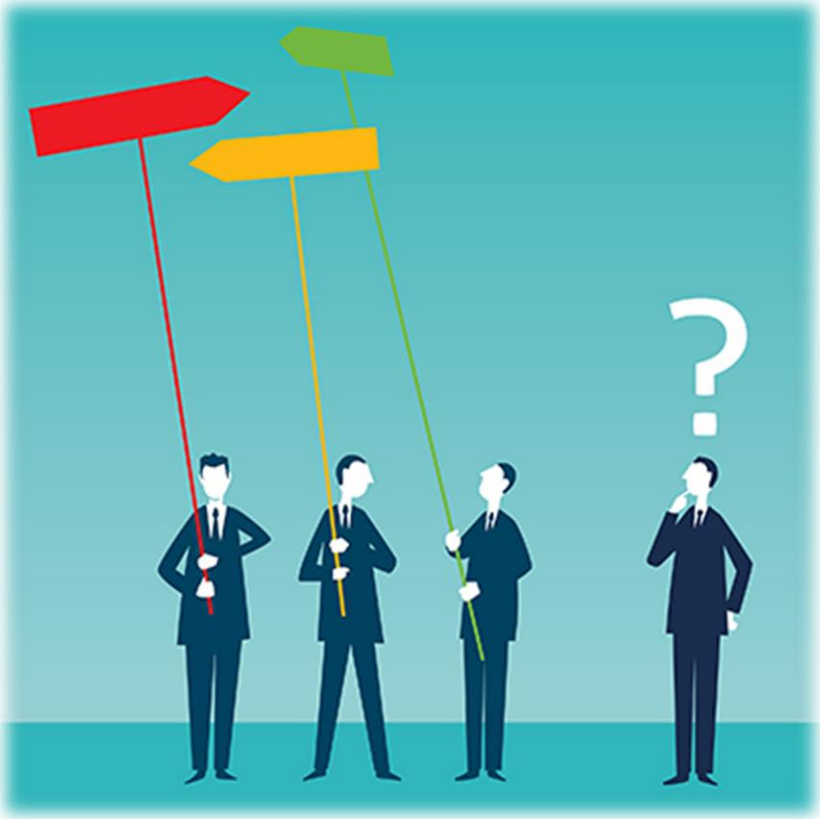
Sybil Detection Using Latent Network Structure

Grant Schoenebeck, Aaron Snook, **Fang-Yi Yu**

UNIVERSITY OF MICHIGAN™

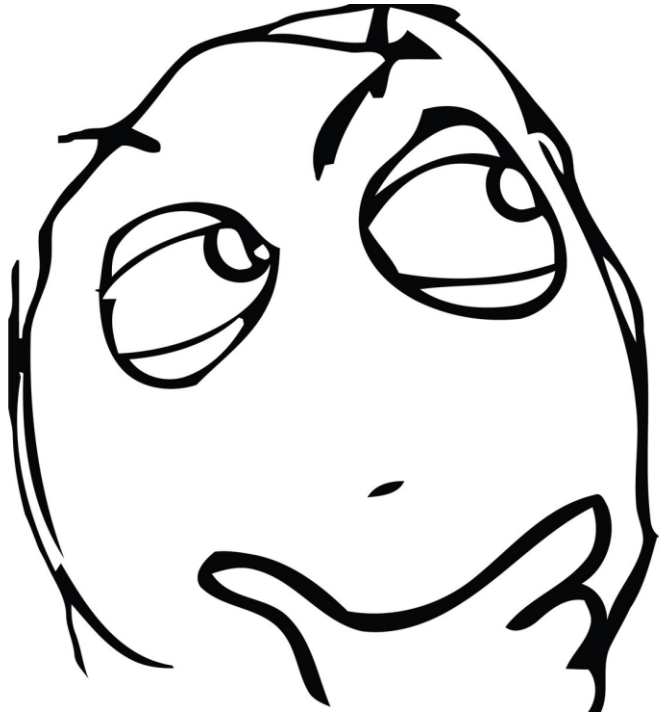
Sybil Attack

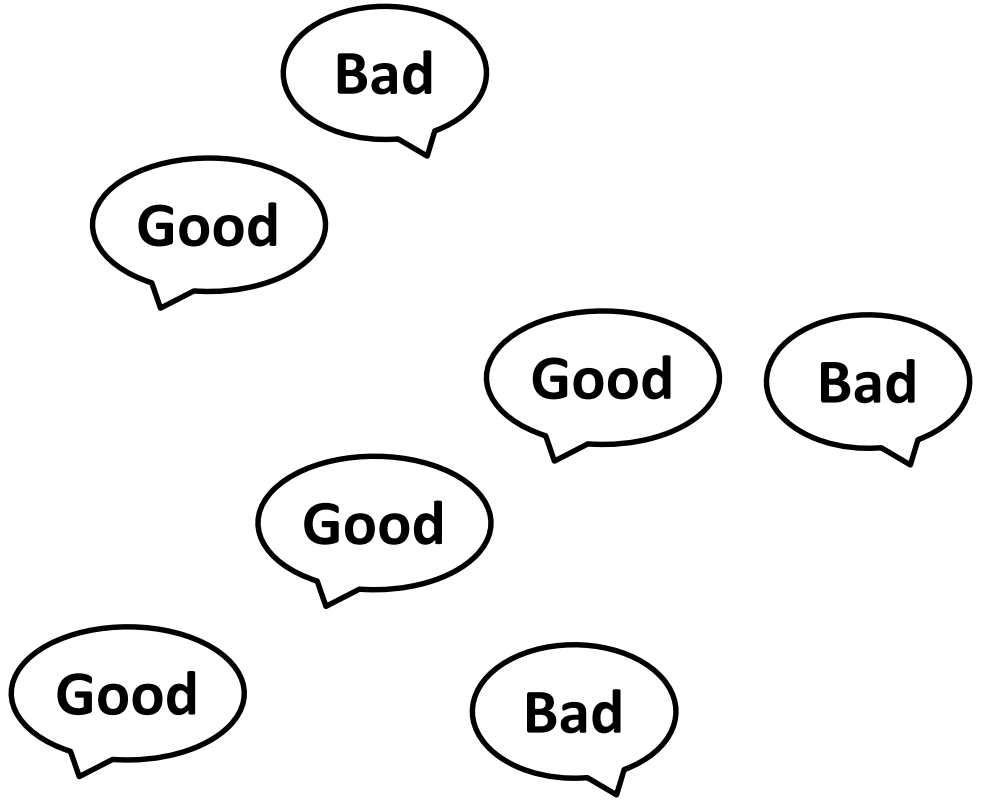
- An **attack** to compromise a recommendation systems by forging identities.



