An analysis of Social Network-based Sybil defenses

Bimal Viswanath[§] Ansley Post[§]

Krishna Gummadi[§] Alan Mislove[¶]

[§]MPI-SWS [¶]Northeastern University

SIGCOMM 2010

Sybil attack

Fundamental problem in distributed systems

Attacker creates many fake identities (Sybils)

- Used to manipulate the system

Many online services vulnerable

- Webmail, social networks, p2p

Several observed instances of Sybil attacks

- Ex. Content voting tampered on YouTube, Digg

Sybil attack

Fundamental problem in distributed systems

Attacker creates many fake identities (Sybils)

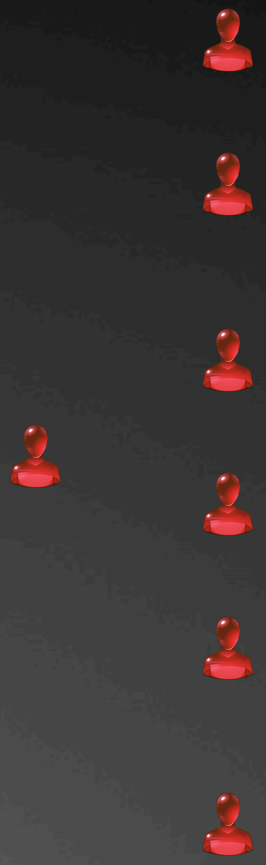
- Used to manipulate the system

Many online services vulnerable

- Webmail, social networks, p2p

Several observed instances of Sybil attacks

- Ex. Content voting tampered on YouTube, Digg



Sybil defense approaches

Tie identities to resources that are hard to forge or obtain

RESOURCE 1 Certification from trusted authorities

- Ex. Passport, social security numbers
- Users tend to resist such techniques

RESOURCE 2 Resource challenges (e.g., cryptopuzzles)

- Vulnerable to attackers with significant resources
- Ex. Botnets, renting cloud computing resources

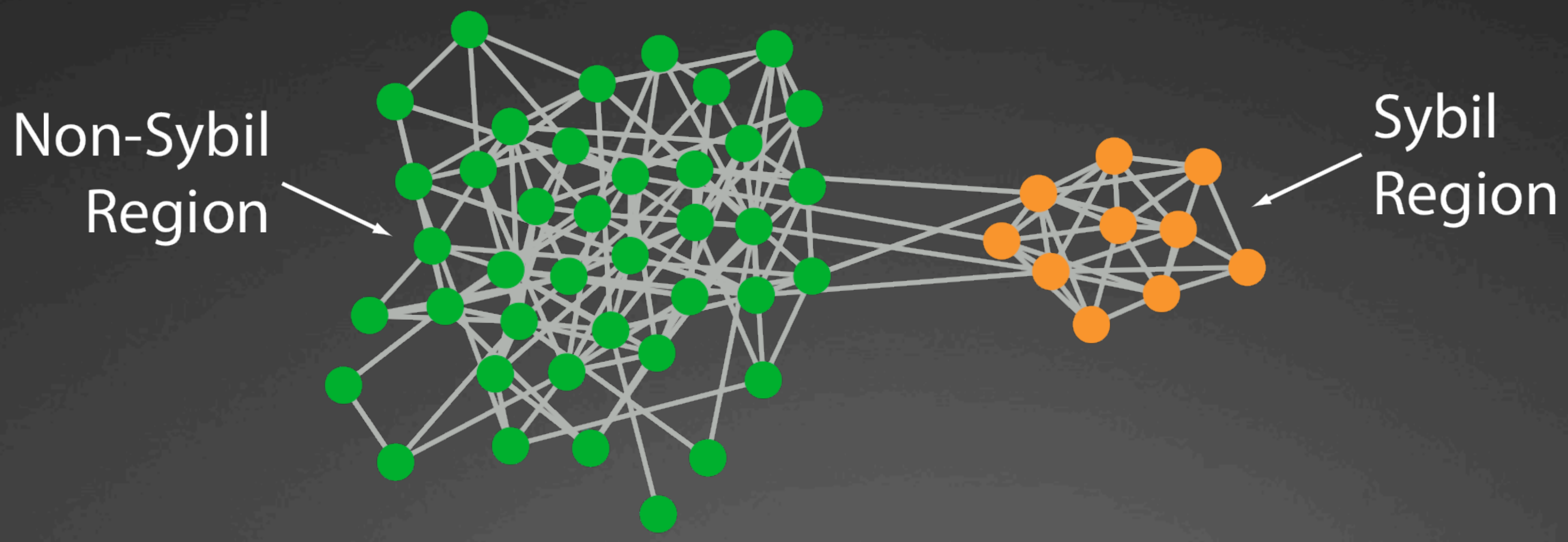
RESOURCE 3 Links in a social network?

