General Threshold Model for Social Cascades

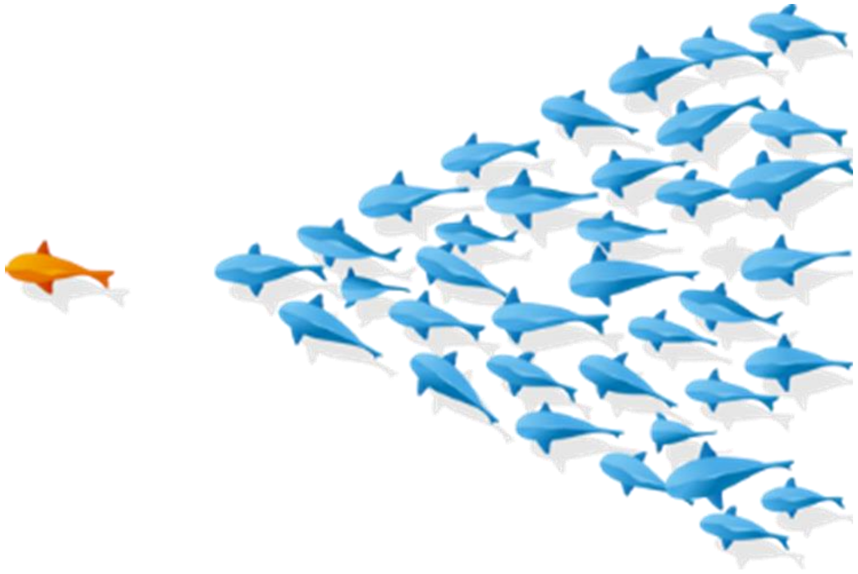
Jie Gao, Golnaz Ghasemiesfeh,
Grant Schoenebeck, **Fang-Yi Yu**

Stony Brook University

UNIVERSITY OF MICHIGAN

Contagions, diffusion, cascade…

- Ideas, beliefs, behaviors, and technology adoption spread through network
- Why do we need to study this phenomena?
  - Better Understanding
  - Promoting good behaviors/beliefs
  - Stopping bad behavior



Outline

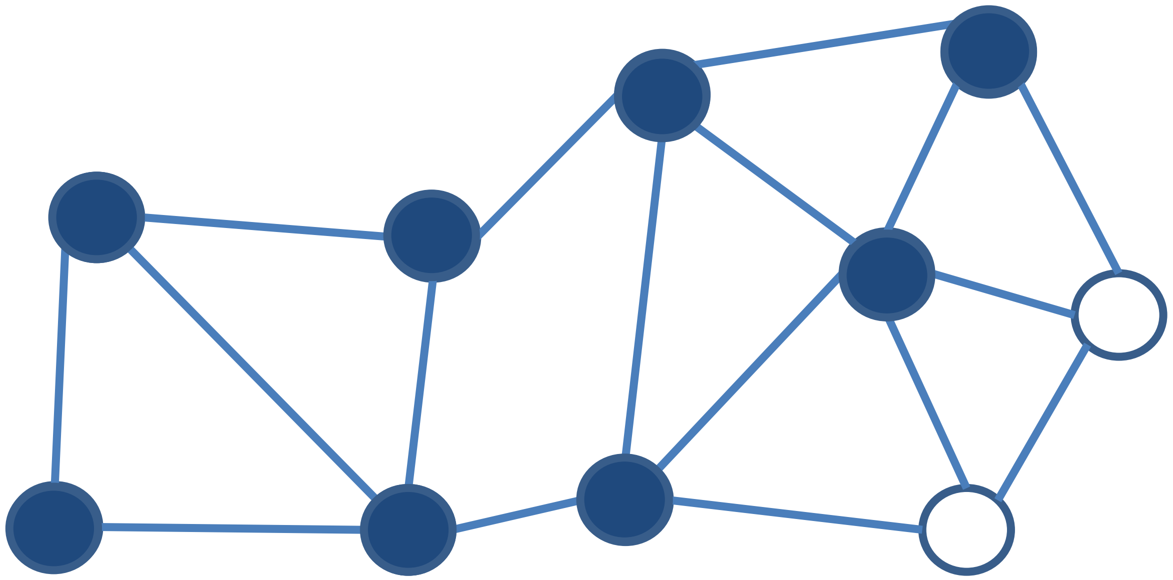
- Cascade Model
- Empirical Results: Testing Network Models
  - Real Data
  - Synthetic Models
- Theoretical Results
  - Directed case
  - Undirected case

Outline

- Cascade Model
- Empirical Results: Testing Network Models
  - Real Data
  - Synthetic Models
- Theoretical Results
  - Directed case
  - Undirected case

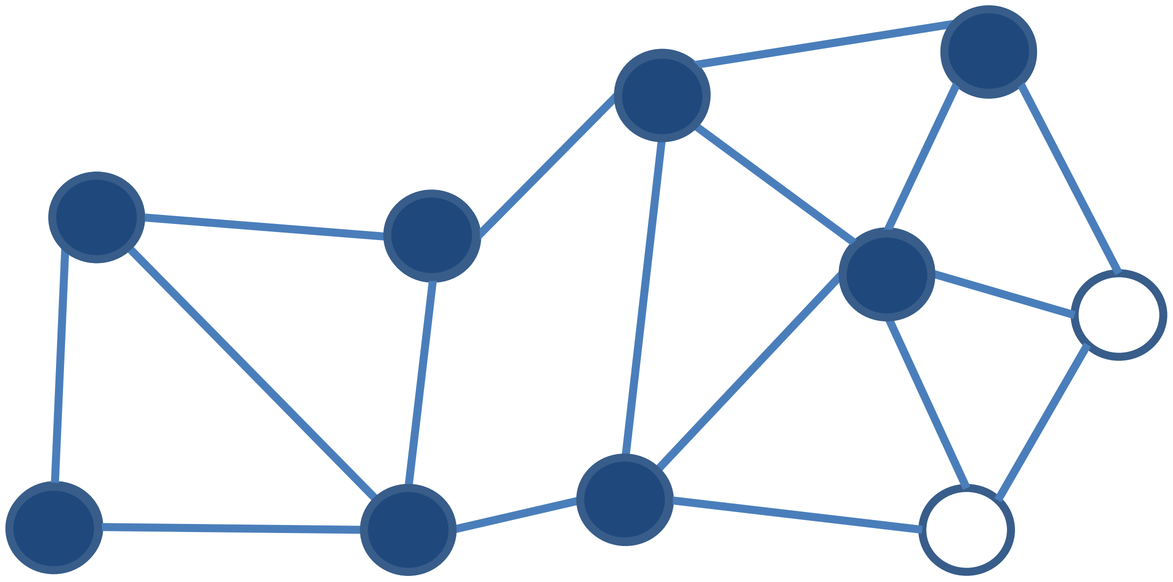
Social Contagion

- Contagion is a chain reaction that starts with early adopters and spreads through the social network