Recommendation System

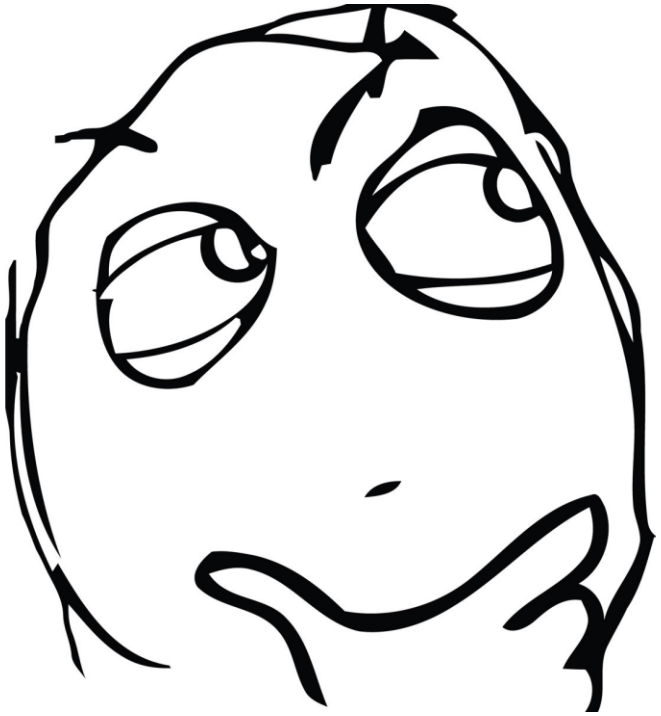
How is that restaurant?

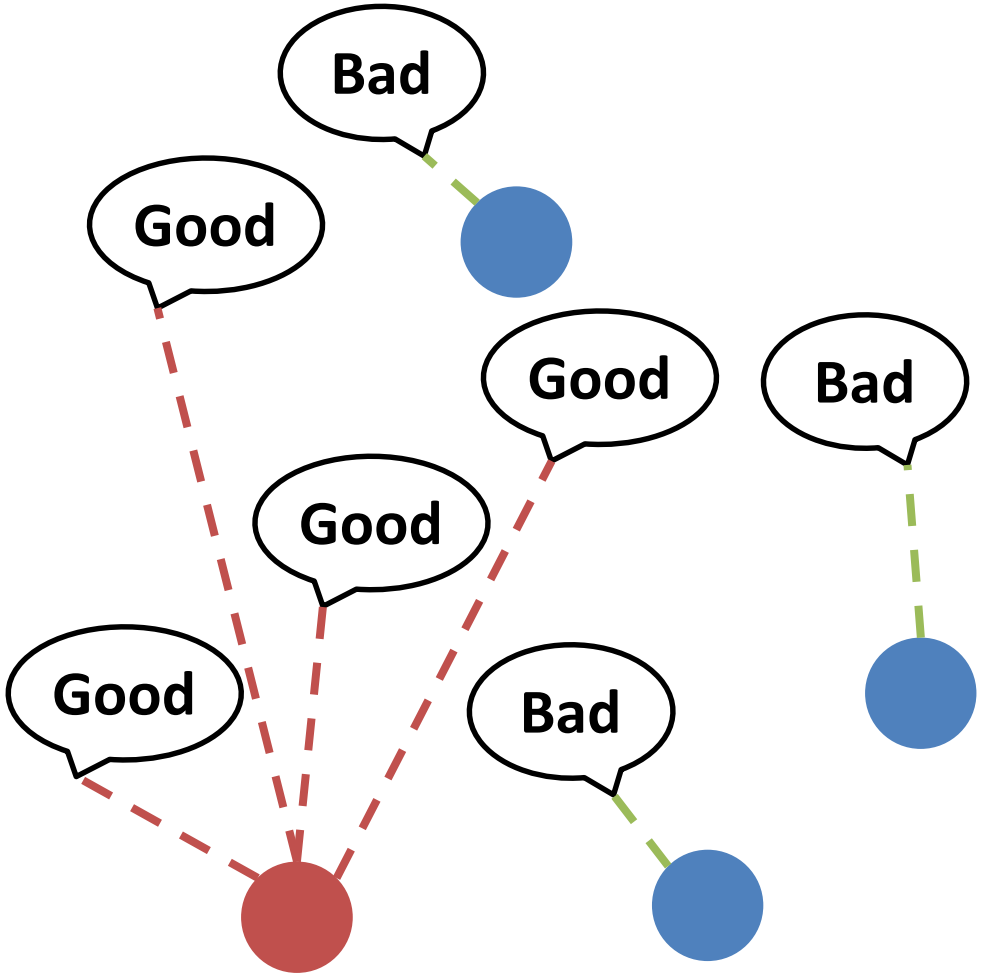




Sybil Can Manipulate the Opinion

How is that restaurant?





