IEEE Symposium on Security & Privacy

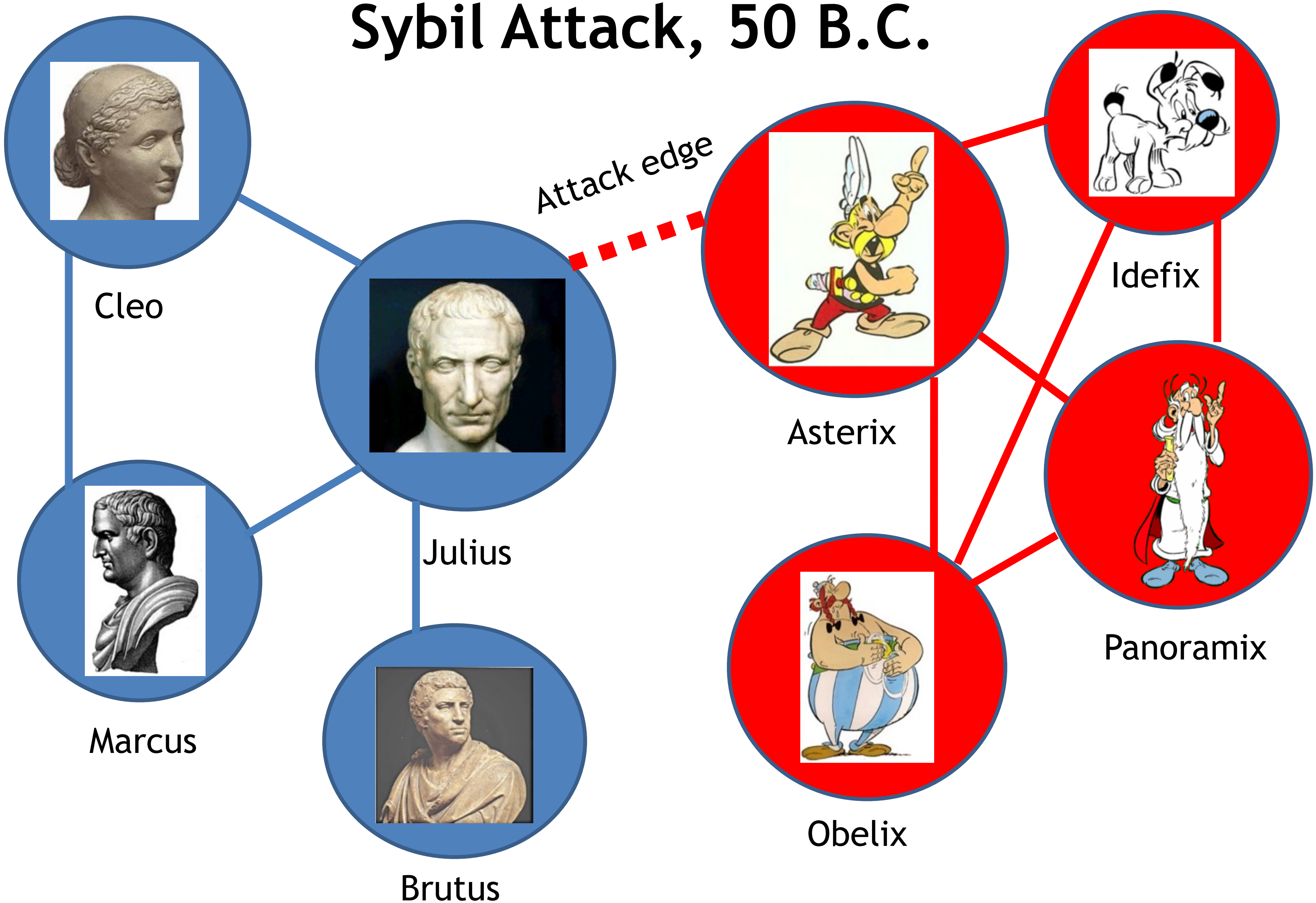
SAN FRANCISCO – 21ST MAY 2013

SoK: The Evolution of Sybil Defense via Social Networks

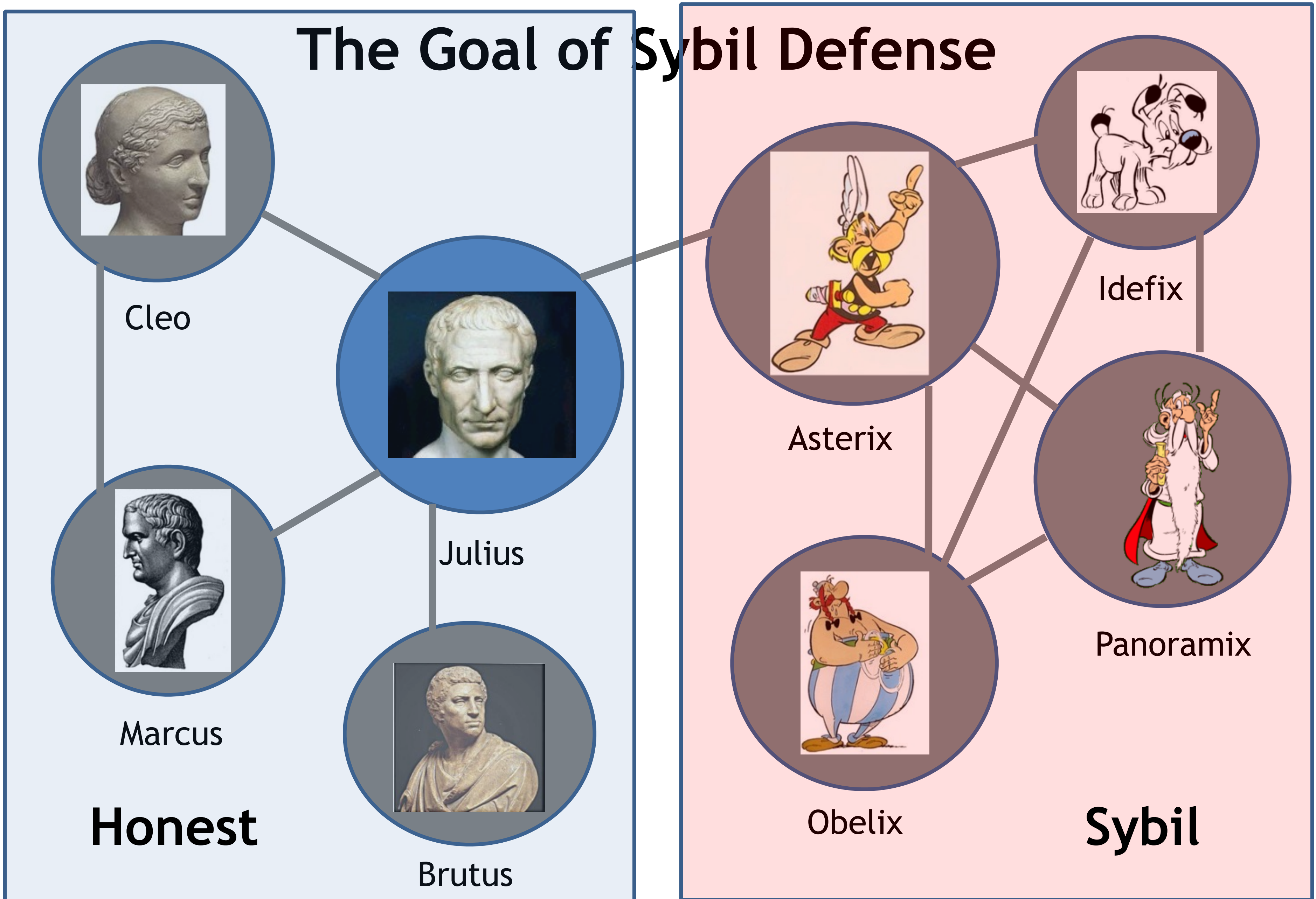
Alessandro Epasto[1]
joint work with L. Alvisi[2], A. Clement[3], S. Lattanzi[4], A. Panconesi[1]

Sapienza U. Rome[1], U. T. Austin[2], MPI-SWS[3], Google Reseach[4]

Sybil Attack, 50 B.C.



