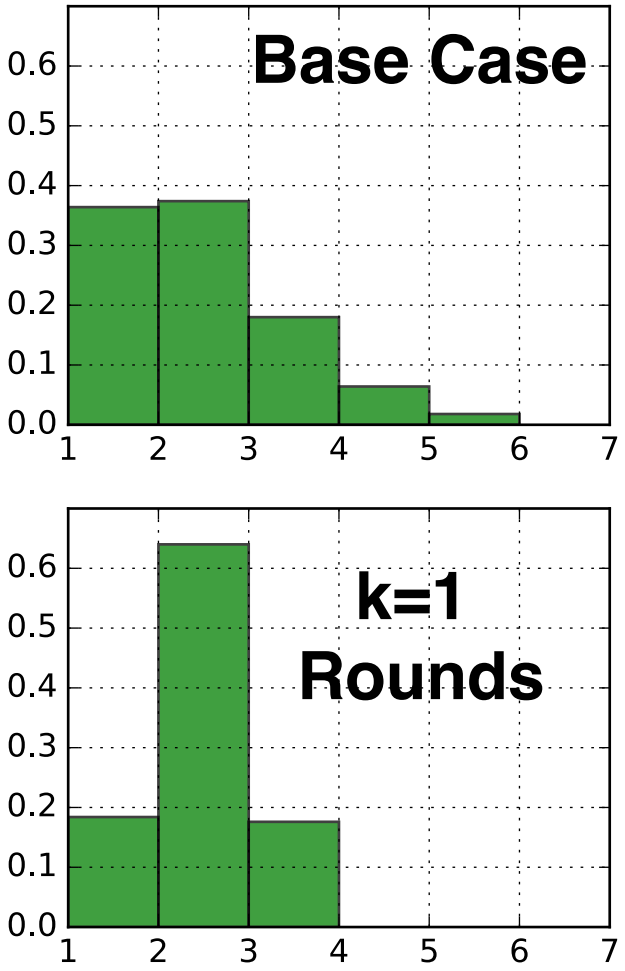
Anonymity graph construction



Base Case

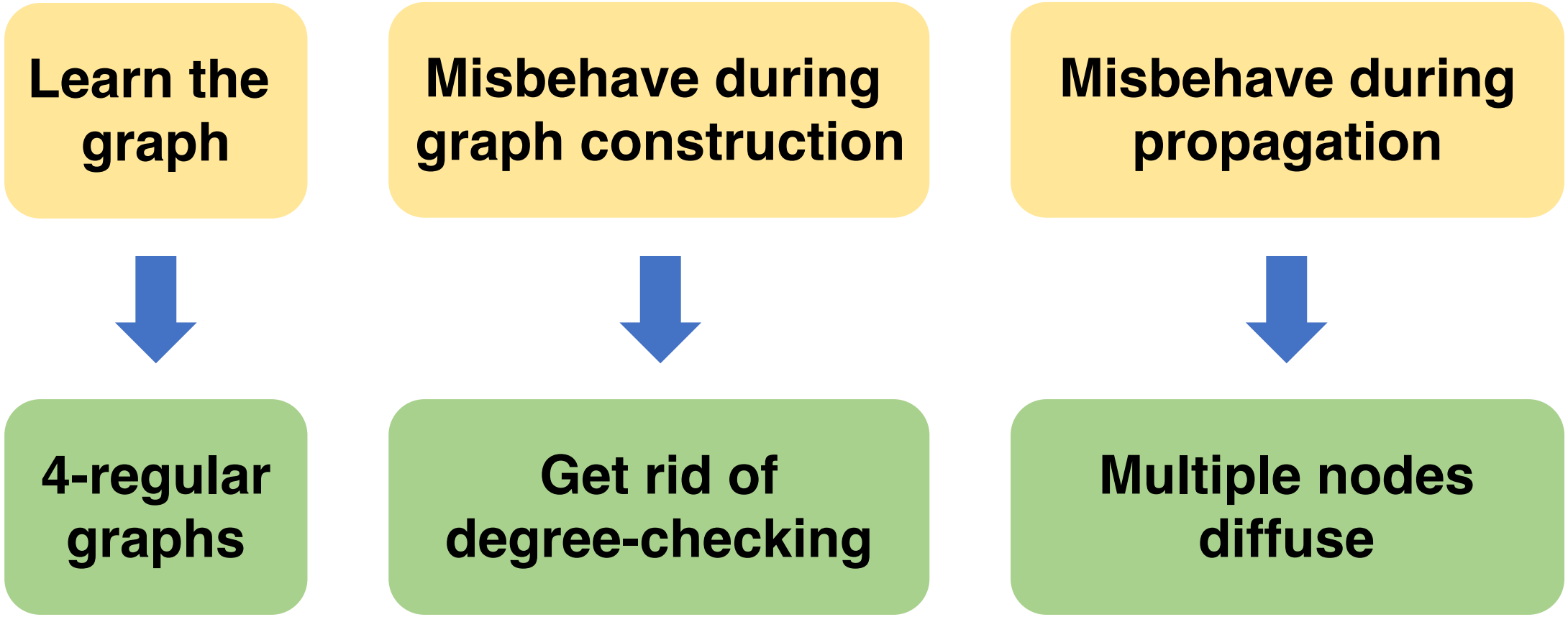