New approach: Use social networks

Assumption: Links to good users hard to form and maintain

- Users mostly link to others they recognize

Attacker can only create limited links to non-Sybil users

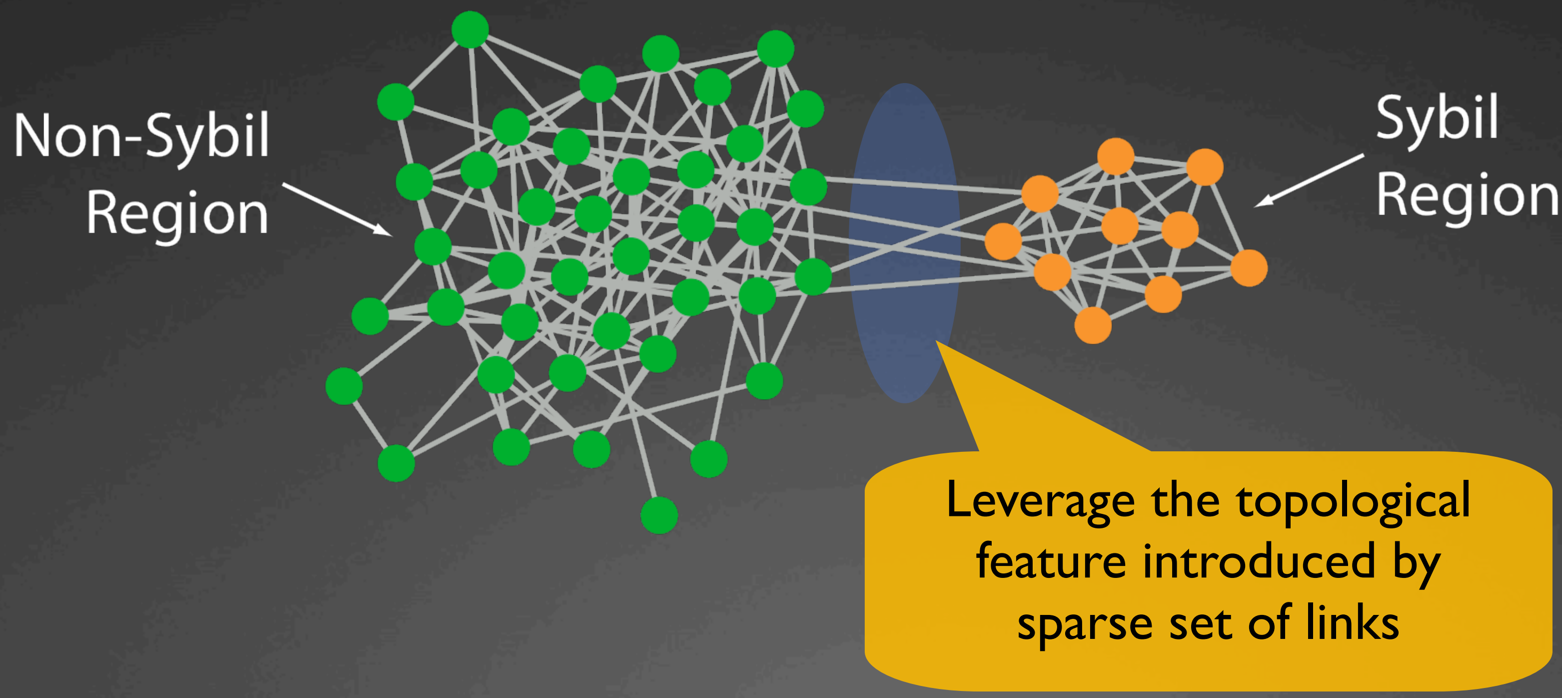


New approach: Use social networks

Assumption: Links to good users hard to form and maintain

Users mostly link to others they recognize

Attacker can only create limited links to non-Sybil users



Social network-based schemes

Social network-based schemes

Very active area of research

- Many schemes proposed over past five years

Examples:

- SybilGuard [SIGCOMM'06]

Social network-based schemes

Very active area of research

- Many schemes proposed over past five years

Examples:

- SybilGuard [SIGCOMM’06]
- SybilLimit [Oakland S&P ’08]

Social network-based schemes

Very active area of research

- Many schemes proposed over past five years

Examples:

- SybilGuard [SIGCOMM’06]
- SybilLimit [Oakland S&P ’08]
- SybilInfer [NDSS’08]

Social network-based schemes

Very active area of research
- Many schemes proposed over past five years

Examples:
- SybilGuard [SIGCOMM'06]
- SybilLimit [Oakland S&P '08]
- SybilInfer [NDSS'08]
- SumUp [NSDI'09]

Social network-based schemes

Very active area of research

- Many schemes proposed over past five years

Examples:

- SybilGuard [SIGCOMM'06]
- SybilLimit [Oakland S&P '08]
- SybilInfer [NDSS'08]
- SumUp [NSDI'09]
- Whanau [NSDI'10]

Social network-based schemes

- Very active area of research
  - Many schemes proposed over past five years

- Examples:
  - SybilGuard [SIGCOMM'06]
  - SybilLimit [Oakland S&P '08]
  - SybilInfer [NDSS'08]
  - SumUp [NSDI'09]
  - Whanau [NSDI'10]
  - MOBID [INFOCOM'10]

But, many unanswered questions

All schemes make same assumptions
- Use only social network

But, schemes work using different mechanisms
- Unclear relationship between schemes

Is there a common insight across the schemes?
- Is there a common structural property these schemes rely on?

Understanding relationship would help
- How well would these schemes work in practice?
- Are there any fundamental limitations of Sybil defense?

This talk

Propose a methodology for comparing schemes

- Allows us to take closer look at how schemes are related

Finding: All schemes work in a similar manner

- Despite different mechanisms

Implications: Hidden dependence on network structure

- Understand the limitations of these schemes

How to compare schemes?

Straightforward approach is to implement and compare

- Treat like a black-box

But, only gives one point evaluation

- Output dependent on scheme-specific parameters

We want to understand HOW schemes choose Sybils

- Interested in underlying graph algorithm

Thus, we had to open up the black-box

- We analyze SybilGuard, SybilLimit, SumUp and SybilInfer

How do schemes work internally?

Take in a social network and trusted node

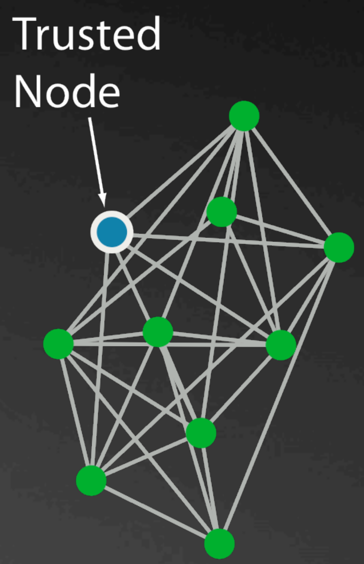
- Declare Sybils from perspective of trusted node

Internally, schemes assign probability to nodes

- Likelihood of being a Sybil

Leverage this to compare schemes?

- View schemes as inducing ranking on nodes
- Easier to compare rankings than full schemes



How do schemes work internally?