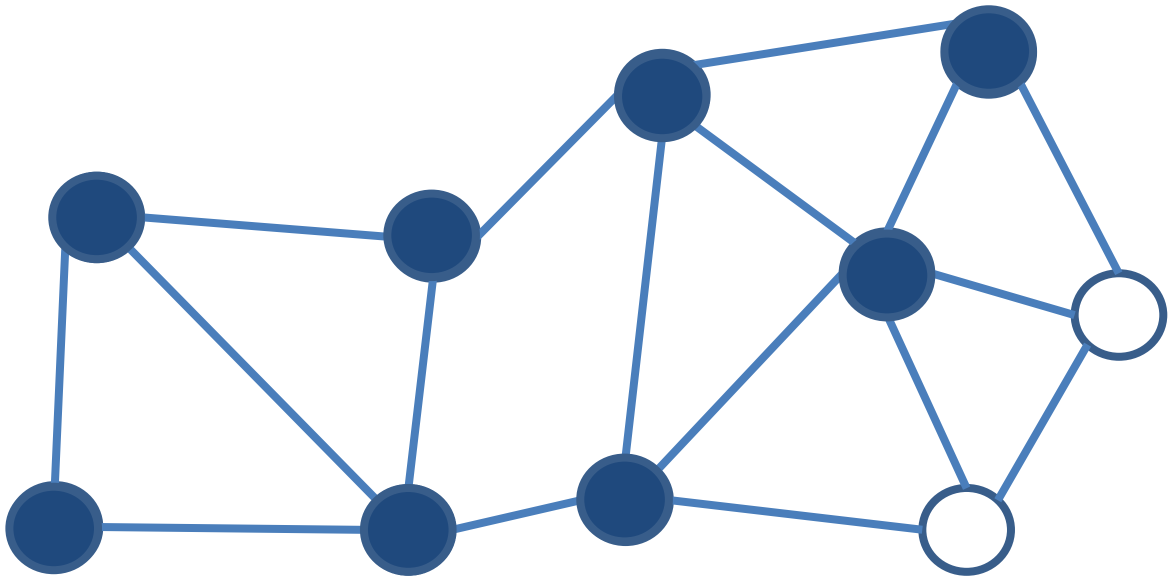
Social Contagion

- Contagion is a chain **reaction** that starts with early adopters and spreads through the social network



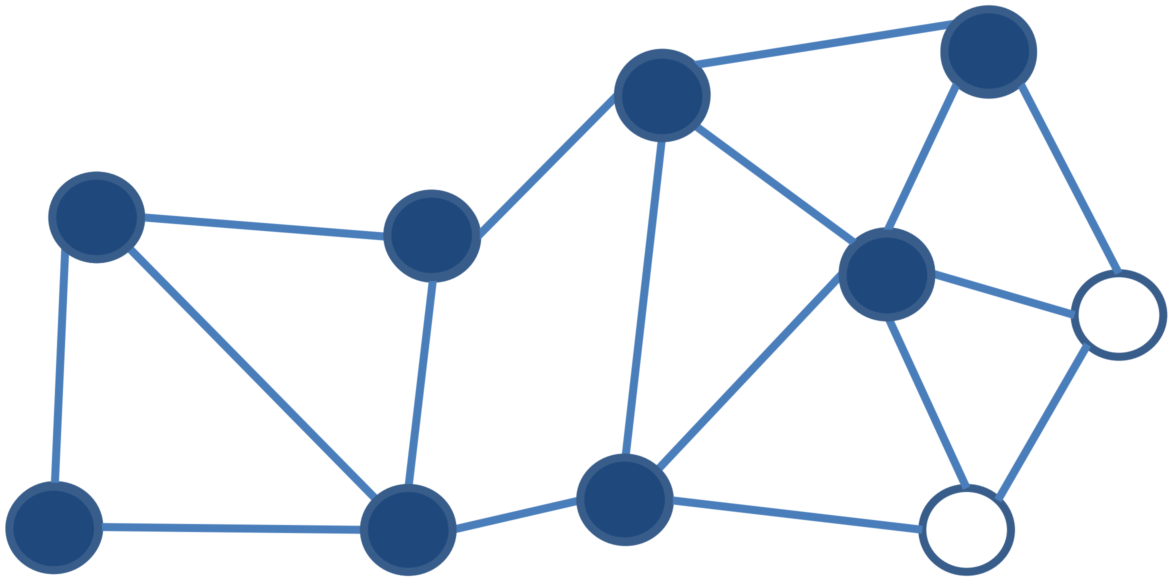
Social Contagion

- Contagion is a chain **reaction** that starts with **early adopters** and spreads through the social network



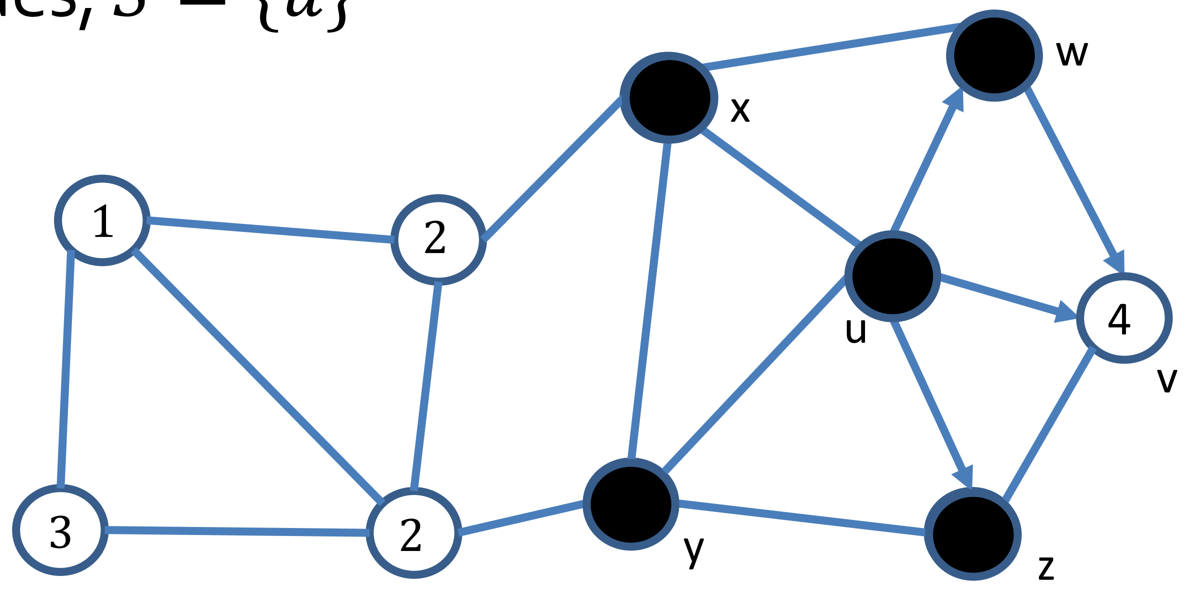
Social Contagion

- Contagion is a chain **reaction** that starts with **early adopters** and spreads through the **social network**



General Threshold Contagion

- General Threshold Contagion GTC(G,D,S) [G 1973; MR 2010]
  - **Social network:** Graph, G
  - **Reaction:** Threshold distribution, $D = U_{\Delta}$
  - **Early adopters**: Seeded nodes, $S = \{u\}$



How general is this model?

- Captures many models as special cases
  - Independent cascade
  - Linear threshold model
  - k-complex contagion