Activities and Profile Characteristics

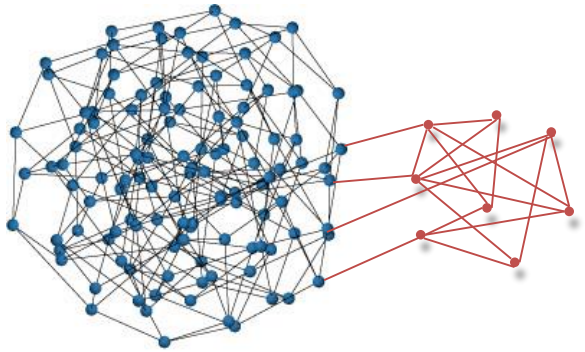
- Pros
  - Proliferating signals to exploit
  - Practical benefits
- Cons
  - Cat and mouse game

Structure of the Social Network

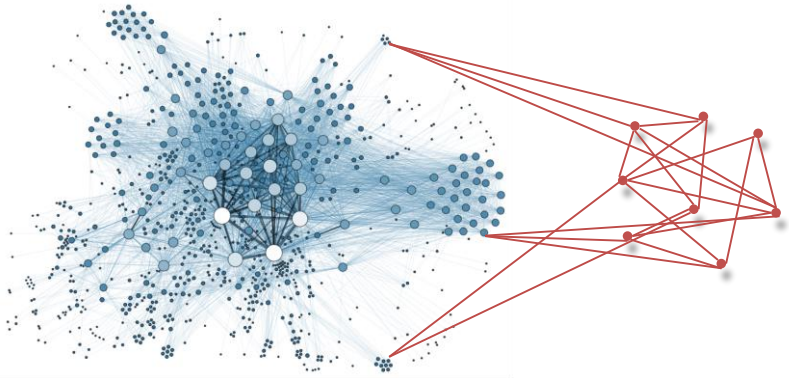
- Pros
  - Expensive signal to forge
- Cons
  - Stringent conditions

Assumptions on Network Topology

- Assuming distinct ability
  - Honest nodes: Well-mixed networks
  - Sybil: Limited connection to the honest



- Empirical results [Alvisi 2013]
  - Social networks don't have fast mixing time
  - Sybil are accepted as friends much higher than anticipated



Alternative Assumptions

Previous Assumptions

- Honest nodes:
  - Well-mixed networks
- Sybil:
  - Limited connection to the honest

Goal

- Recover all honest agents

Our Assumptions

- Honest nodes:
  - 'locally' dense in low dimensional space
- Sybil:
  - relax to constant fraction of honest agent would be compromisable

Goal

- core space: a **whitelist** of nodes

Alternative Assumptions

Previous Assumptions

- Honest nodes:
  - Well-mixed networks
- Sybil:
  - Limited connection to the honest

Goal

- Recover all honest agents

Our Assumptions

- Honest nodes:
  - 'locally' dense in low dimensional space
- Sybil:
  - relax to constant fraction of honest agent would be compromisable

Goal

- core space: a **whitelist** of nodes

Low Dimensional Latent Metric Space

- Intuition
  - Metrics space encodes the similarity between agents
- Well-regarded network models
  - Watts-Strogatz model: **ring**
  - Kleinberg's small world model: **lattices**
  - Low distortion multiplex social network [Abraham2013]

Our Low Dimensional Assumptions

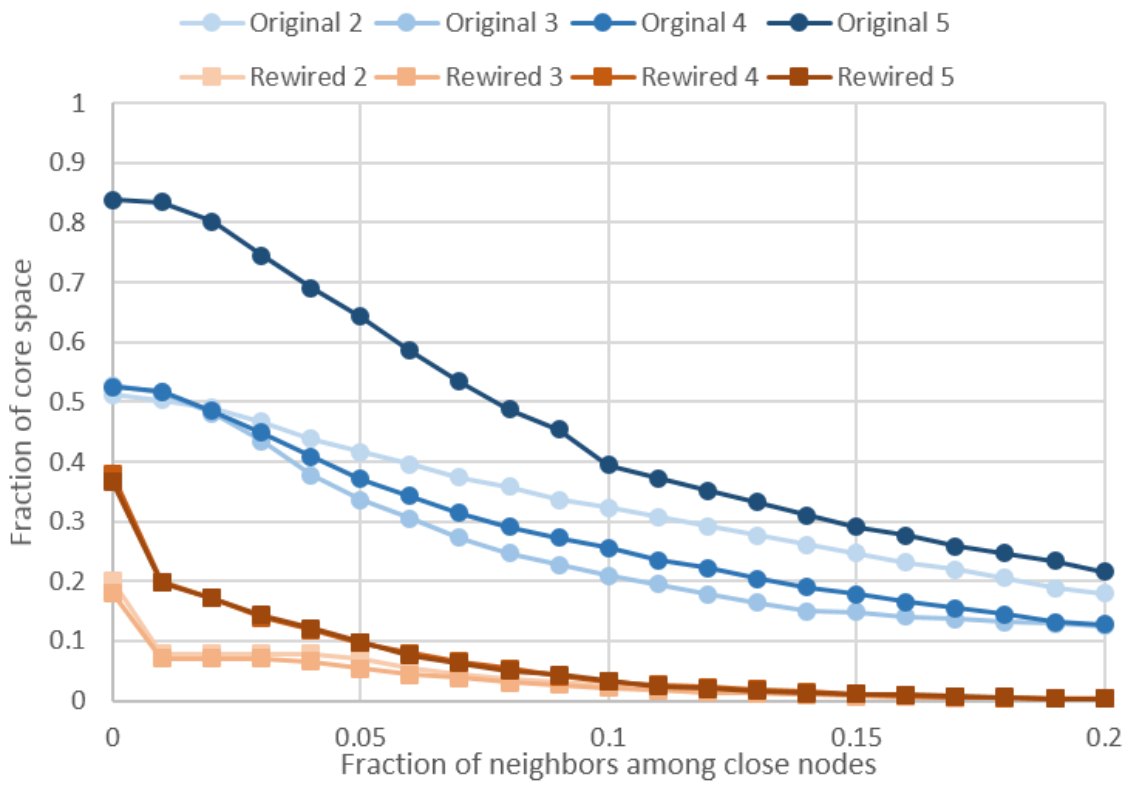
- Dimensionality
  - Graph with pairwise distance
  - Requiring low doubling dimension having $\mathbb{R}^d$ as special cases
- Density
  - Every local region contains a random graph
  - Only require of constant fraction of nodes
- **How realistic are our assumptions**