Take in a social network and trusted node

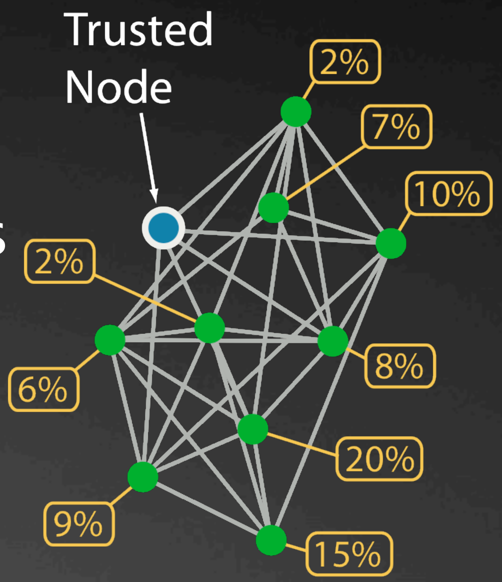
- Declare Sybils from perspective of trusted node

Internally, schemes assign probability to nodes

- Likelihood of being a Sybil

Leverage this to compare schemes?

- View schemes as inducing ranking on nodes
- Easier to compare rankings than full schemes



How do schemes work internally?

Take in a social network and trusted node

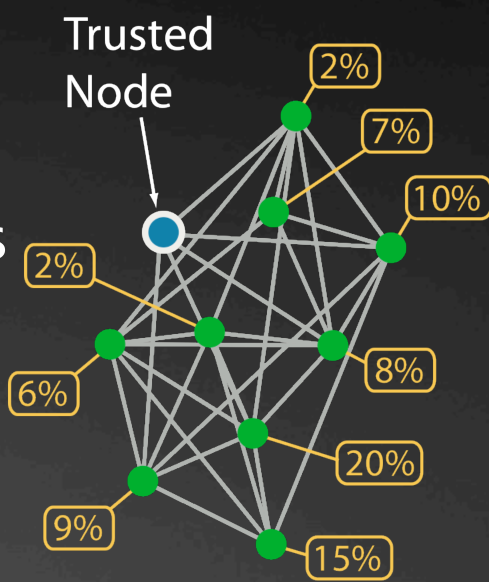
- Declare Sybils from perspective of trusted node

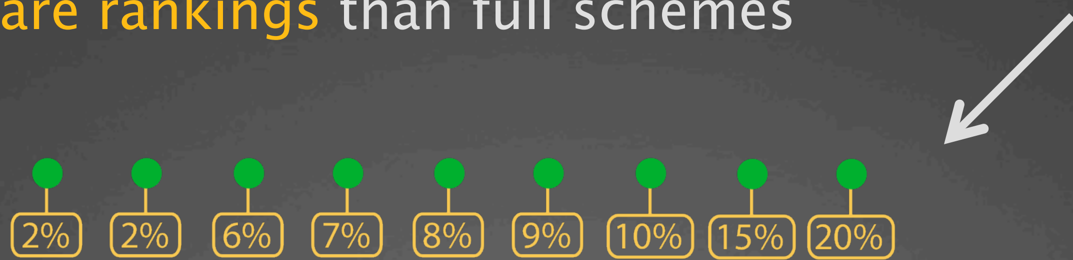
Internally, schemes assign probability to nodes

- Likelihood of being a Sybil

Leverage this to compare schemes?

- View schemes as inducing ranking on nodes
- Easier to compare rankings than full schemes

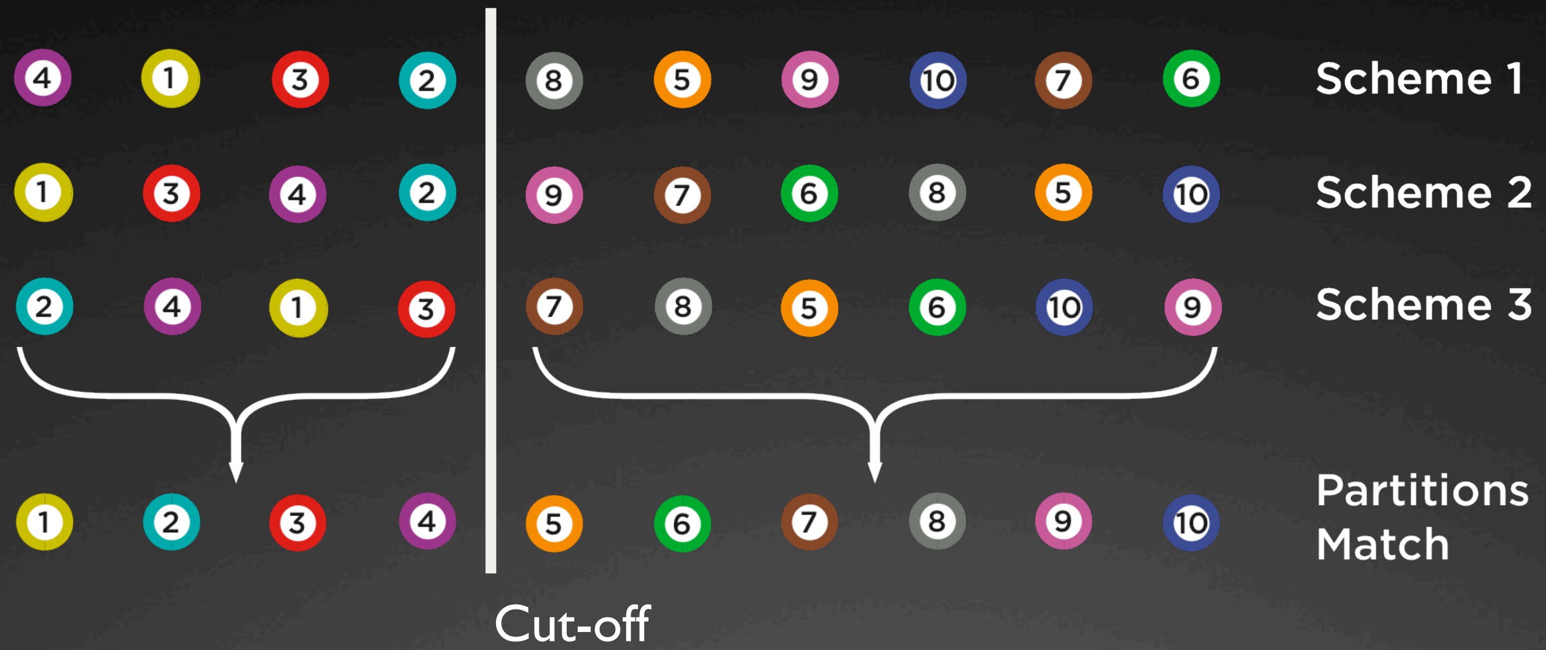




How do the rankings compare?