Outline

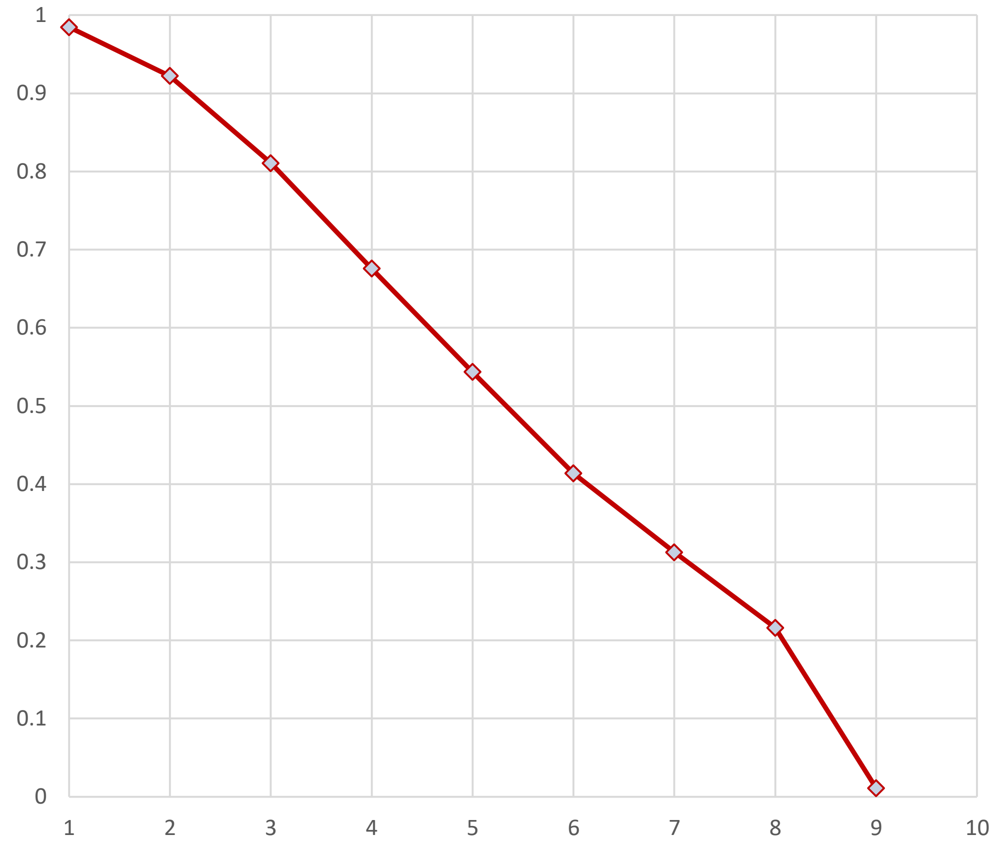
- Cascade Model
- Empirical Results: Testing Network Models
  - Real Data
  - Synthetic Models
- Theoretical Results
  - Directed case
  - Undirected case

Experiment Setups

- G: graph
  - DBLP co-authorship network with 317,080 nodes
  - Stanford web graph with 281,903 nodes
- D: threshold ~ Poisson distribution with different mean λ
- S: The 'earliest' 25 nodes

Contagion on DBLP Database

- G: DBLP co-authorship network
  - 317,080 nodes 1,049,866 edges
  - **3.3** average degree
- D: Poisson distribution
- S: The 'earliest' 25 nodes



Outline

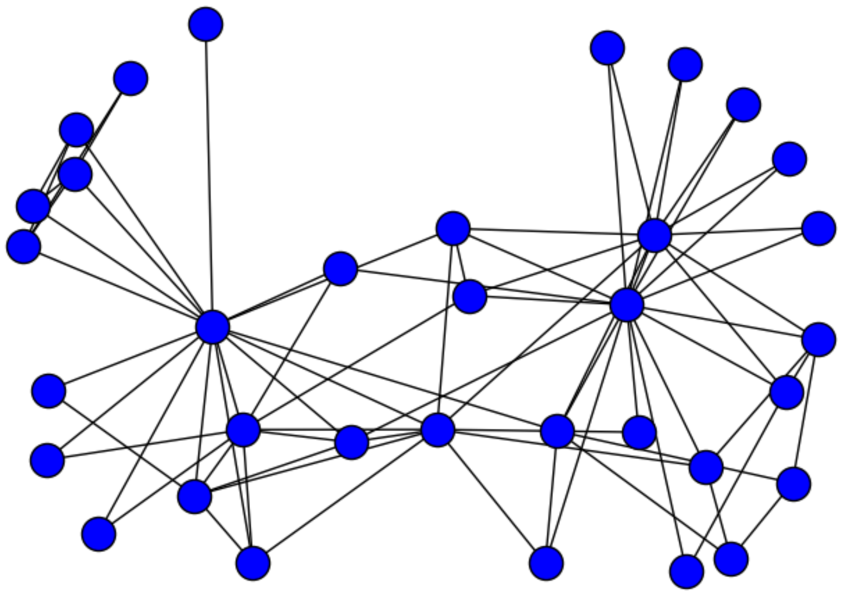
- Cascade Model
- Empirical Results: Testing Network Models
  - Real Data
  - Synthetic Models
    - Configuration Model
    - Stochastic Attachment Model
- Theoretical Results
  - Directed case
  - Undirected case

Social Networks

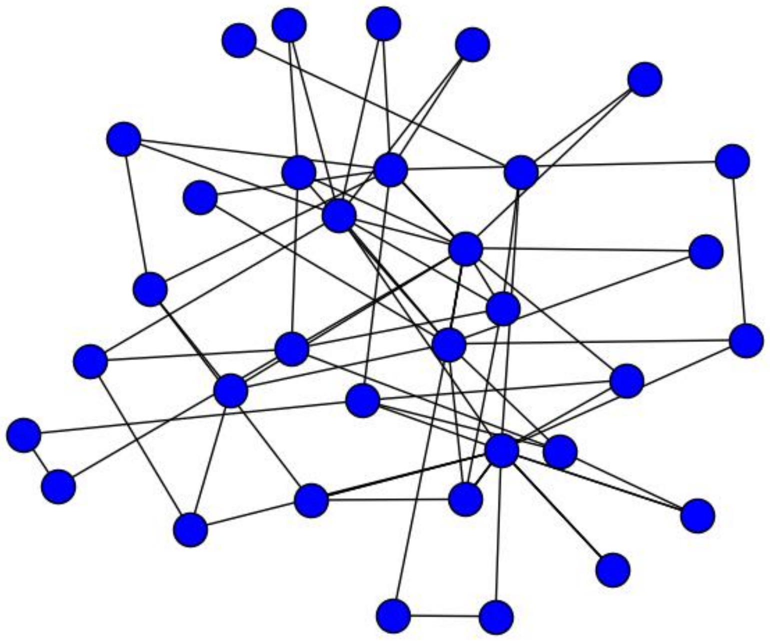
- Can we **generate** synthetic but “realistic” graphs?
  – Configuration models
  – Preferential attachment networks
  – …

Configuration Model

Original Graph (Karate Club)