Experiment Setups

- Dataset Description
  - Facebook
  - Twitter
  - Wiki-vote
  - Epinion
- Implementation
  - Use Spectrum embedding
  - Compute the core space

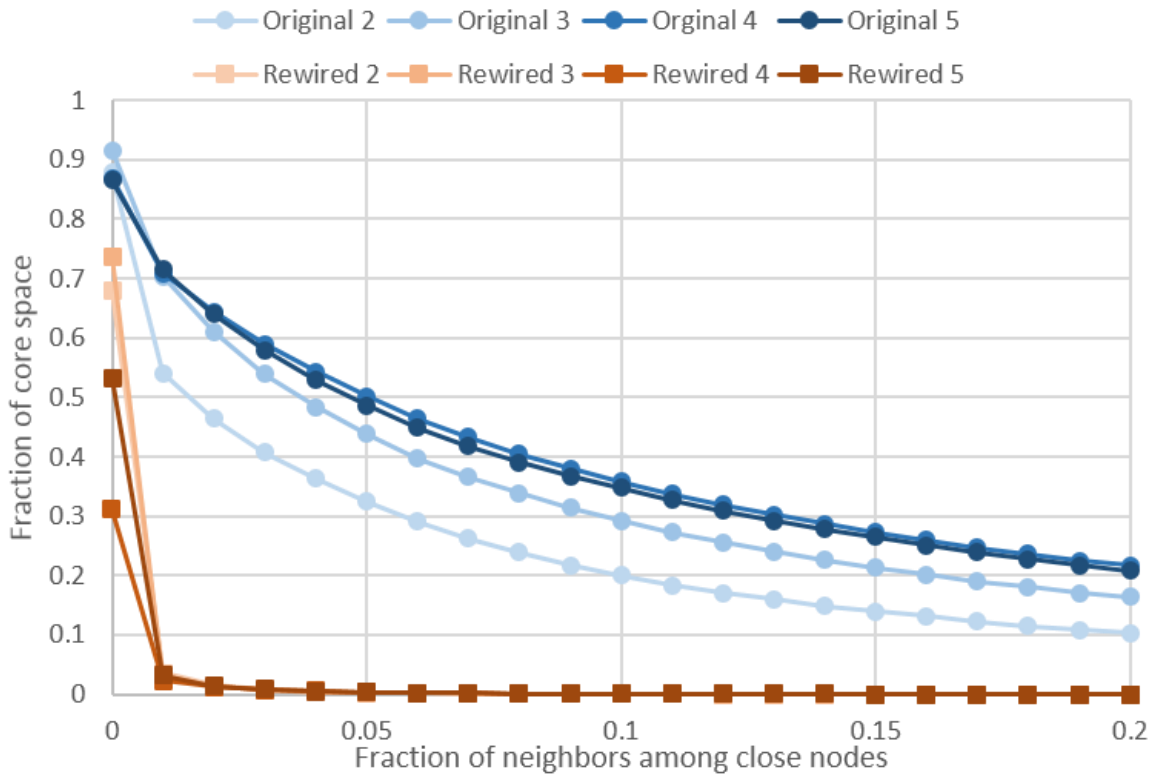
Core Space in Facebook

- Graph properties
  - 4,039 nodes, 88,234 edges
  - Average degree 21.8
- Core space
  - Density > 10
  - Connect to p fraction of nearby nodes



Core Space in Twitter

- Graph properties
  - 81,306 nodes, 1,768,149 edges
  - Average degree 21,75
- Core space
  - Density > 10
  - Connect to p fraction of nearby nodes



Alternative Assumptions

Previous Assumptions

- Honest nodes:
  - Well-mixed networks
- Sybil:
  - Limited connection to the honest

Goal

- Recover all honest agents

Our Assumptions

- Honest nodes:
  - 'locally' dense in low dimensional space
- Sybil:
  - relax to constant fraction of honest agent would be compromisable

Goal

- core space: a **whitelist** of nodes

Compromisable Agents

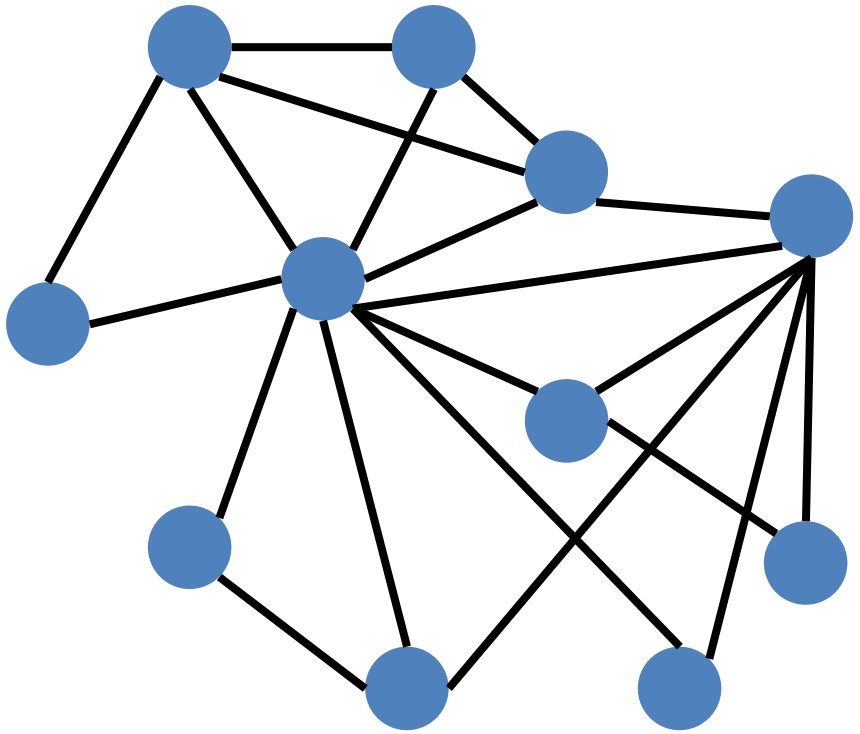
- Idea
  - Someone might accept all the friend requests
- Honest nodes
  - Most of the nodes are trustworthy
  - A random portion of nodes are compromisable
- Sybils
  - Cannot connect to trustworthy nodes