The Goal of Sybil Defense



Motivation

- Fundamental security issue in **any** open system.
- Real impact:
  - **>500k** sybils in RenRen.
  - Manual checking is expensive (Tuenti).

Social Sybil Defense

- **Key idea**: leverage social structure
  - Friendship is hard to fake!

Our contributions

- A perspective on the past of social sybil defense
  - Unifies two distinct trends
    - Random-walk based methods
    - Community detection
- A program for the future of sybil defense
  - All sybil defense is local
- A concrete first step on the new road
  - First community detection algorithm with provable sybil defense guarantees

How can we leverage the structure of the social graph?

A thought experiment

- Given a social network, is it under sybil attack?
- Which property to use?

~~Small world phenomena~~

~~Clustering coefficient~~

~~Popularity distribution~~

Conductance

Conductance

- Conductance measures how **well connected** a graph is.
- (Intuitively) A graph has **high conductance** only if there are no sets of nodes sparsely connected with the rest of the graph.
- Our analysis shows that conductance is by far the most **resilient** property

Why random walks?

- **Conductance** is intimately related to the intuitive concept of **mixing time:**
  - (Roughly): length of random walk to hit truly random node.

- **Fast mixing** networks (mixing time is $O(\log(n))$

- Further justification of random walk approach introduced by Yu et al. (2006).

Random walk based defenses

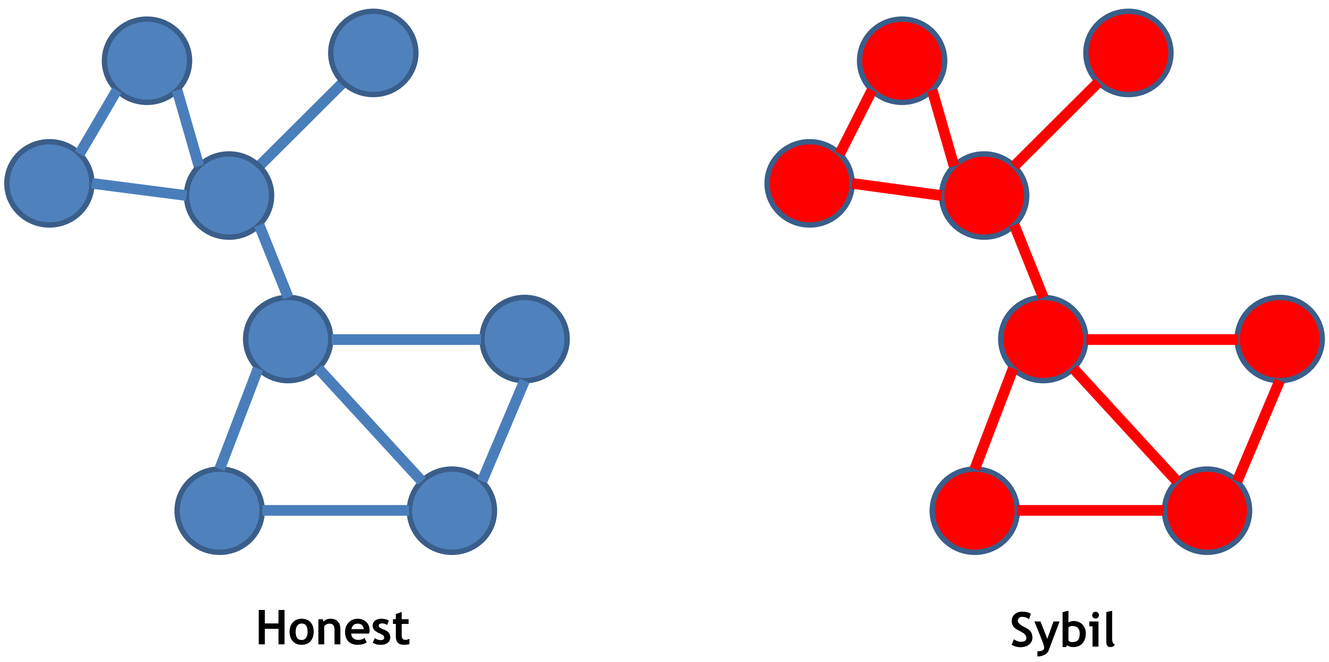
- **Many state of the art solutions use random walks:**
  - SybilGuard, Yu et al., SIGCOMM 2006
  - SybilLimit, Yu et al., SP 2008
  - SybilInfer, Danezis et al., NSDD 2006
  - SybilRank, Cao et al, NSDI 2012

- **Our contribution:** A unified view of these techniques based on random walk theory.

Random Walks: the intuition

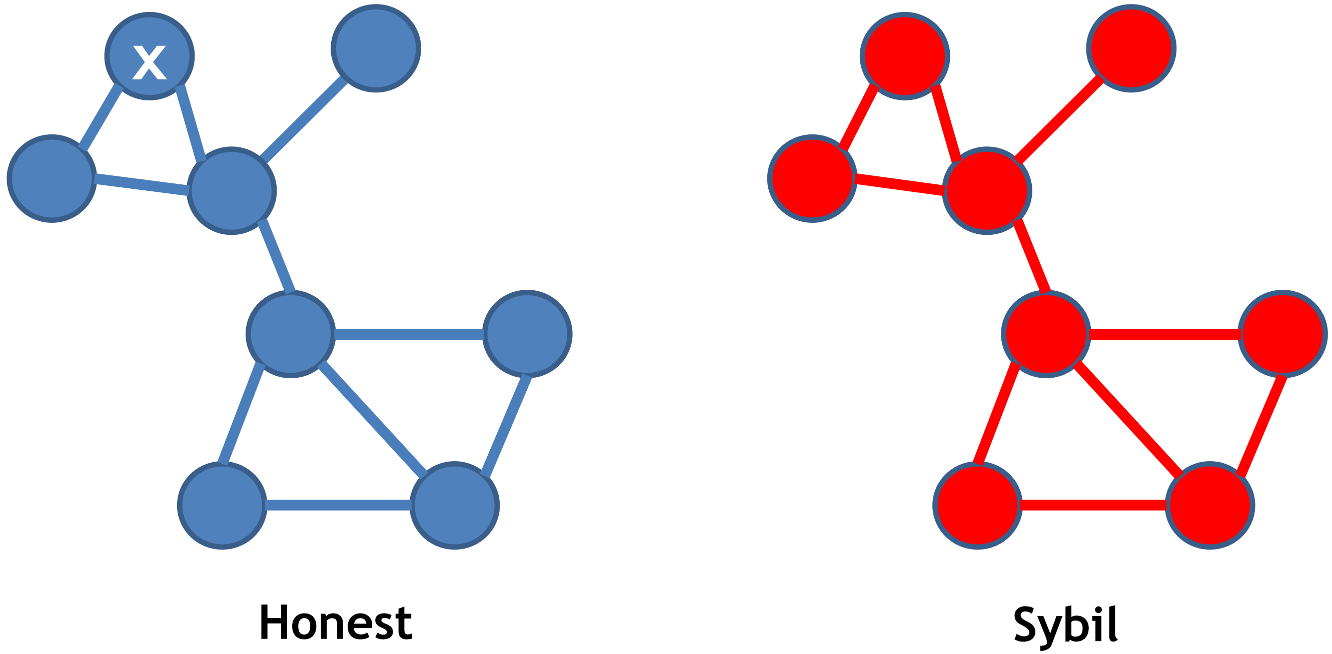
A toy problem

- Consider the following simplified problem:
  - Two disjoint graphs. **No** attack edges.