4	1	3	2	8	5	9	10	7	6	Scheme 1
1	3	4	2	9	7	6	8	5	10	Scheme 2
2	4	1	3	7	8	5	6	10	9	Scheme 3

How do the rankings compare?



All schemes observed to have distinct cut-off point

What is going on at this cut-off point?

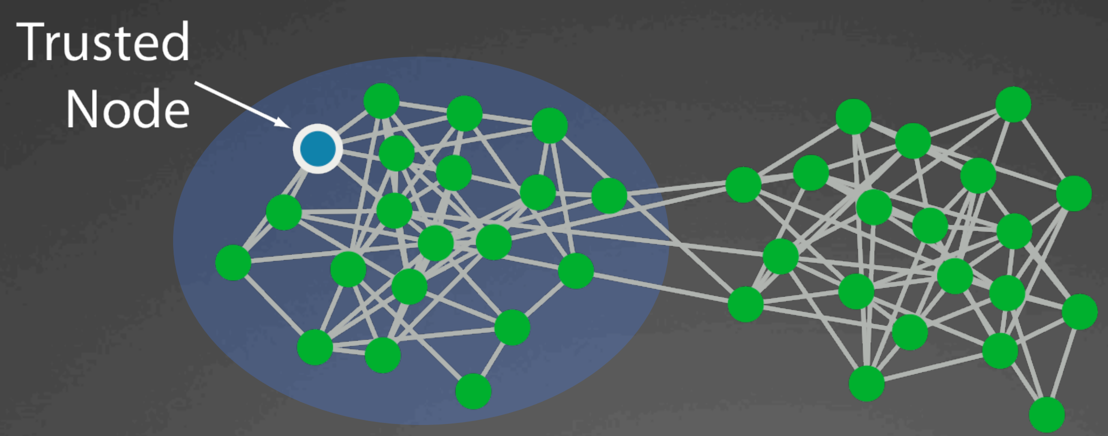
Where do the rankings match?

The cut-off point at the boundary of the local community

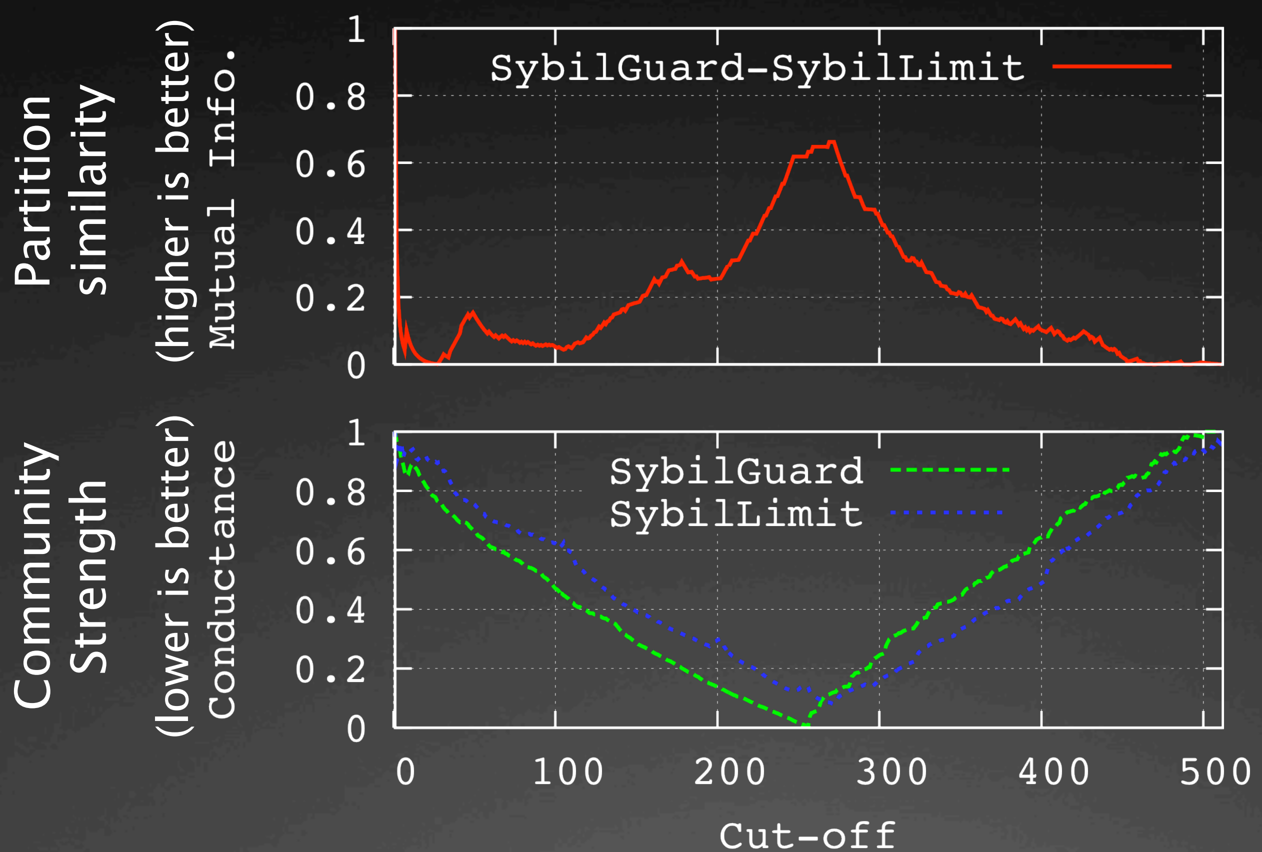
- Around the trusted node

Community well-defined in paper

- Roughly, set of nodes more tightly knit than surrounding graph



Investigating the cut-off point



Peak in similarly corresponds to boundary of local community

Details, more results in paper

Common insight across schemes

All schemes are effectively detecting communities

- Nodes in the local community are ranked higher
- Ranking within and outside community in no particular order

Implications

Leveraging community detection

Community detection is a well-studied topic

- Wealth of algorithms available

Can leverage existing work on community detection

- To design new approaches to detect Sybils

Also, better understand the limitations

What are the limitations?