Assumptions Summary

Assumptions	Social network	Sybils
Previous Works	Well-mixed	Bounded connection to honest nodes
Our Work	Locally dense in low-dimensional space	Only connection to compromisable nodes

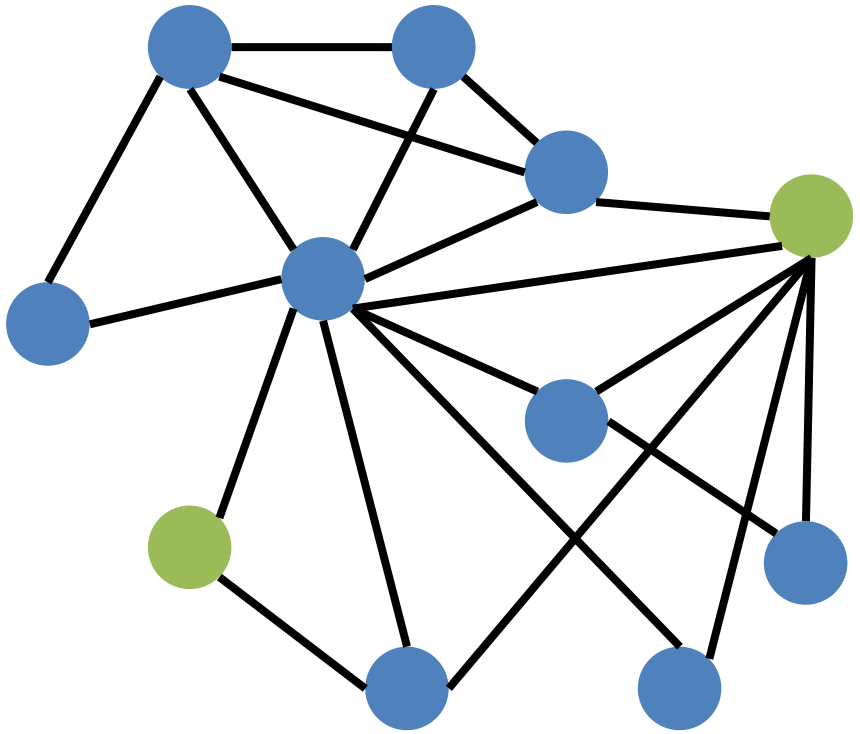
Detection Game

- Original Graph



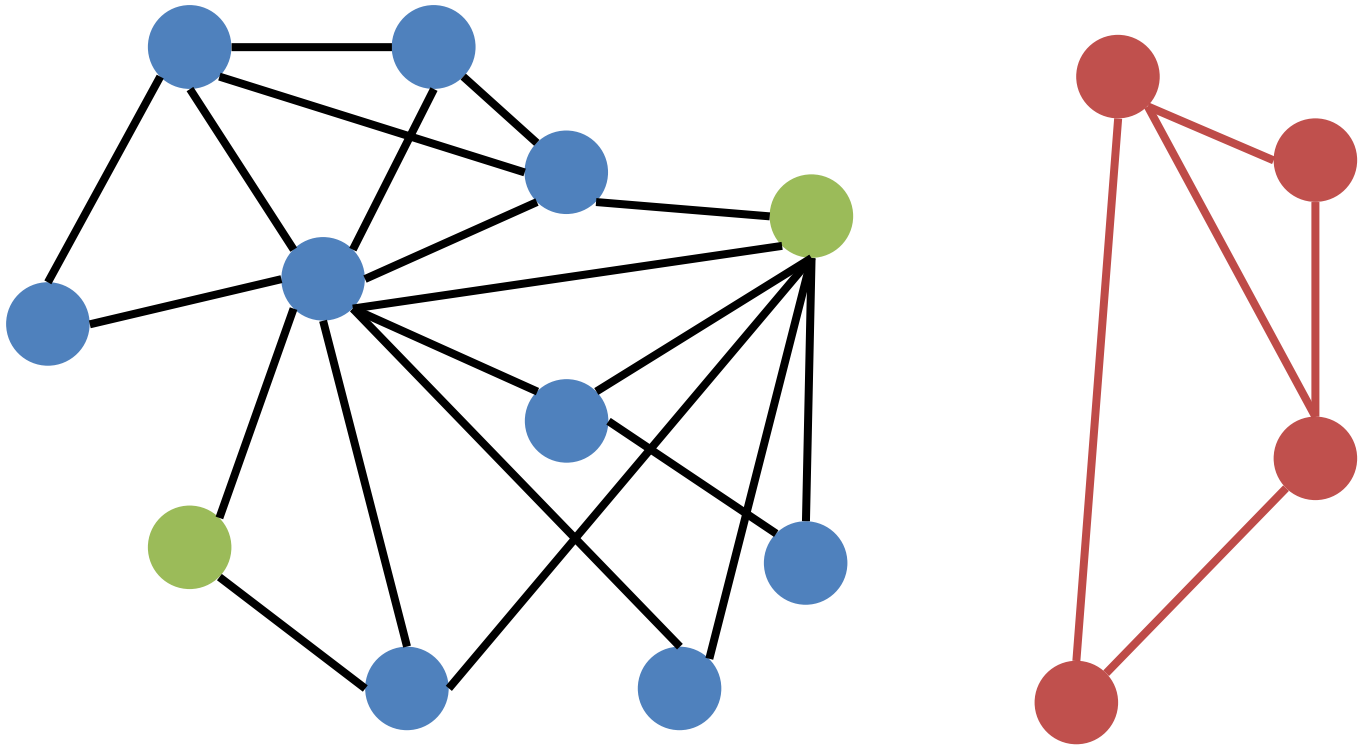
Detection Game

- Reveal the trustworthy and compromisable nodes



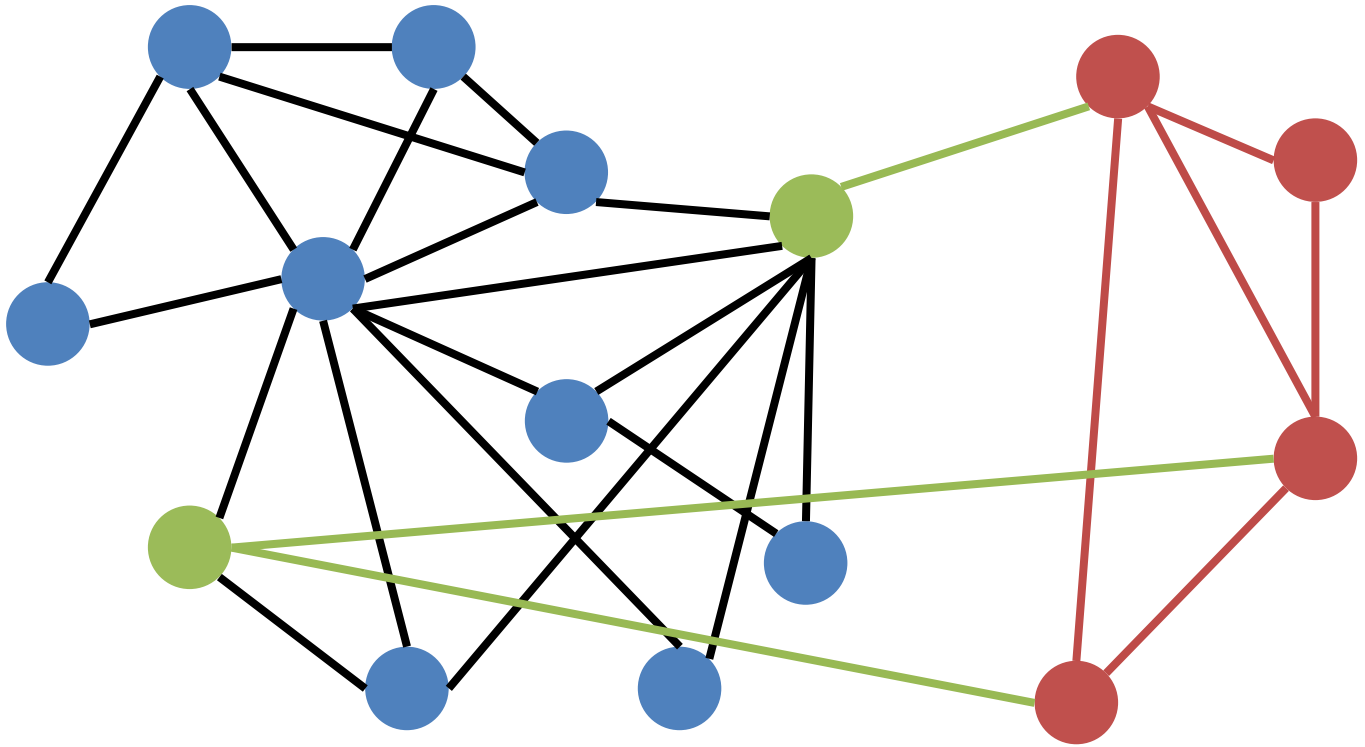
Detection Game

- Adversary try to add Sybil nodes into the networks



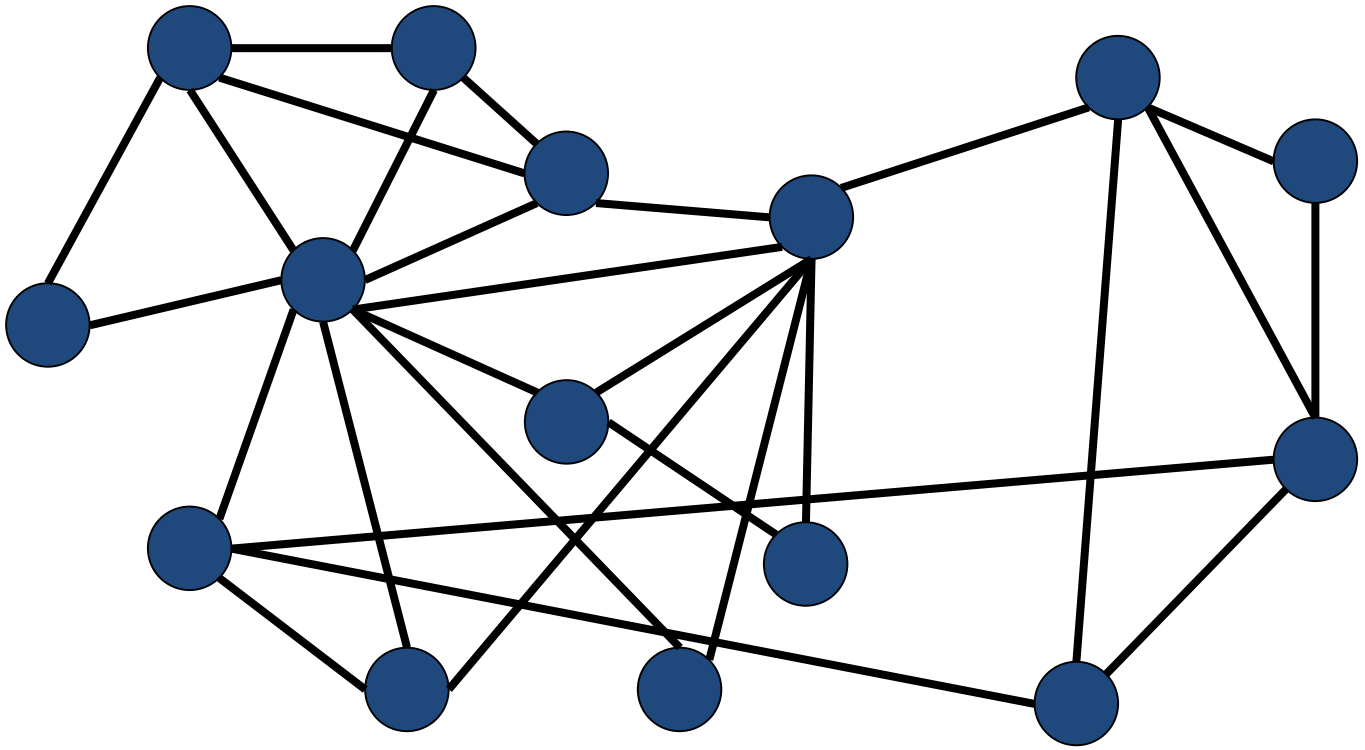
Detection Game