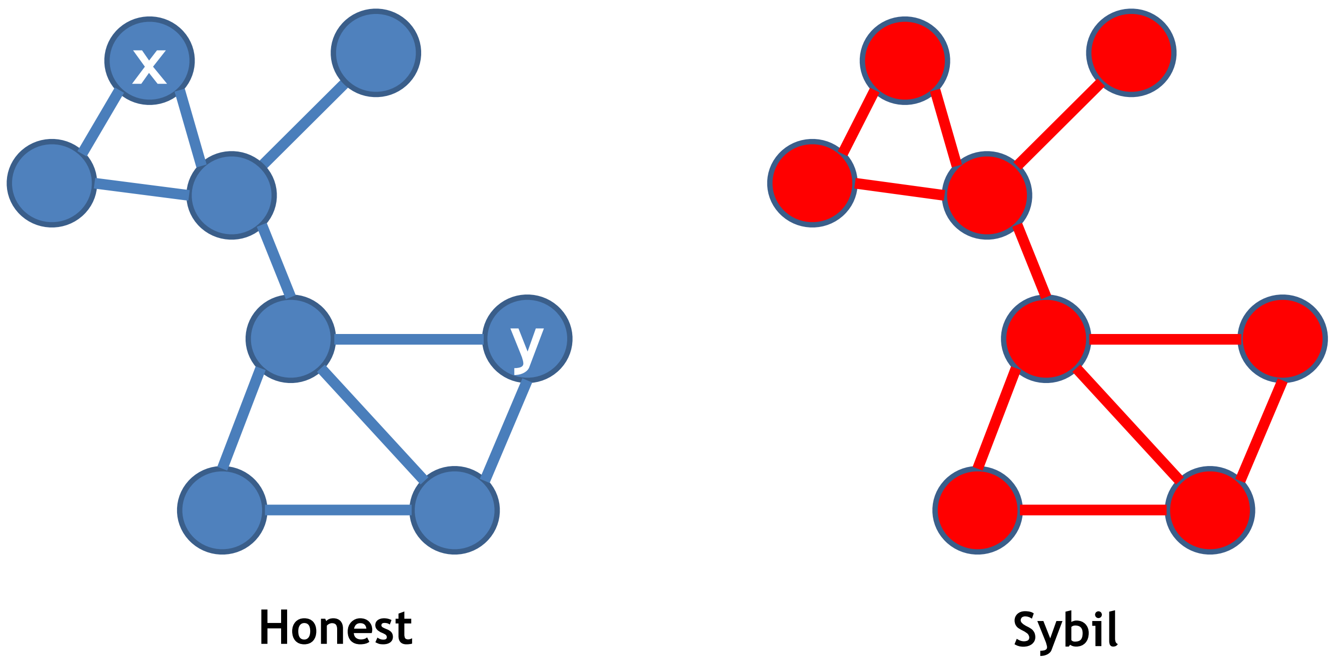
A toy problem

- Consider the following simplified problem:
  - Two disjoint graphs. **No** attack edges.
- How can a node decide who to trust in a distributed way?



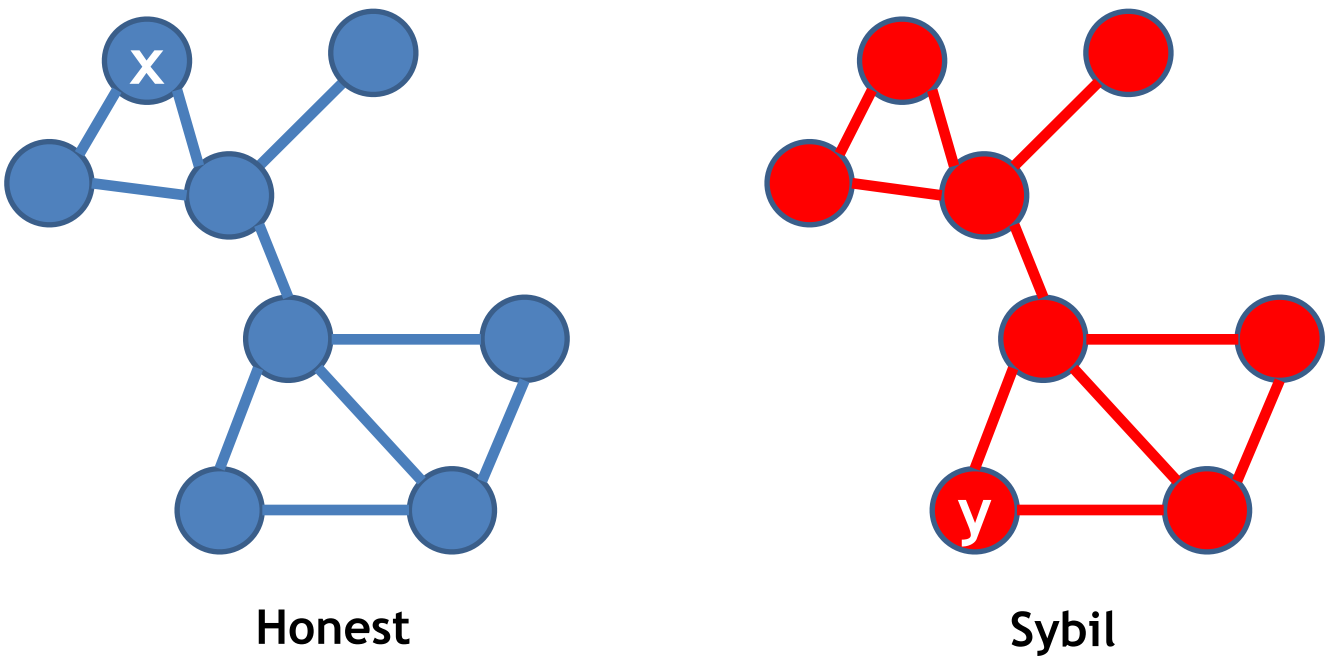
A toy problem

- Consider the following simplified problem:
  - Two disjoint graphs. **No** attack edges.
- How can a node decide who to trust in a distributed way?