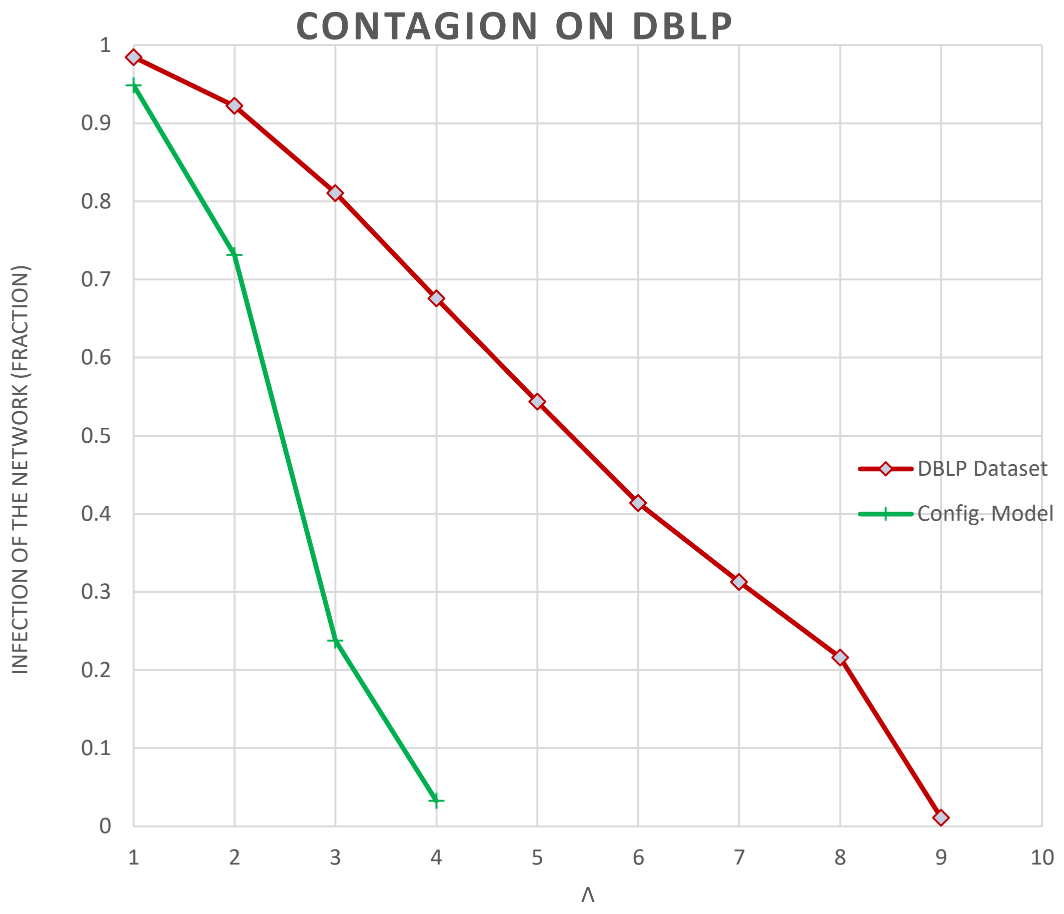


Configuration model



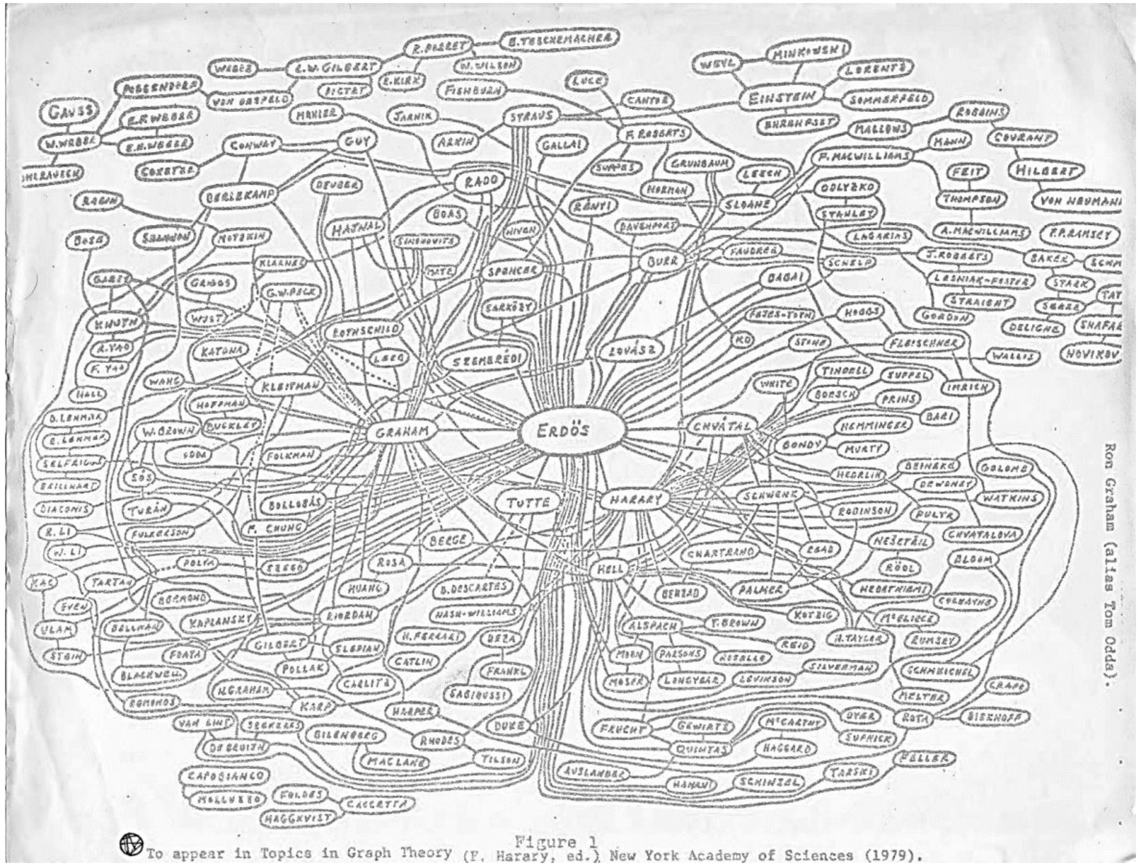
Real Network and Configuration Model

- Graph
  - DBLP
  - Configuration Model
- D: Poisson distribution
- S: The ‘earliest’ 25 nodes



Having better model for DBLP

- Time evolving graphs?
  - A growing network in which newcomers connect to old nodes.



Having better model for DBLP

- Preferential attachment network
  - Add a new node, create **m** out-links to old nodes
  - Connect old nodes with attachment rule $\mathbb{A}$
    - Preferentially with probability α
    - Uniformly random otherwise
- How can we model DBLP by PA?

Having better model for DBLP

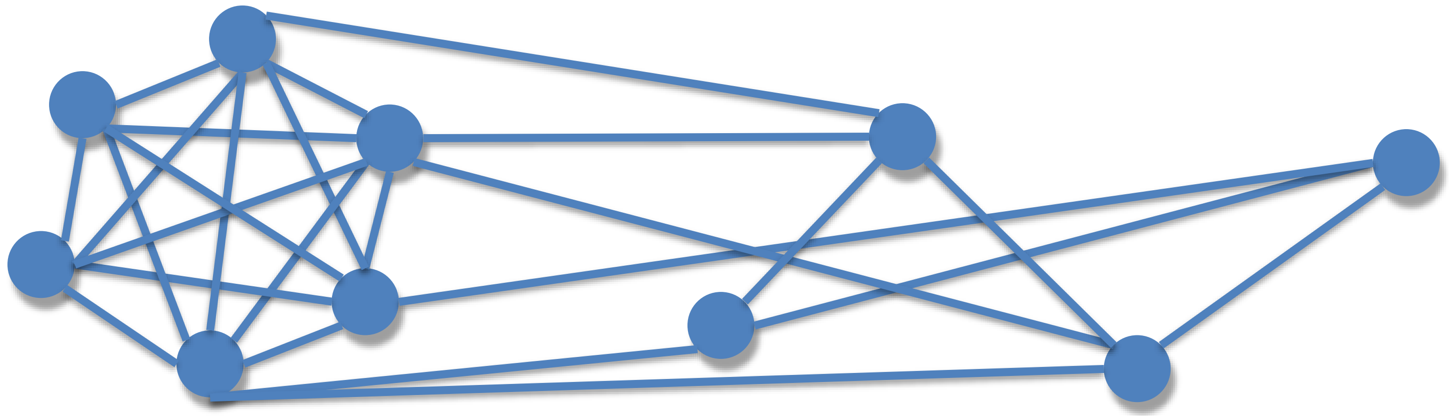
- ~~Preferential attachment network~~
  - Add a new node, create $\cancel{m}$ out-links to old nodes
  - Connect old nodes with attachment rule $\mathbb{A}$
    - Preferentially with probability α
    - Uniformly random otherwise
- How can we model DBLP by PA?