- Adversary try to add Sybil nodes into the networks



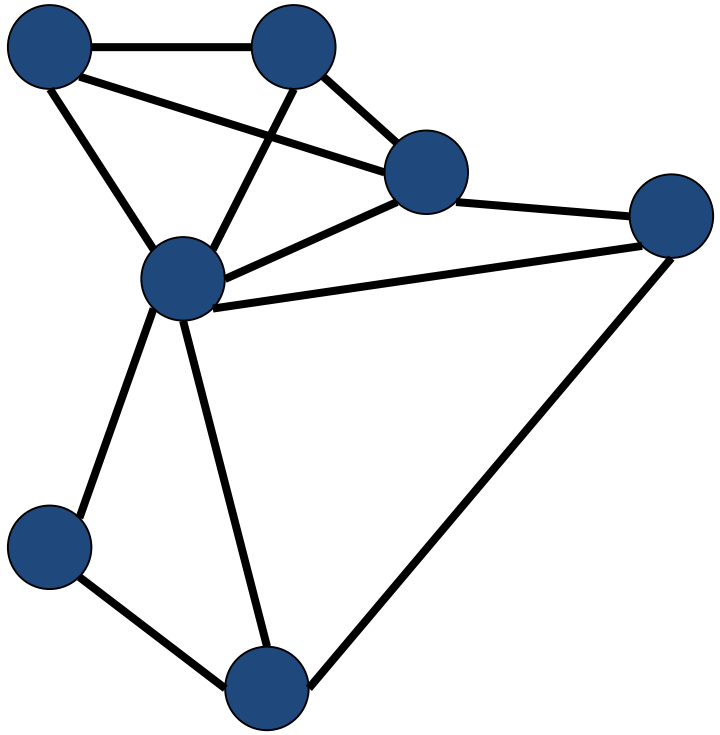
Detection Game

- Detection algorithm return a **whitelist**



Detection Game

- Detection algorithm return a **whitelist**



Theorem

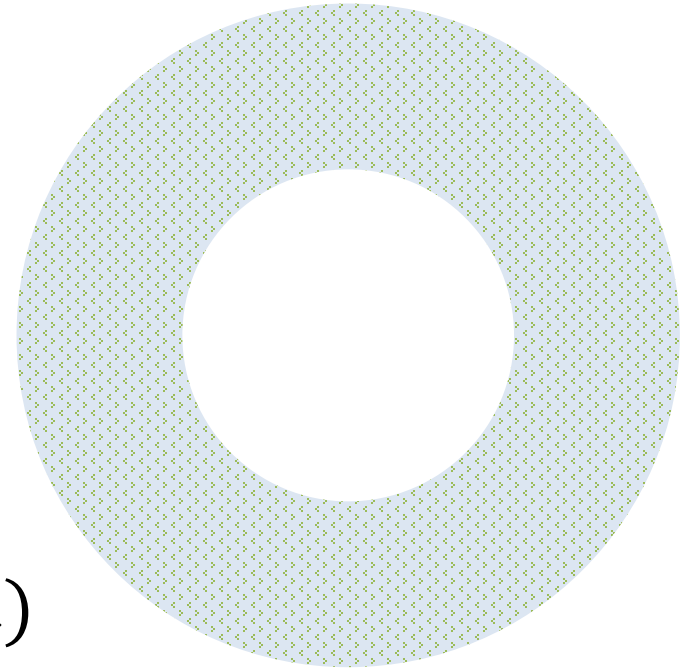
- If the total number of Sybil nodes and Compromisable nodes is smaller than some constant fraction the honest nodes, and the graph can be imbedded into locally dense low dimensional space,
  in the **detection game** for any adversary the detection algorithm can return a large whitelist without any Sybil