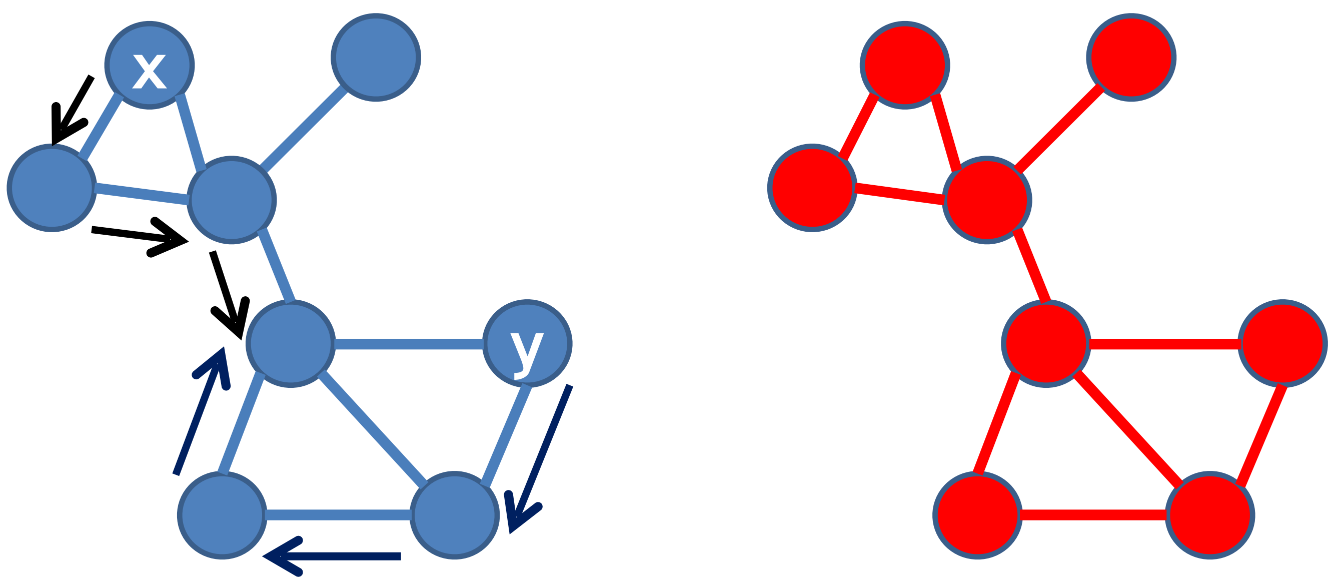
A toy problem

- Consider the following simplified problem:
  - Two disjoint graphs. **No** attack edges.
- How can a node decide who to trust in a distributed way?



Random walks

- Intuition: perform a **random walk** from each node
- Two node trust each other if there is any intersection.



Honest

Sybil

Properties of the protocol

- **Safety:** sybil nodes are never accepted
- **Liveness:** boost probability of accepting honest nodes by using many random walks (still computationally efficient)

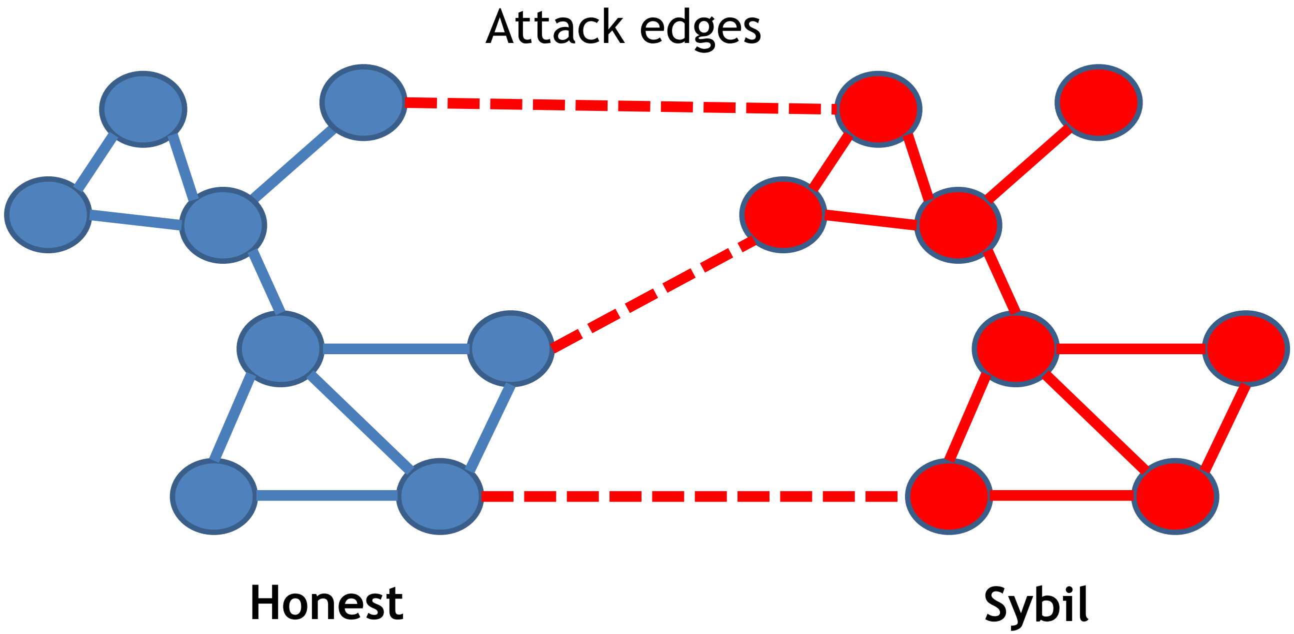
Implementation of the protocol

- How long should this random walk be?
  - As short as possible
  - Cover uniformly the honest region

- The answer is the **mixing time**, $O(\log(n))$ if the graph is **fast mixing**

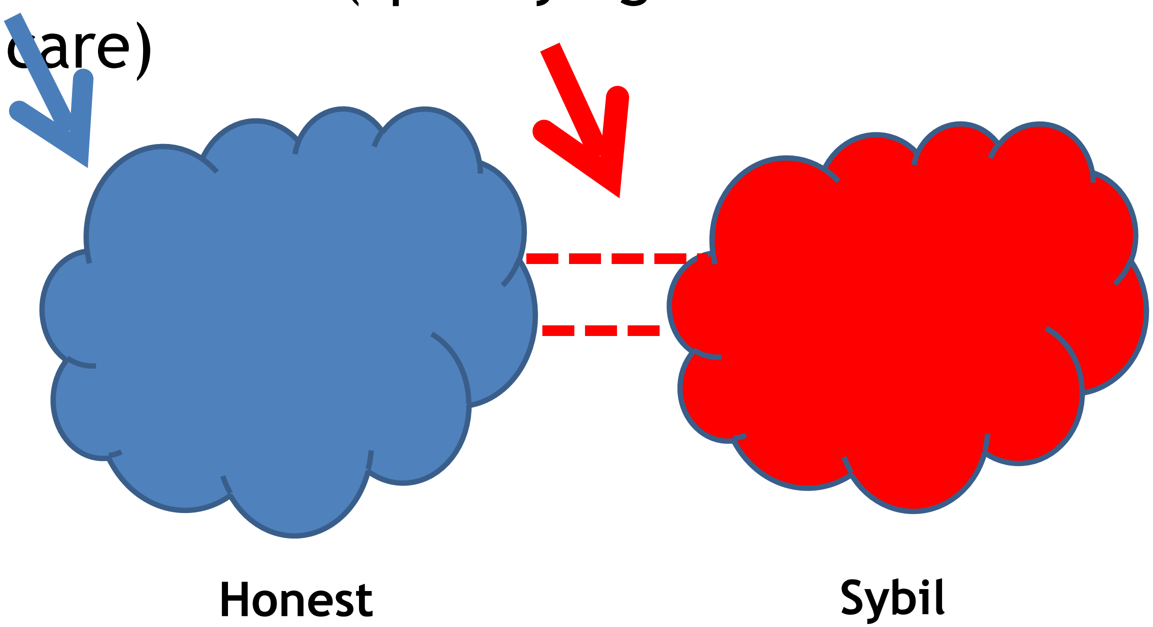
Back to the real world

- The two graphs are not disjoint.
- With **few** attack edges and **short** walks it still works.
- **Note:** Precise theoretical guarantees are based on **conductance.**