Recall, schemes effectively finding local communities

Suggests dependence on graph structural properties

- Size, location, characteristics of local community

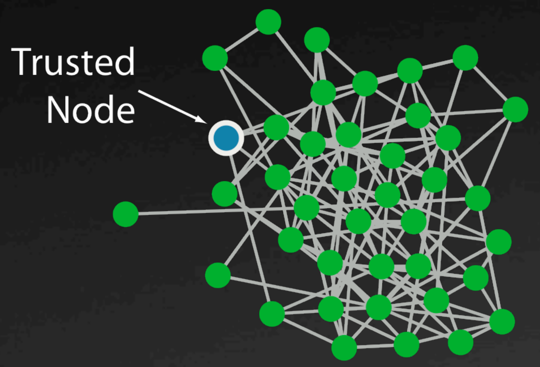
Explore two implications:

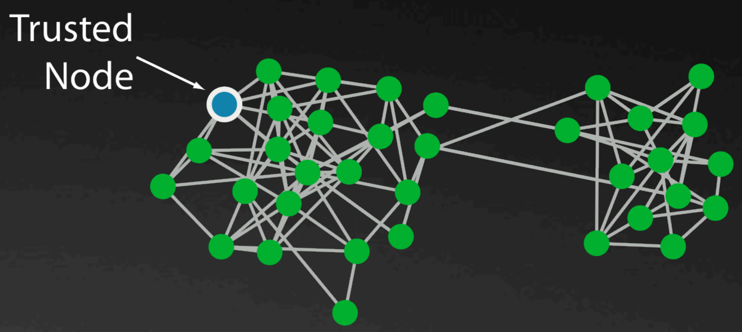
IMPLICATION 1 Are certain network structures more vulnerable?

IMPLICATION 2 What happens if the attacker knows this?

- Are more intelligent attacks possible?

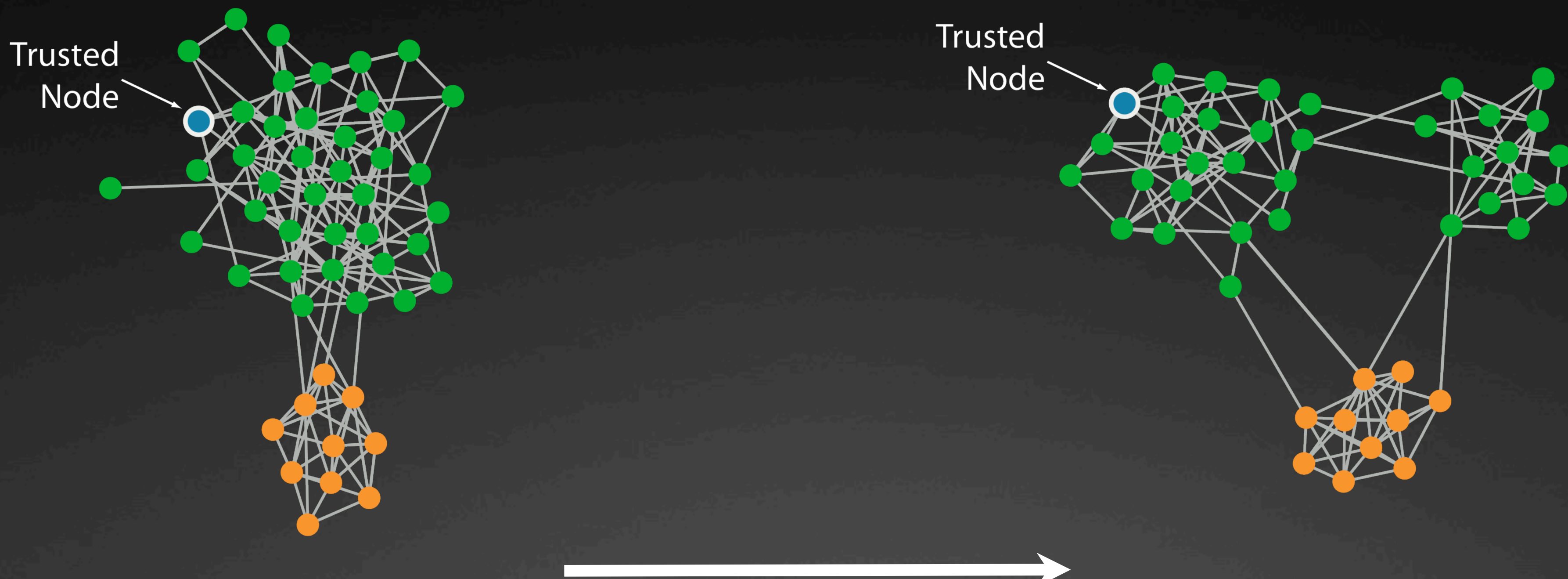
Certain network structures vulnerable?