k=1 rounds of Degree-Checking



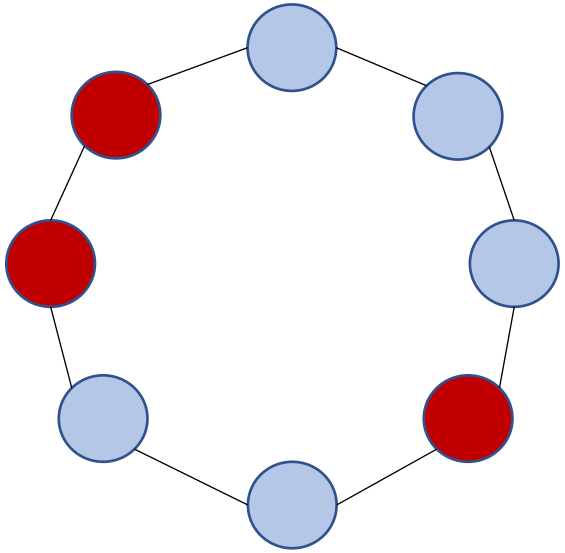
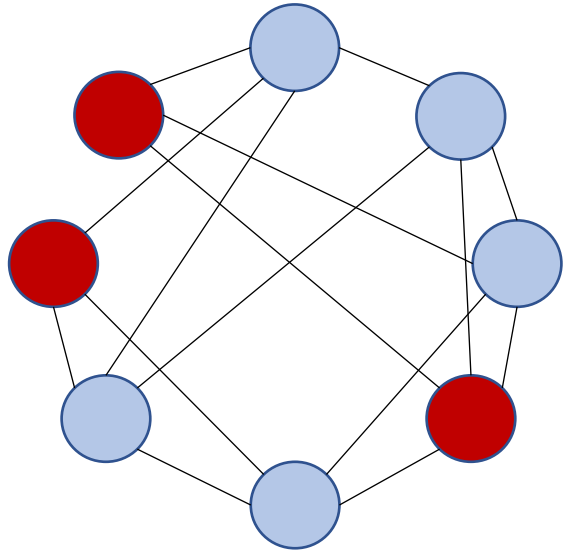
Degree