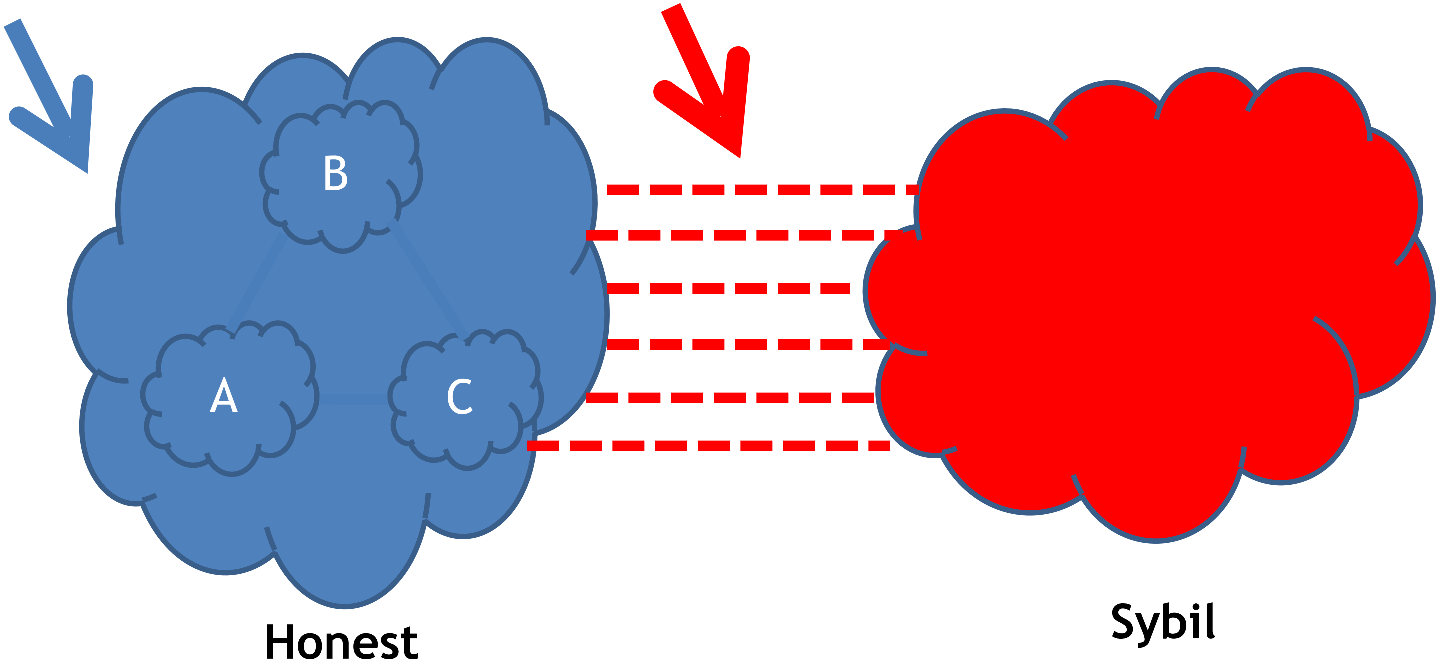
Central assumptions

- The method works provided that two assumptions are met:
  1. **Sparse cut** between honest and sybils;
  2. The honest region is **fast mixing.**
- Then: it works (specifying in which sense requires some care)



However…

The two assumptions do not hold



The cut is not as sparse as assumed (Bilge et al. WWW 200

The honest region is not fast mixing (Mohaisen, et al. IMC 2

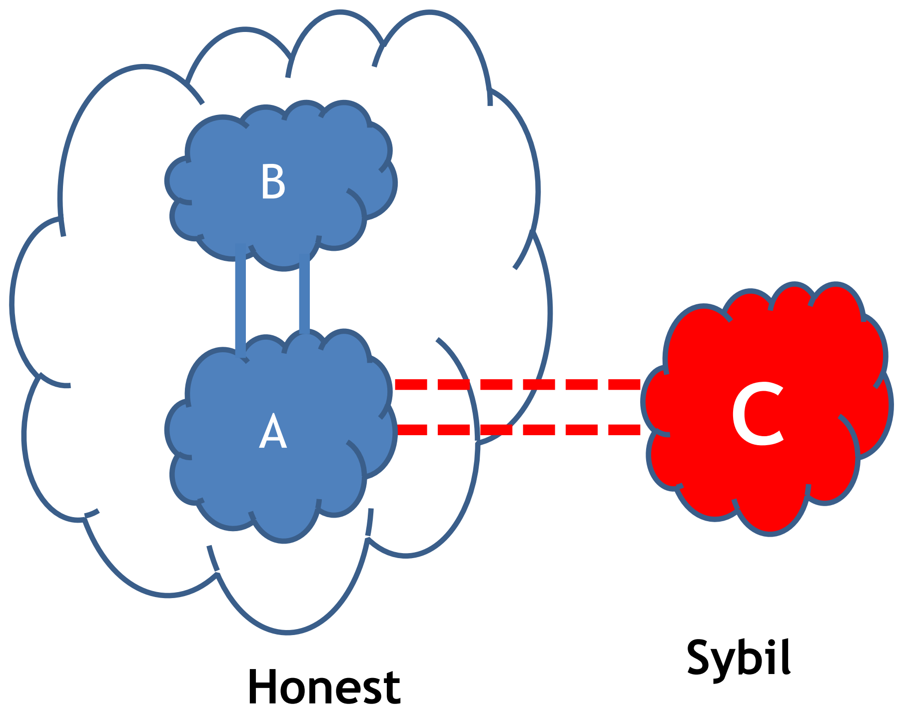
Global sybil defense is unrealistic

Traditional sybil defense depends
on
assumptions that are **too strong…**

What can we *realistically* do?

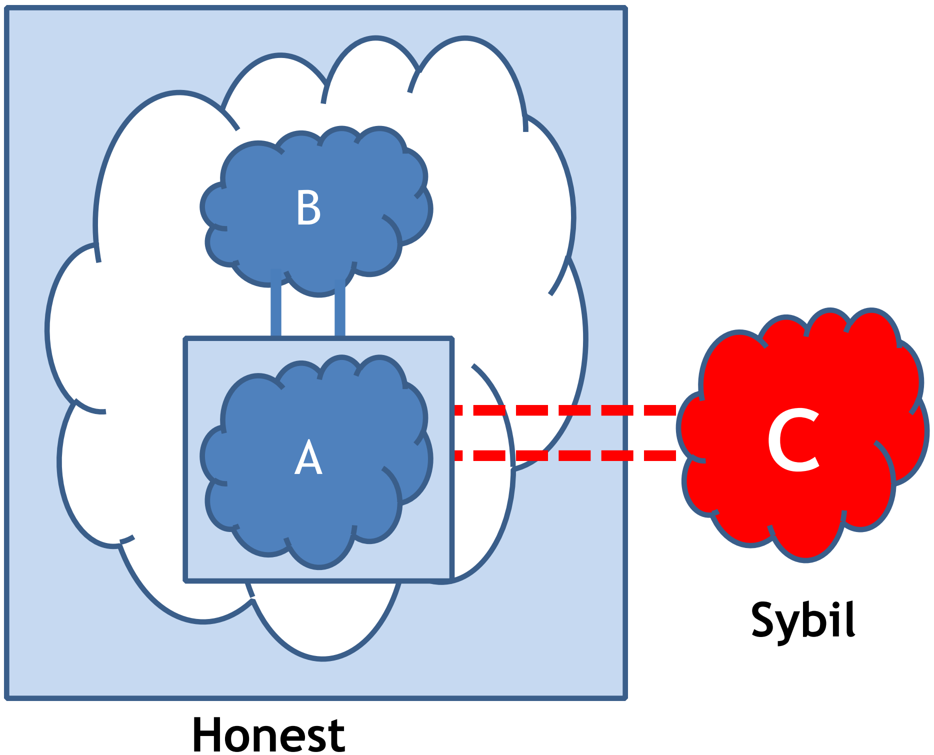
From global to local sybil defense

Sybil defense in real networks



- **A** can not distinguish between **B** and **C**

A new goal for sybil defense



- White-list the nodes in **A**'s community
  - Practically useful
  - Attainable

Sybil Defense & Community Detection

- Sybil defense as community detection (Viswanath et. al, SIGCOMM 2010).
  - Must identify correct and sybil communities
- … but with no provable guarantees!