Stochastic Attachment Model (SA)

- Model
  - Add a new node, create m out-links from distribution M to the old nodes
  - Connect old nodes with attachment rule $\mathbb{A}$
    - Preferentially with probability α
    - Uniformly random otherwise

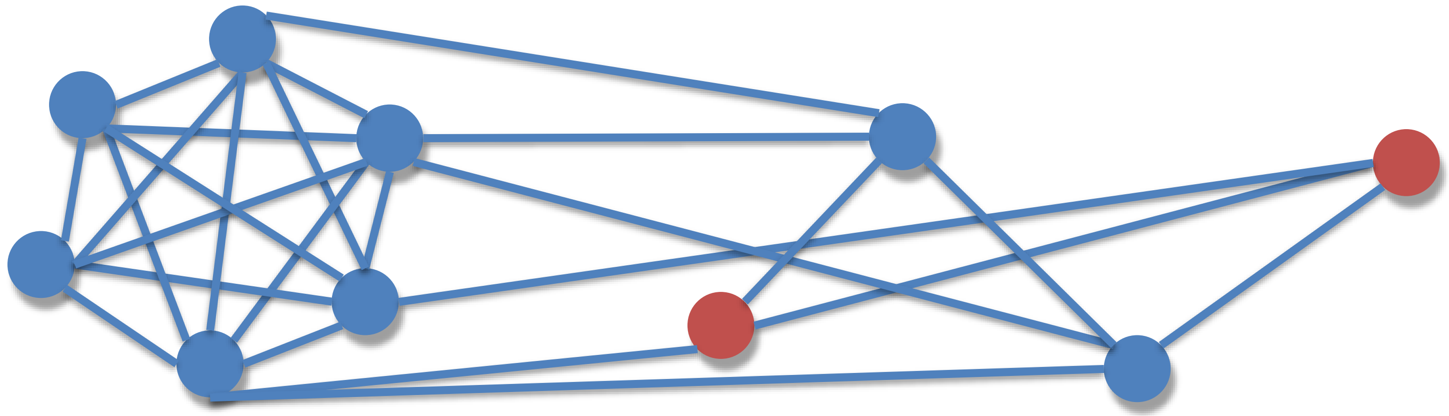
Parameters for the SA

- Learn parameters from real social network
  - Learn M by iteratively remove the minimal degree node



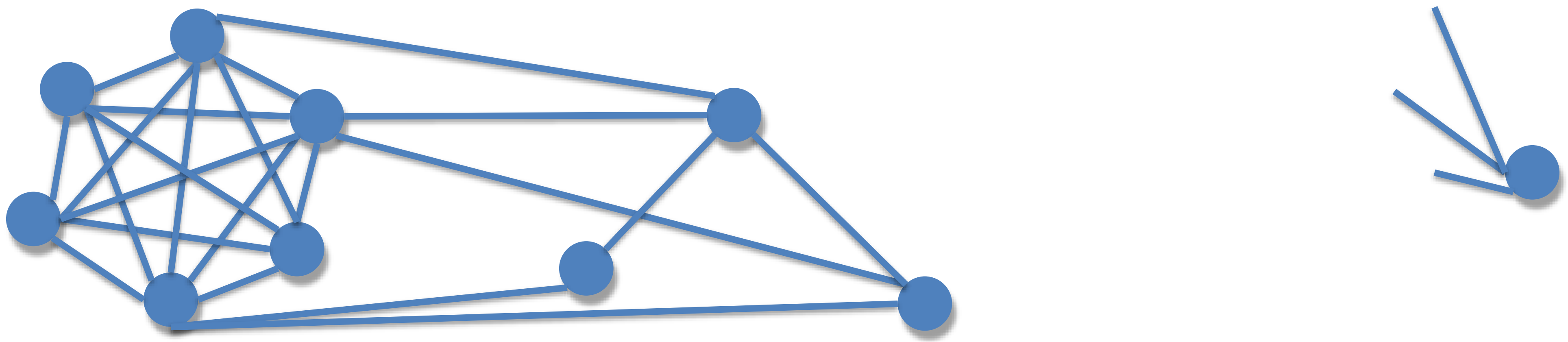
Parameters for the SA

- Learn parameters from real social network
  - Learn M by iteratively remove the minimal degree node



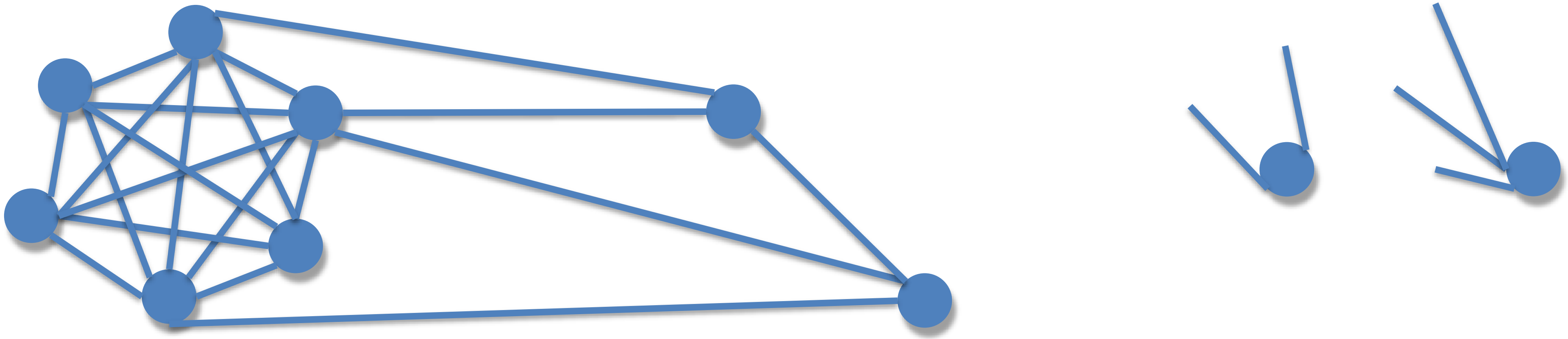
Parameters for the SA

- Learn parameters from real social network
  - Learn M by iteratively remove the minimal degree node



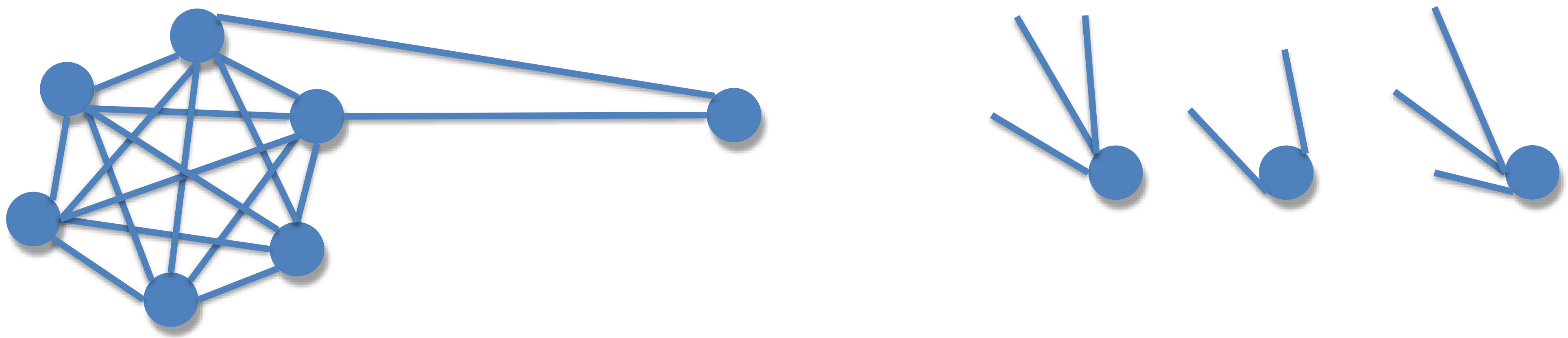
Parameters for the SA

- Learn parameters from real social network
  - Learn M by iteratively remove the minimal degree node



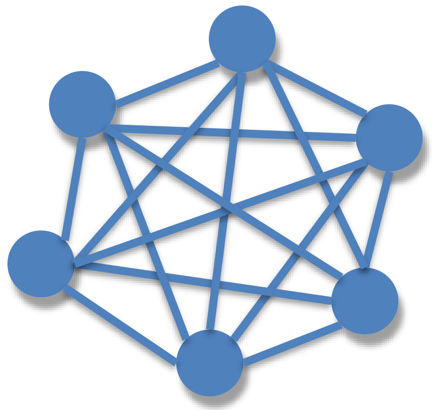
Parameters for the SA

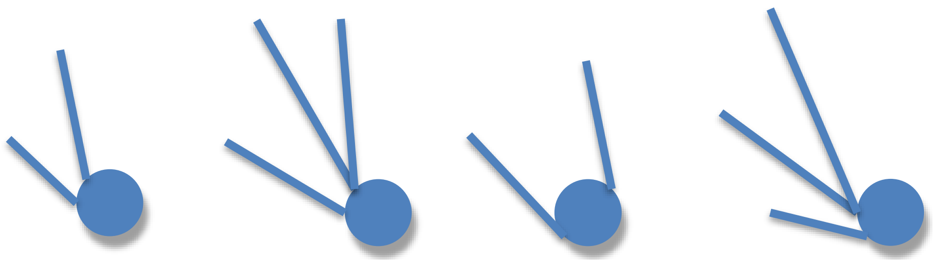
- Learn parameters from real social network
  - Learn M by iteratively remove the minimal degree node