Dealing with stronger adversaries

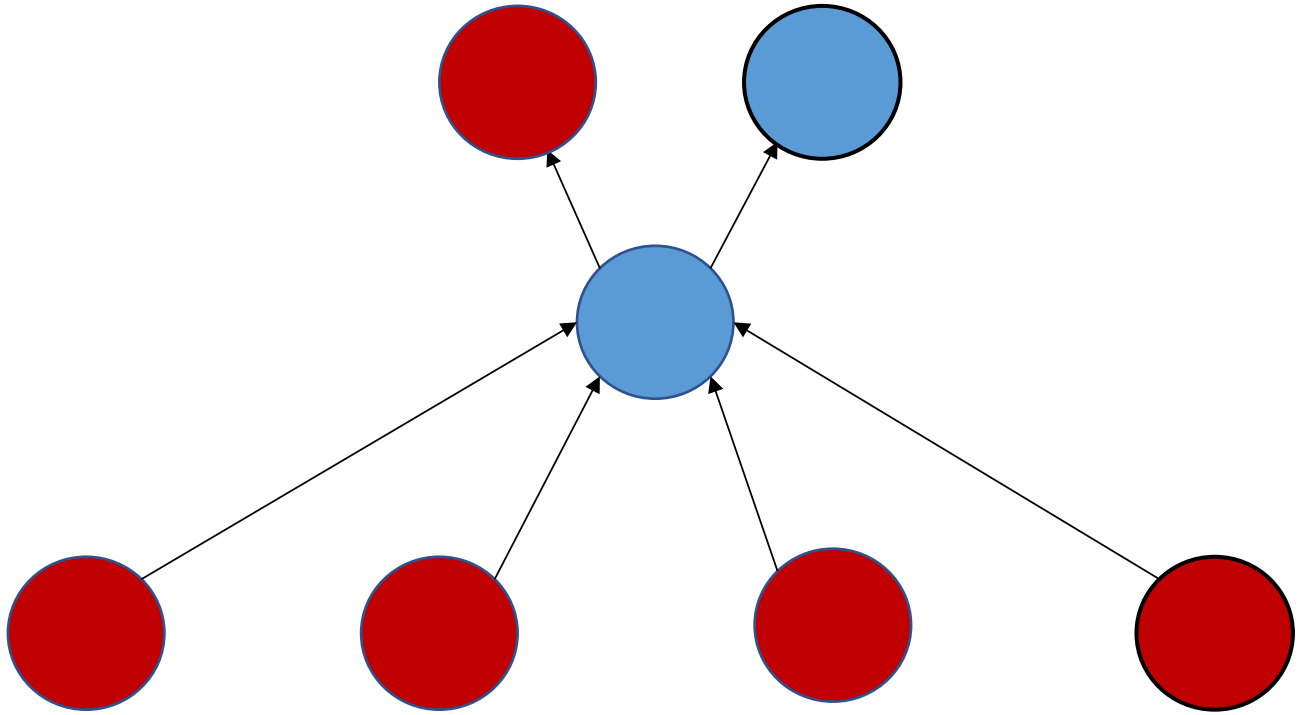


Learning the anonymity graph

Precision

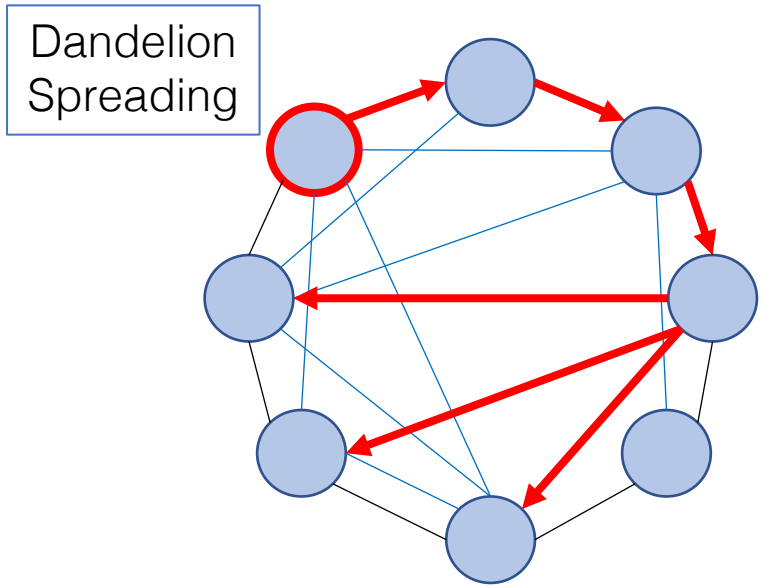
	Line	Random regular
		
Graph unknown	$O\left(\mathrm{p}^2 \log\left(\frac{1}{p}\right)\right)$	?
Graph known	$\Omega(p)$	

Manipulating the anonymity graph



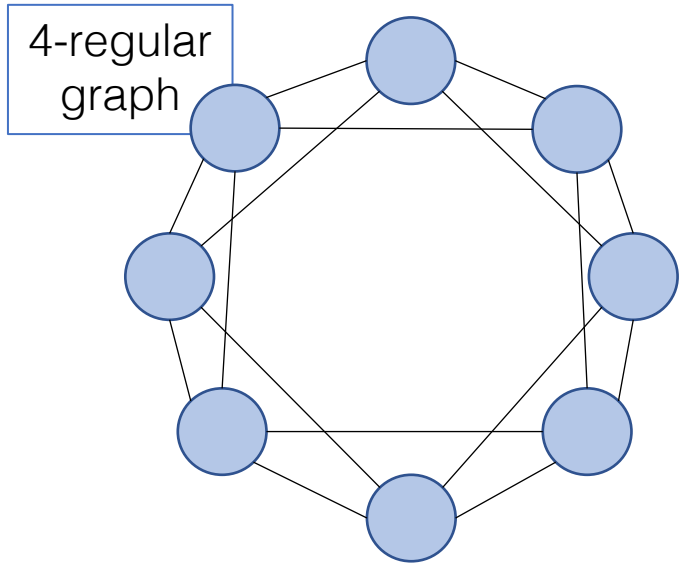
DANDELION++ Network Policy

Spreading Protocol



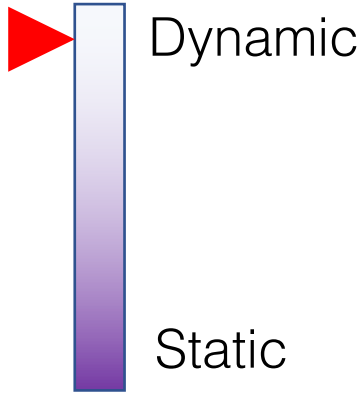
Given a graph, how do we spread content?

Topology



What is the anonymity graph topology?

Dynamicity



How often does the graph change?