Our contribution:

A community detection algorithm
with provable sybil defense guarantees

- The keys once again are **conductance** and **random walks**

Random Walks Revisited: ACL

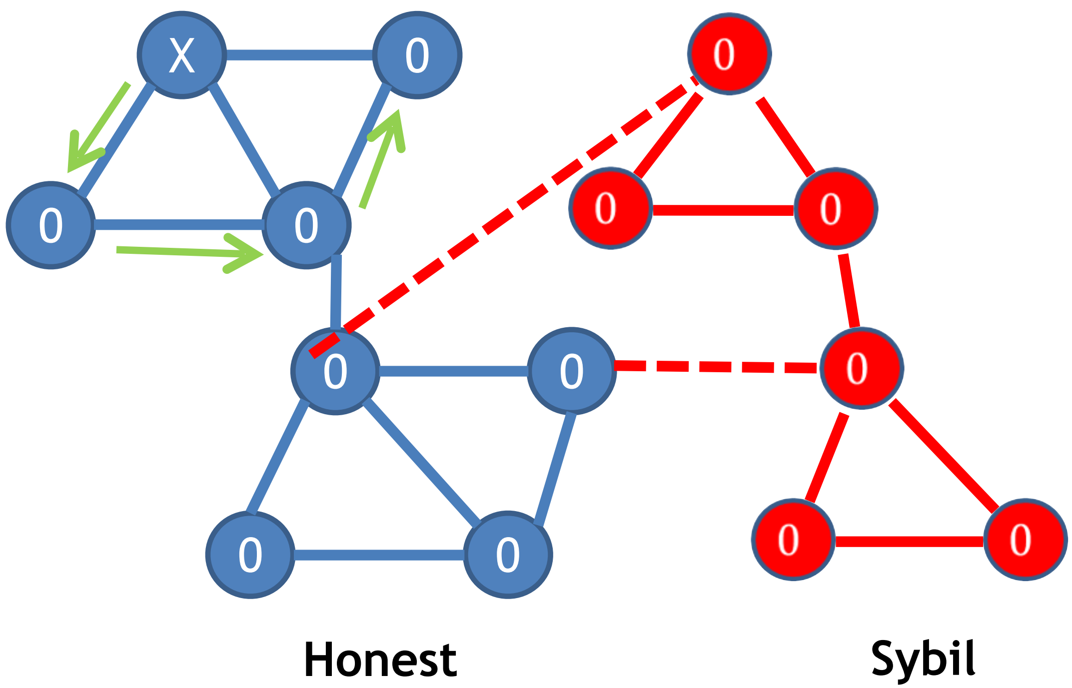
- How to find the community of given node?
  - Random walks with a bias on the community of the seed
  - Assign higher score to nodes inside the community

- Leverage community detection literature:
  - **ACL** (Andersen, et al. 2006)
    - Provable sybil defense guarantees.

Random Walks Revisited: ACL

- **Personalized PageRank:** variable length random walks

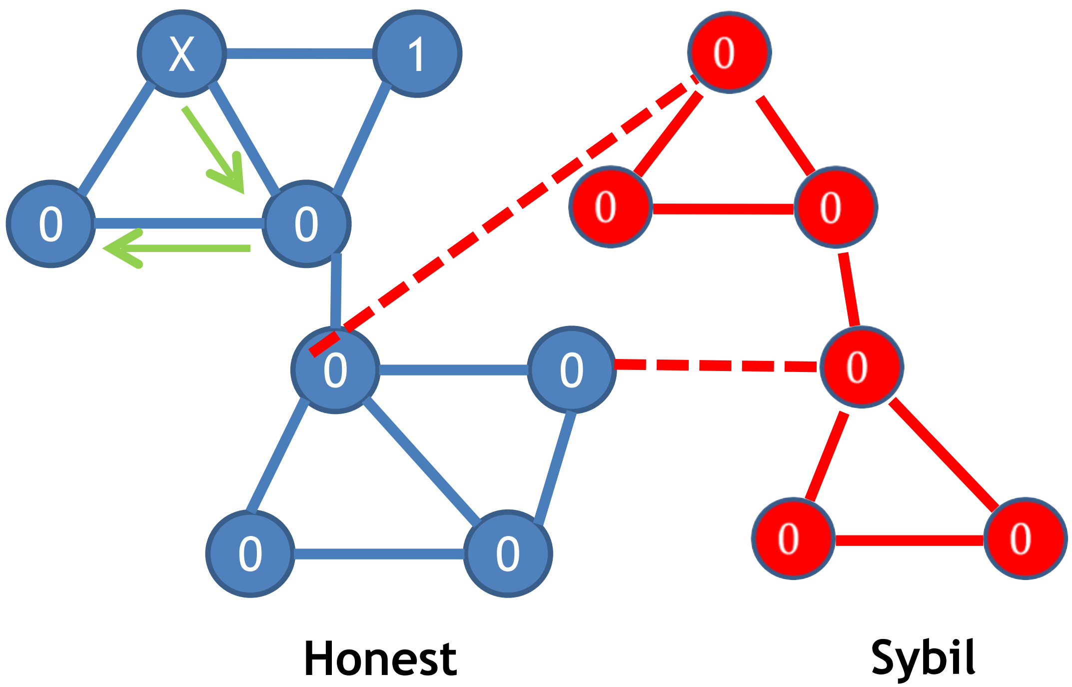
3 Steps



Random Walks Revisited: ACL

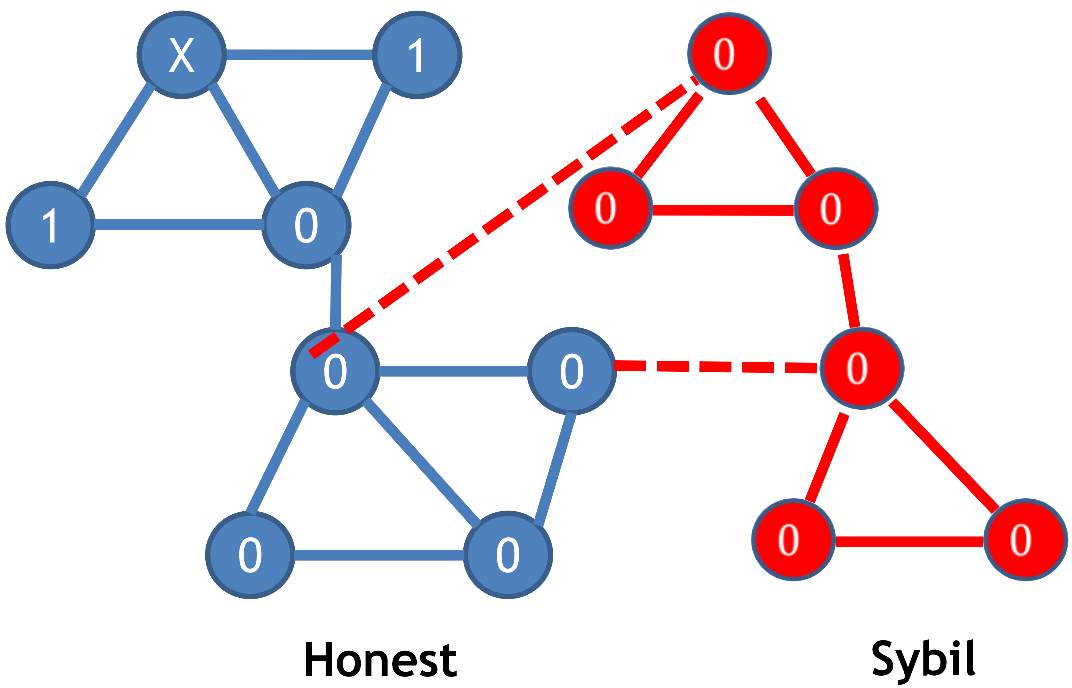
- **Personalized PageRank:** variable length random walks