Parameters for the SA

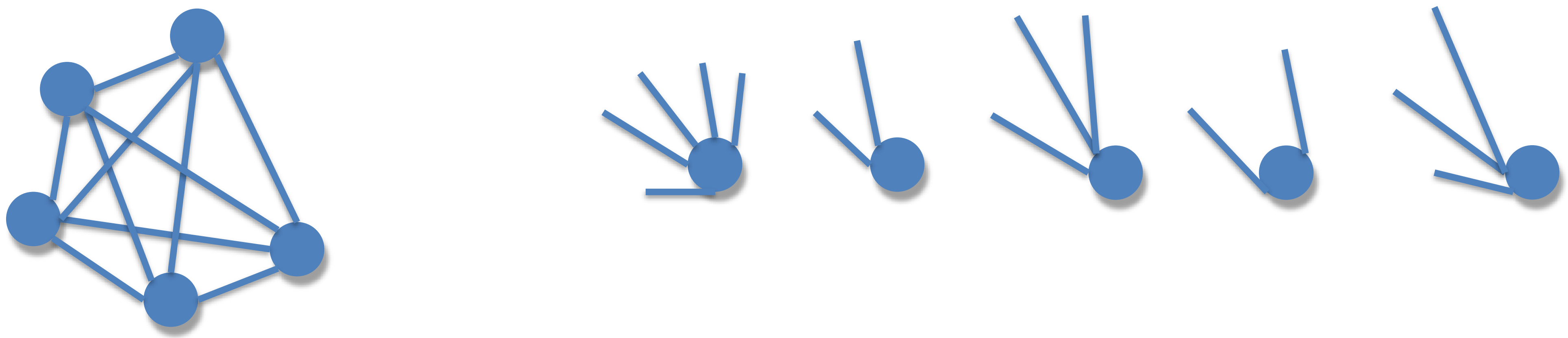
- Learn parameters from real social network
  - Learn M by iteratively remove the minimal degree node





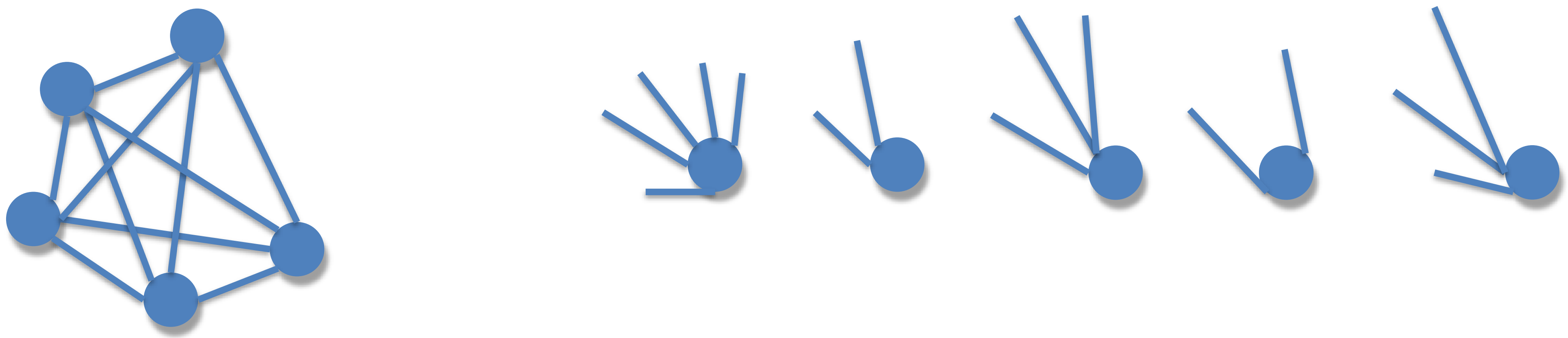
Parameters for the SA

- Learn parameters from real social network
  - Learn M by iteratively remove the minimal degree node



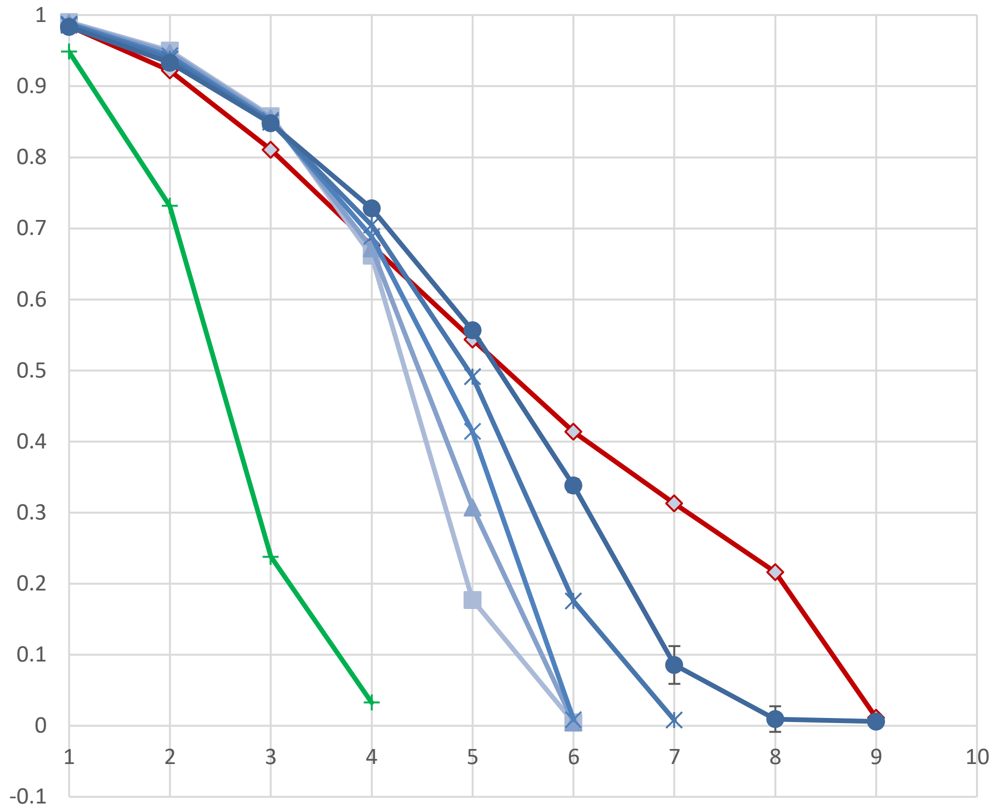
Parameters for the SA

- Learn parameters from real social network
  - Learn M by iteratively remove the minimal degree node
  - Try different α: 0, 0.25, 0.5, 0.75, 1