A Toy Model

- Network of honest nodes
  - 1 dimensional unit circle
  - n nodes uniformly placed
  - Well-connected within distance $\frac{1}{\log n}$
- Limitation of Sybils
  - Connects to Sybil or compromisable node
  - #Sybil $= O(n)$, #the Compromisable $= O(n)$

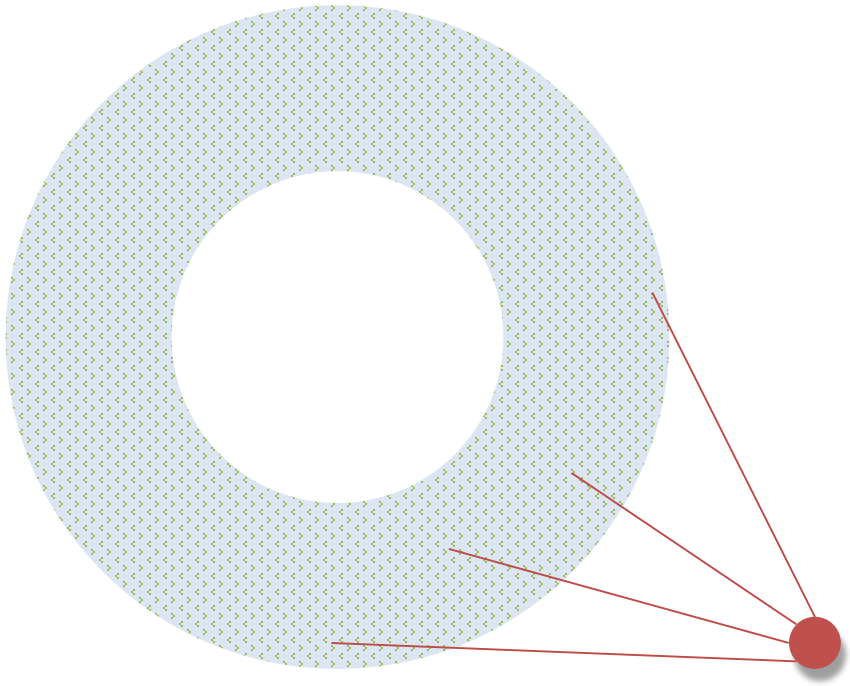
A Toy Model

- Network of honest nodes
  - 1 dimensional unit circle
  - n nodes uniformly placed
  - Well-connected within distance $\frac{1}{\log n}$
- Limitation of Sybils
  - Connects to Sybil or compromisable node
  - #Sybil $= O(n)$, #the Compromisable $= O(n)$