2 Steps



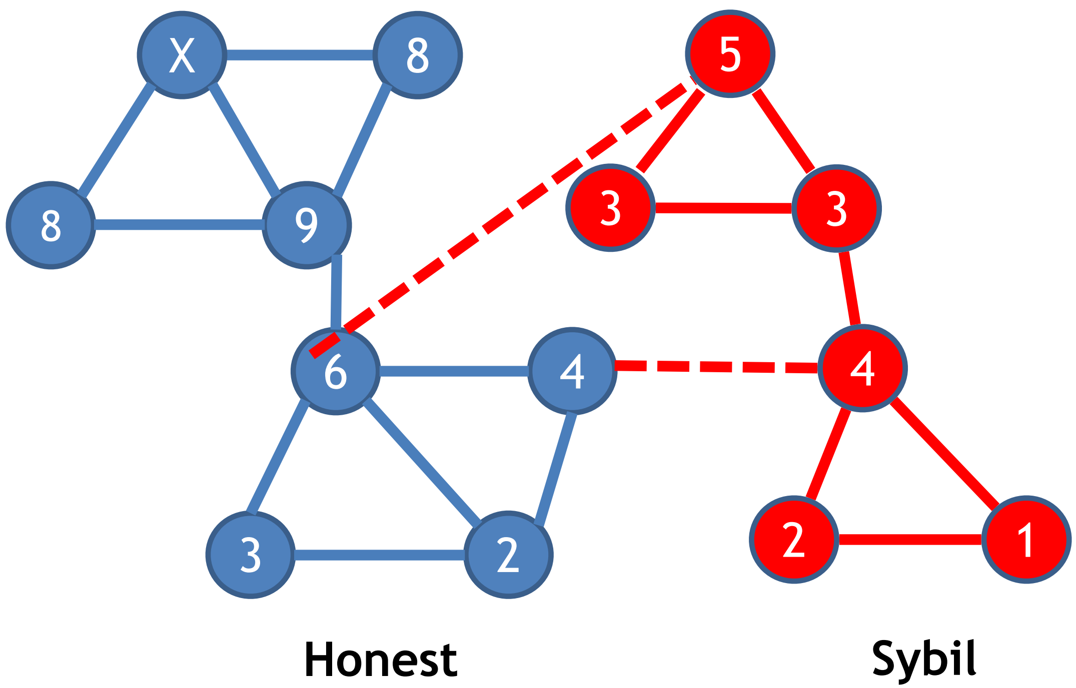
Random Walks Revisited: ACL

- **Personalized PageRank:** variable length random walks
  - After **many** walks…



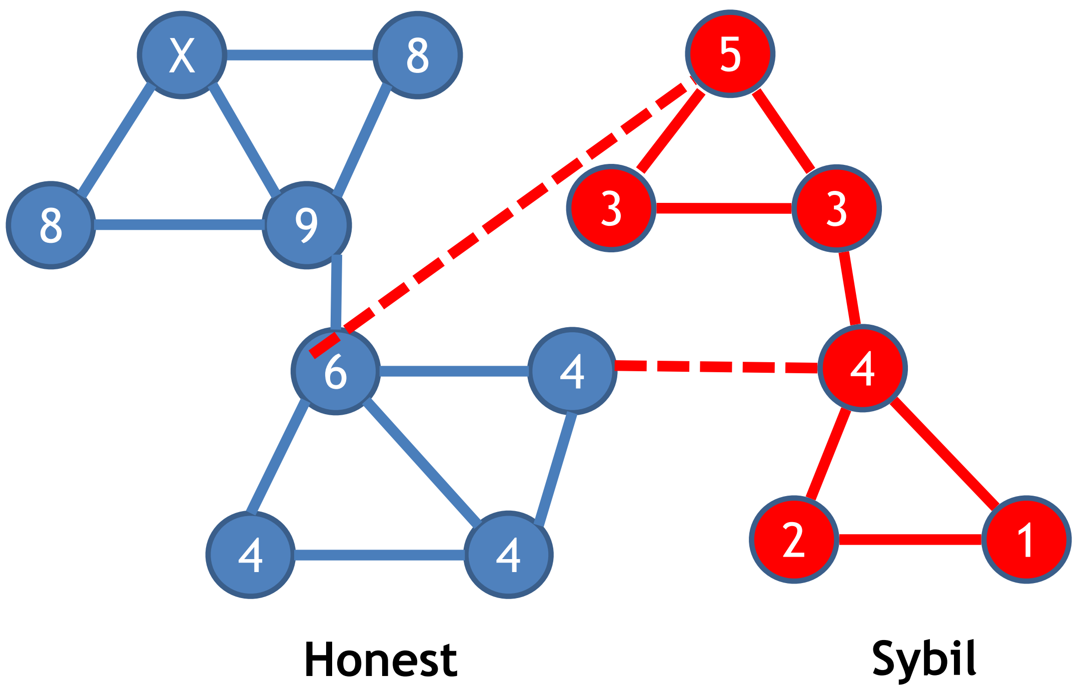
Random Walks Revisited: ACL

- **Personalized PageRank:** variable length random walks
  - After **many** walks…
  - Node's score = how frequently node is visited