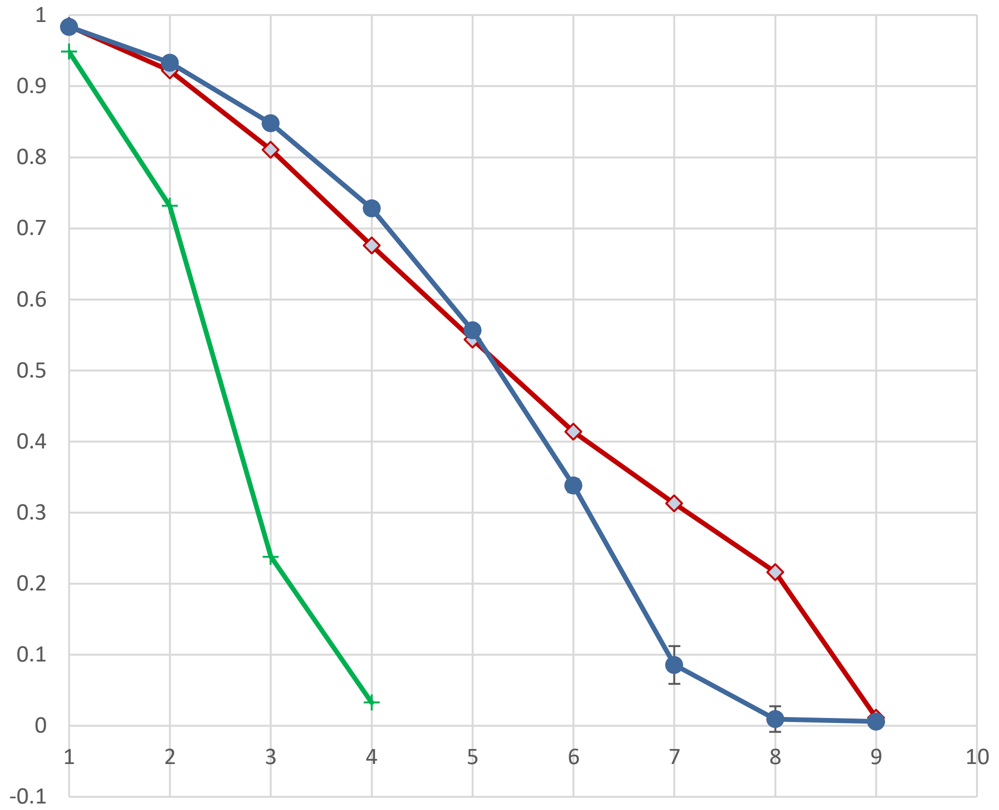
Stochastic Attachment and Contagions

- Graph:
  - DBLP
  - Configuration Model
  - Stochastic Attachment Network
- D: Poisson distribution
- S: The ‘earliest’ 25 nodes



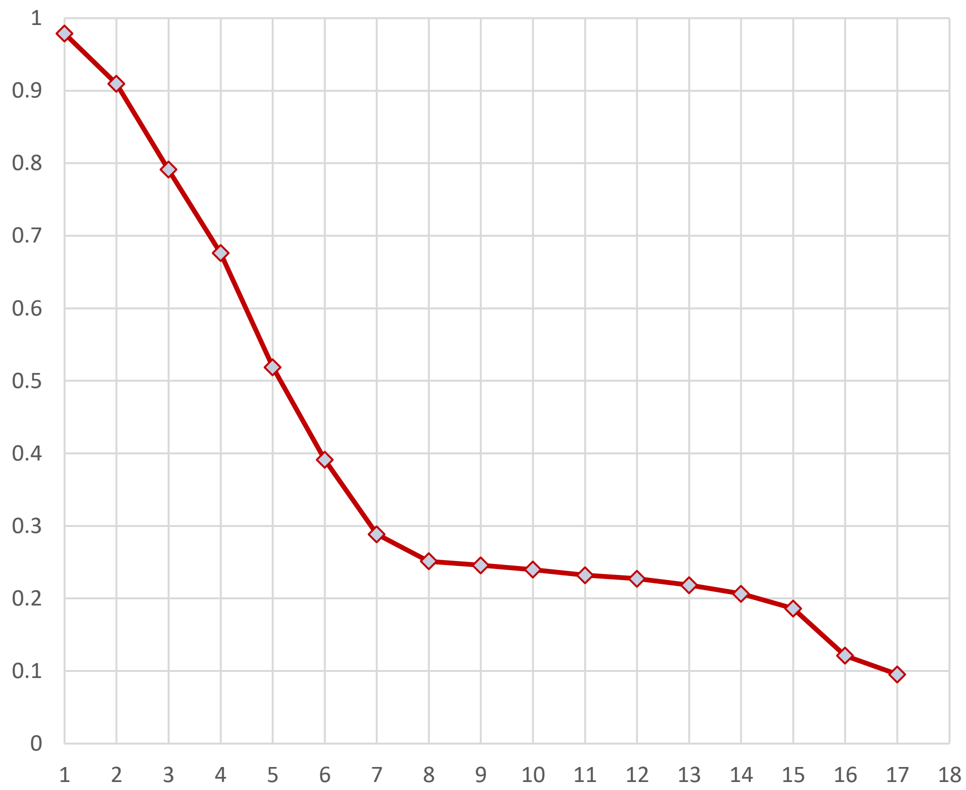
Stochastic Attachment and Contagions

- Graph:
  - DBLP
  - Configuration Model
  - Stochastic Attachment Network
- D: Poisson distribution
- S: The ‘earliest’ 25 nodes



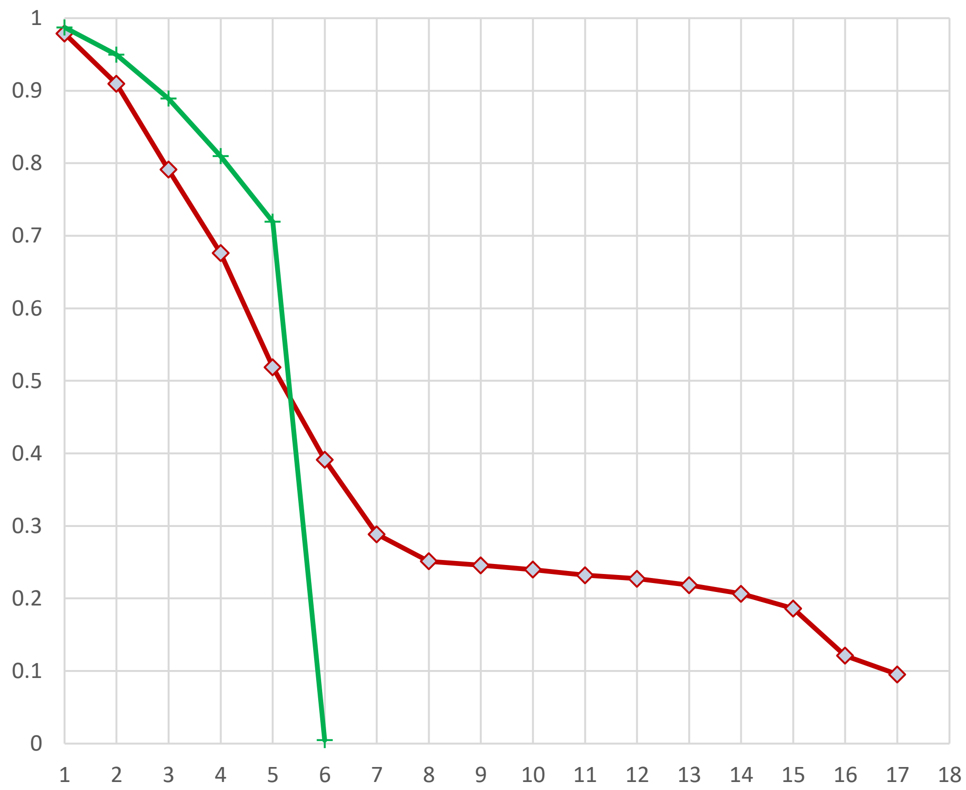
Contagion on Stanford Web Graph

- Graph: Stanford Web Graph
  - 281,903 nodes 2,312,497 edges
  - **7.3** average degree
- D: Poisson distribution
- S: The ‘earliest’ 25 nodes