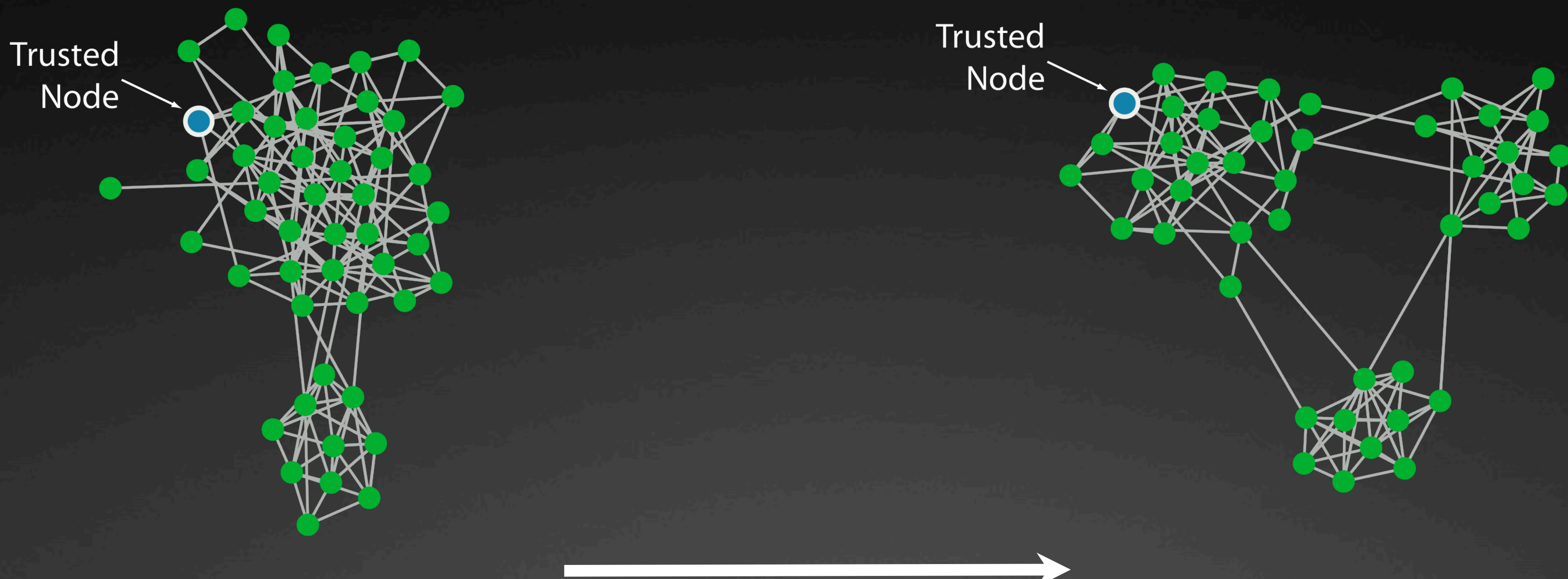
Increasing community structure of honest region

Certain network structures vulnerable?



Increasing community structure of honest region

Certain network structures vulnerable?



Increasing community structure of honest region

Hypothesis: Community structure makes identifying Sybils harder

Testing community structure hypothesis

Selected eight real-world networks

- Online social networks: Facebook (2)
- Collaboration networks: Advogato, Wikipedia, co-authorship
- Communication networks: Email

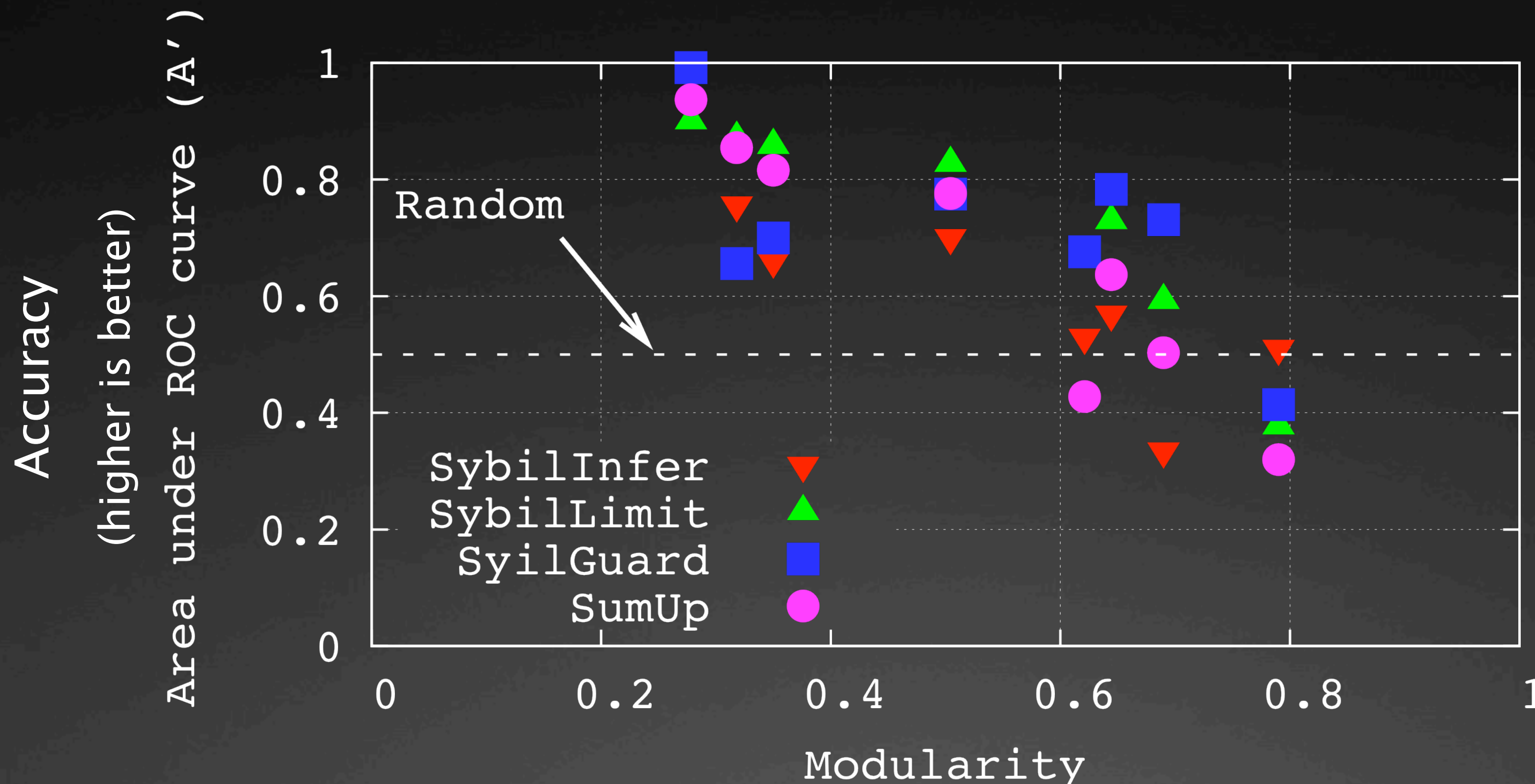
Simulated attack by consistently adding Sybils

- Similar strength attacker, despite different network sizes
- 5% attack links, 25% Sybil nodes

Measure accuracy using ranking

- Accuracy: Probability Sybils ranked lower than non-Sybils
- Fair comparison across schemes, networks

Impact of community structure?