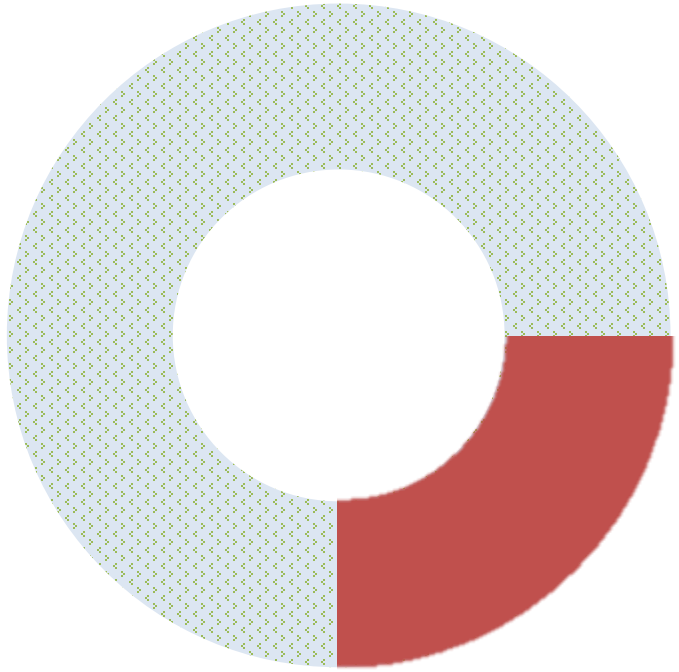


What can Sybil do?

Connect to the compromisable

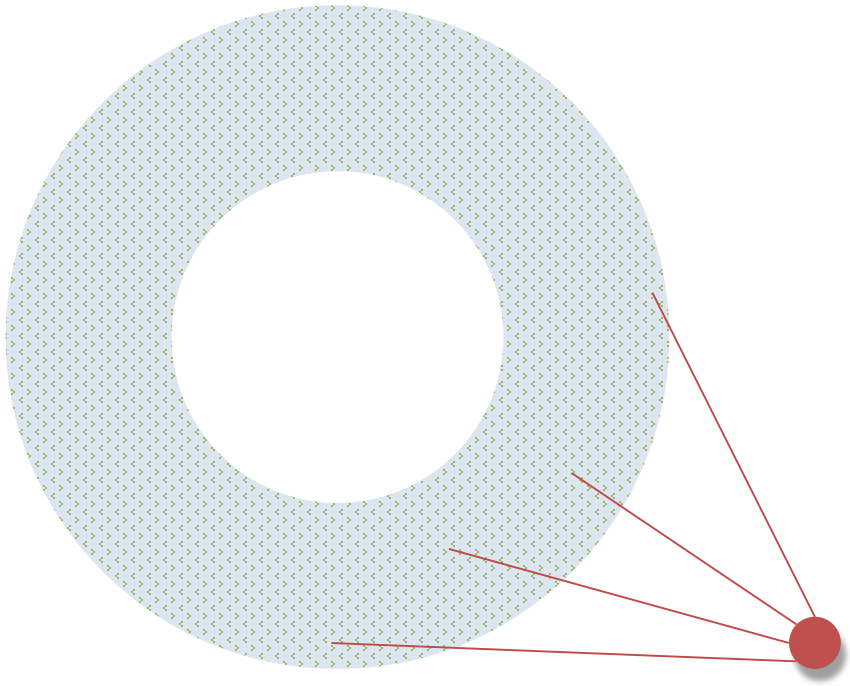


Form its own network

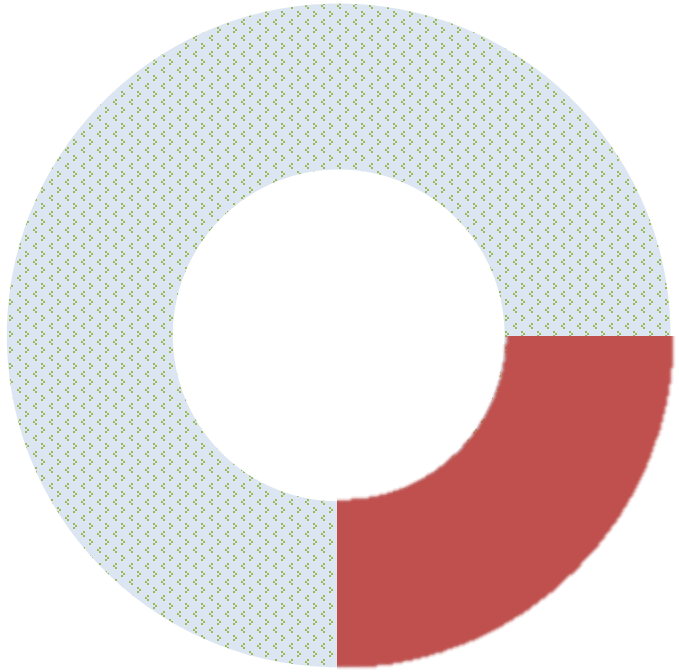


What should detection algorithm do?

Remove non-local edges



Remove low degree nodes



Future Work

- Can we do better if we have information of compromisable nodes?