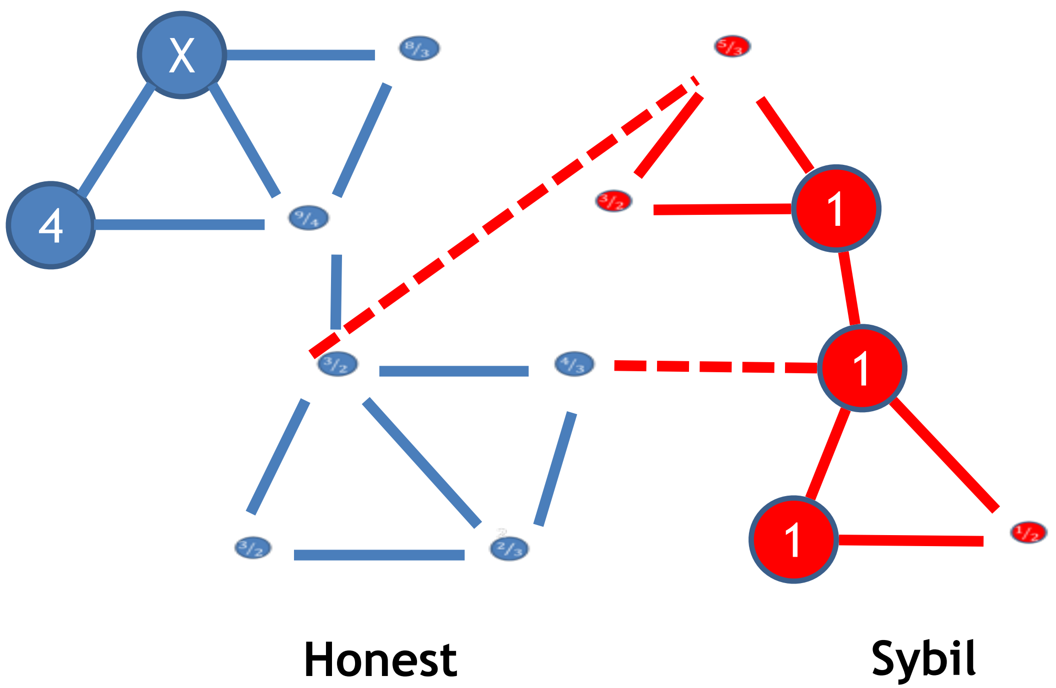
Random Walks Revisited: ACL

- **High degree** nodes can achieve disproportionate score



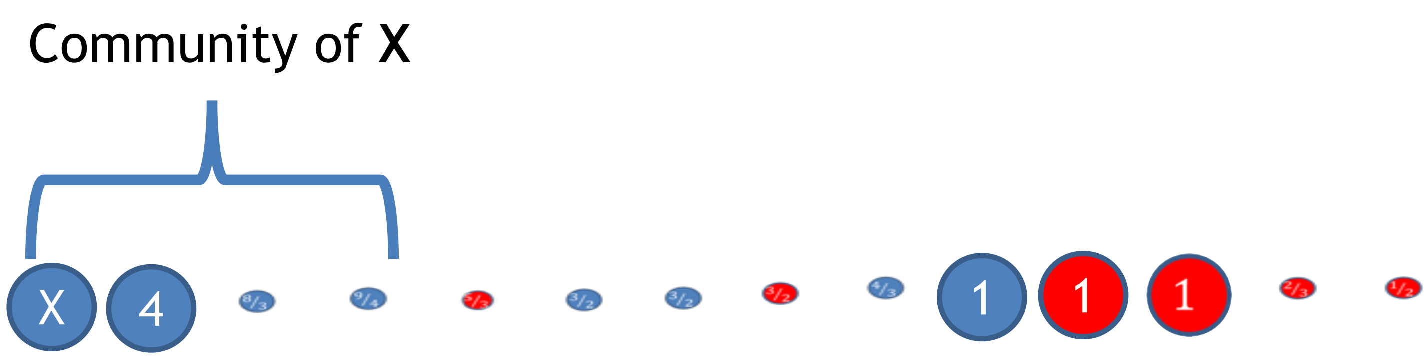
Random Walks Revisited: ACL

- **High degree** nodes can achieve disproportionate score
- Node’s trustworthiness = score normalized by degree



Random Walks Revisited: ACL

- Nodes are ranked by their trustworthiness
- Ranking has **strong bias** on the seed's community



The Guarantee

- The intuition can be formalized in a **theorem**:

> Select a u.a.r. honest node in a fast mixing community
> **C** with fewer than o(n/log(n)) attack edges:
>
> The **ACL** ranking contains 1-o(1) honest nodes in the first **|C|** positions.

- We confirm this result with an **experimental** evalutation.

Experimental evaluation

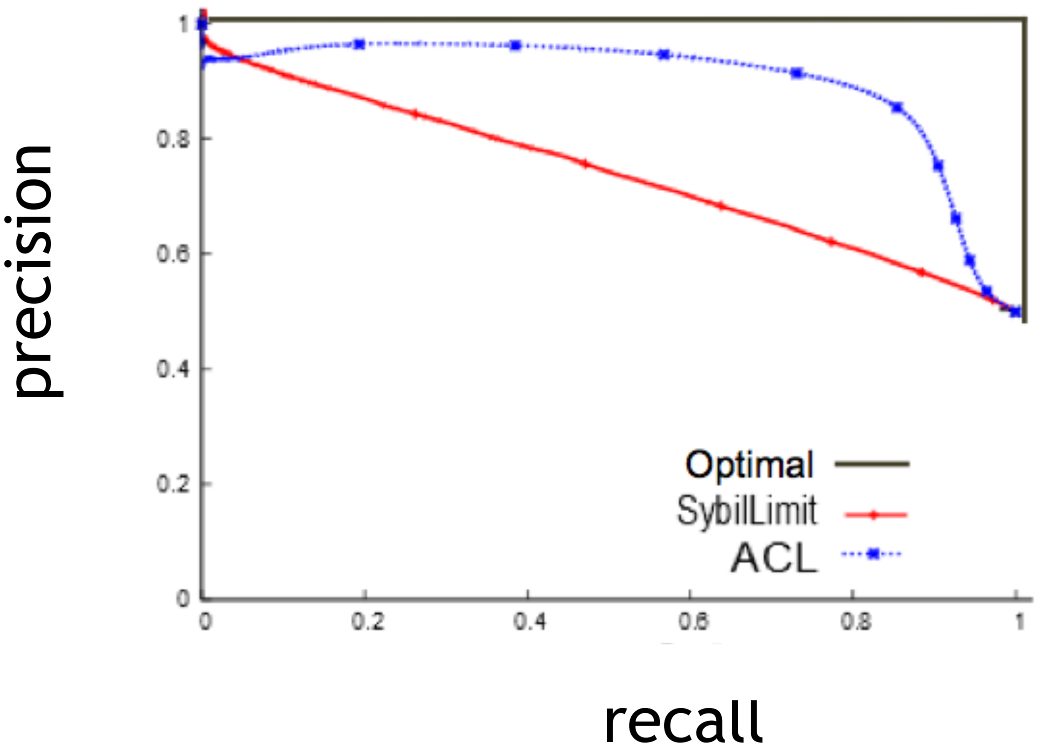
- We compared the performance of **ACL** with several state-of-the-art algorithms: **SybilGuard**, **SybilLimit**, **Gatekeeper** and **Mislove's** community detection algorithm.

- Attack models:
  - Traditional attack model (Danezis et al., NSDD 2006)
  - New attack model with interesting theoretical properties

- The results were **consistent** across the different models and datasets.

Performance

Precision vs Recall in Facebook (new attack model)



Facebook (New Orleans)
Viswanath et al. 2009

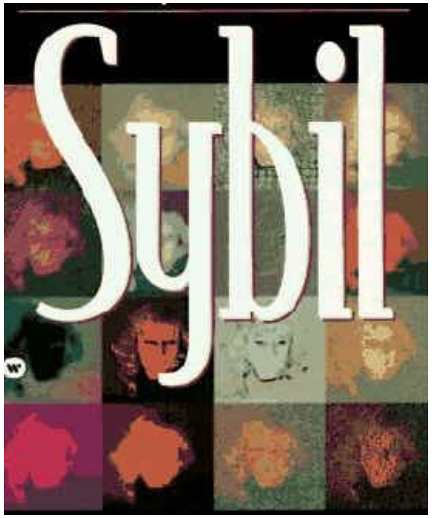
Nodes: 63k
Edges: 816k

ACL vs SybilLimit

Similar results are obtained in all our datasets

Conclusions

- Unified view of social network based sybil defense: random walks and community detection
- New goal for sybil defense
- Community detection can provide secure sybil defense schemes.



Thank you for your attention