Amount of community structure (modularity)

(higher is more community structure)

More community structure makes Sybils indistinguishable

Can attacker exploit this dependence?

Attacker's goal is to be higher up in the rankings

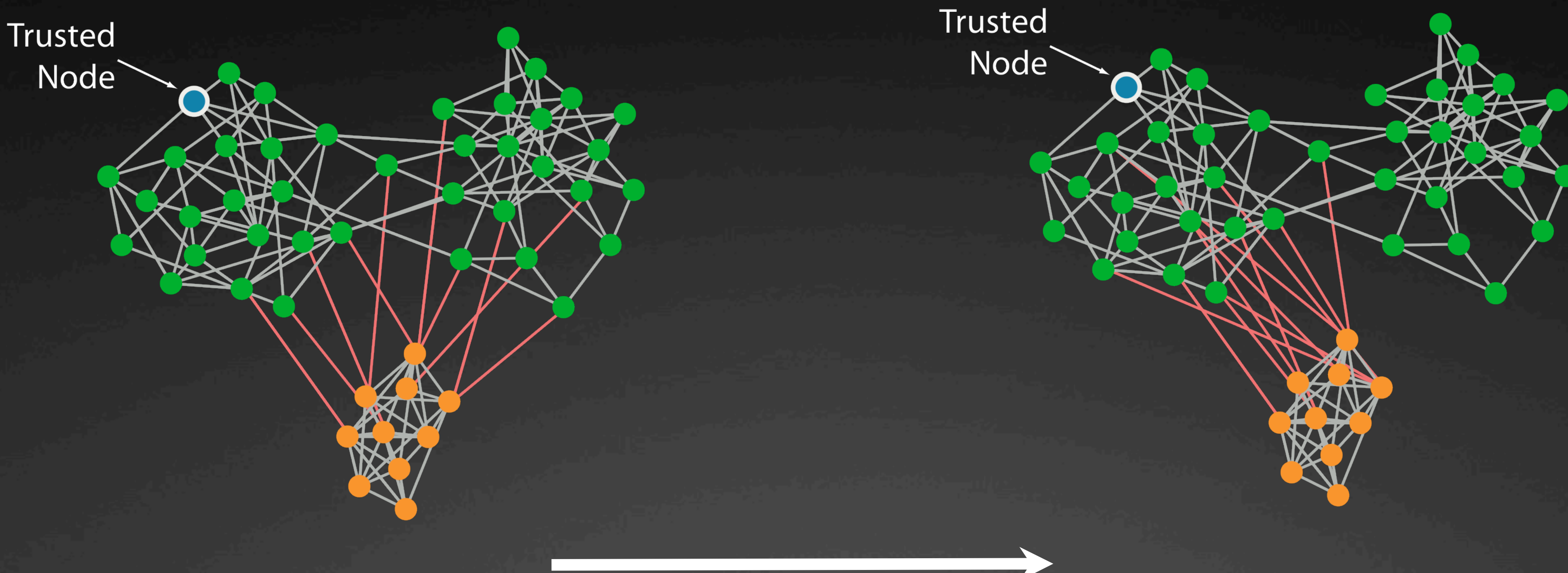
- Increases likelihood of being "accepted"

Existing Sybil schemes tested with "random" attackers

- Links placed to random non-Sybils

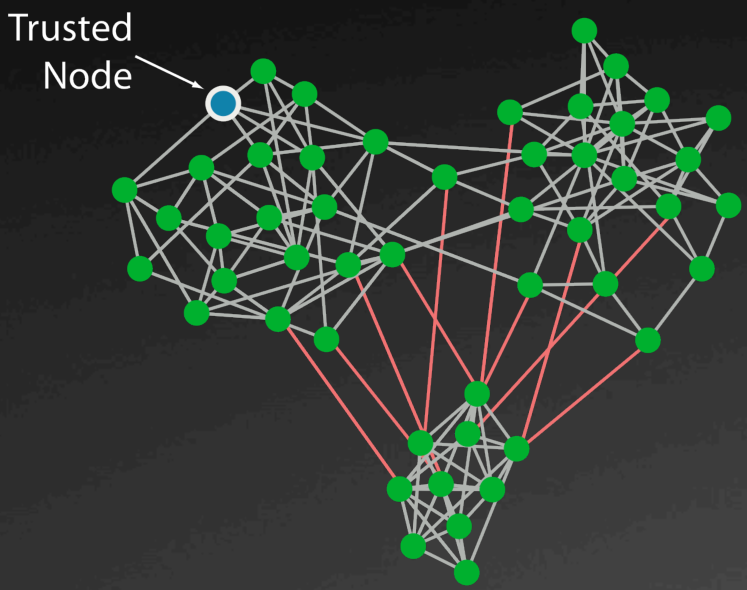
What happens if attacker given slightly more power?

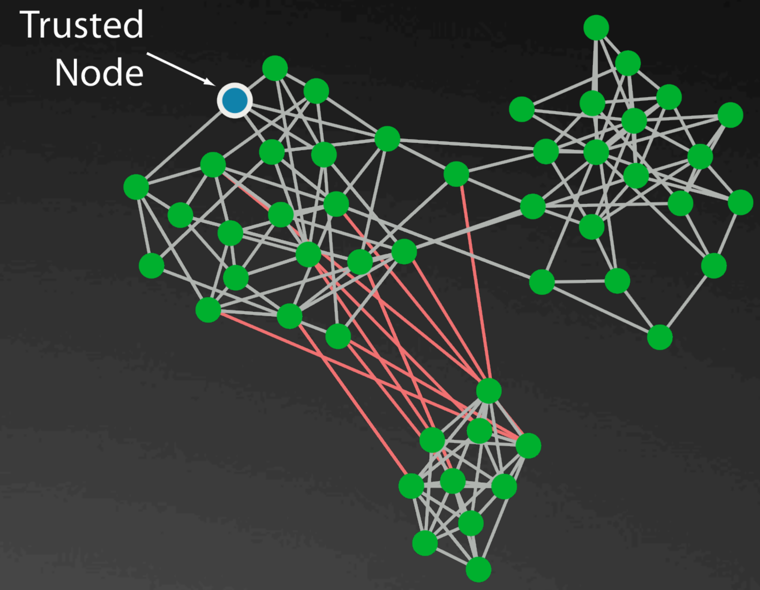
Changing attacker strength



Links placed closer to trusted node

Changing attacker strength





Links placed closer to trusted node

Hypothesis: Closer links makes Sybils harder to detect

Testing strong attacker hypothesis

Simulated attack by consistently adding Sybils

- Same strength as before

Allow attacker more flexibility in link placement

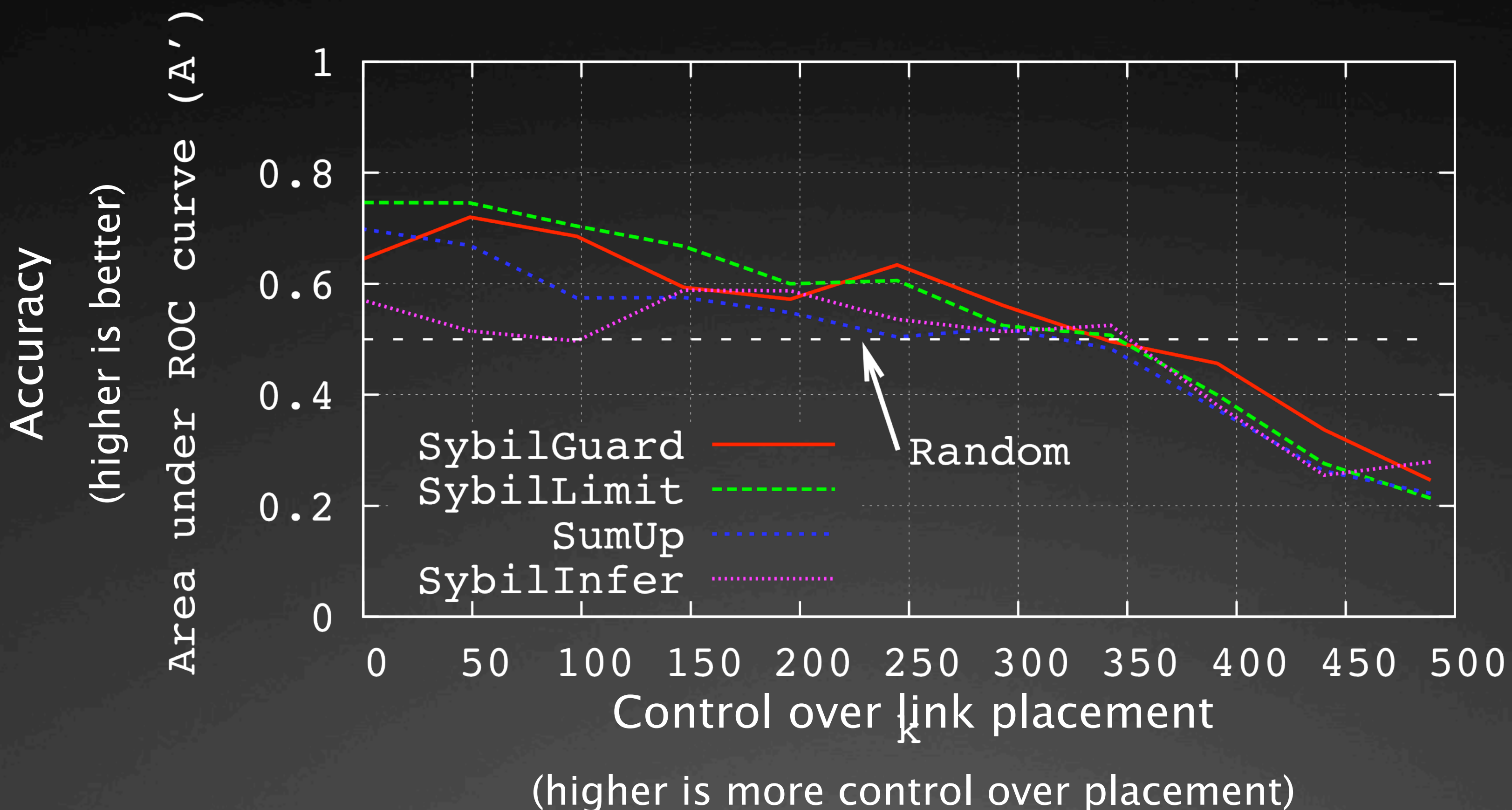
- Place links randomly among top N nodes; vary N
- Lower N represents more control

Present results on the Facebook network

- Tested other networks as well

What happens as Sybils given more control?

Impact of targeted links?



Attack becomes much more effective

Sybils ranked higher than non-Sybils (accuracy << 0.5)

Summary

Many social network-based Sybil defense schemes proposed

- All use very different mechanisms
- Hard to understand relationship, fundamental insight

Are they doing the same thing?

Developed methodology to compare schemes

- Found they are all detecting local communities

Significant implications of this finding

- Can leverage community detection for Sybil defense
- Certain networks more difficult to defend
- Attacker can exploit this to spend effort more wisely

Moving forward

Is social network-based Sybil defense always practical?

- Certain real networks have significant communities
- Could be still useful for white-listing small number of nodes

Is more information beyond graph structure helpful?

- More information about Sybil/non-Sybil nodes is useful
- Other information from higher layers eg. interaction

Questions?

Thank You!