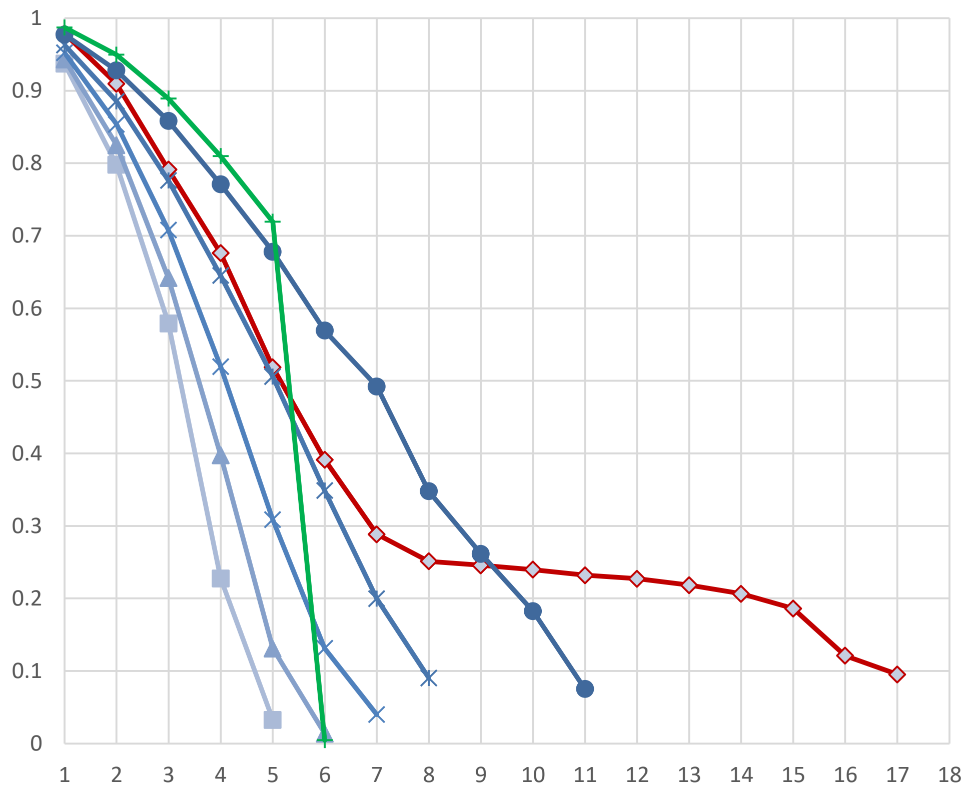
Contagion on Real Network

- Graph
  - Stanford Web Graph
  - Configuration Model
- D: Poisson distribution
- S: The ‘earliest’ 25 nodes



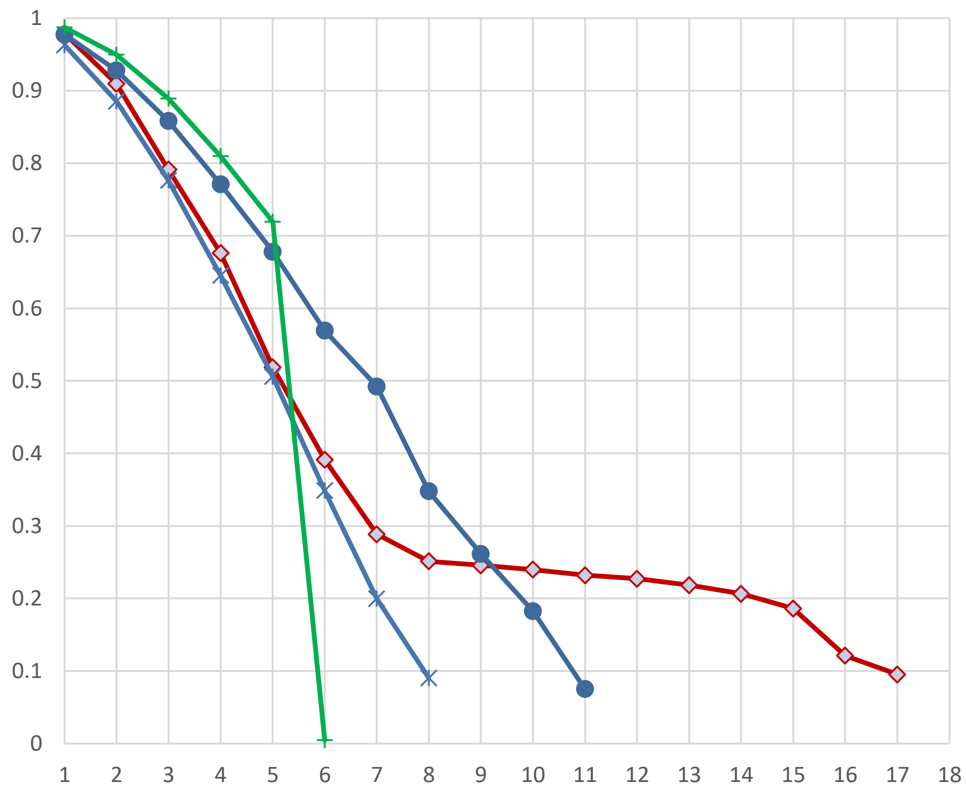
Contagion on Real Network

- Graph
  - Stanford Web Graph
  - Configuration Model
  - Stochastic Attachment Network
- D: Poisson distribution
- S: The ‘earliest’ 25 nodes



Contagion on Real Network

- Graph
  - Stanford Web Graph
  - Configuration Model
  - Stochastic Attachment Network
- D: Poisson distribution
- S: The 'earliest' 25 nodes

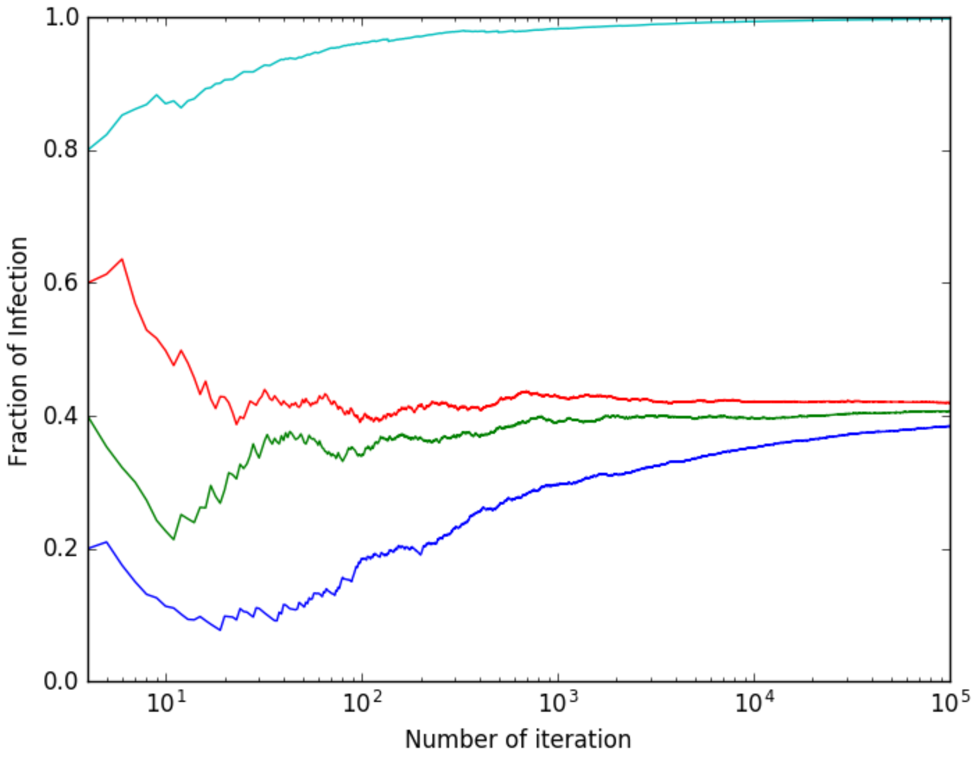


Outline

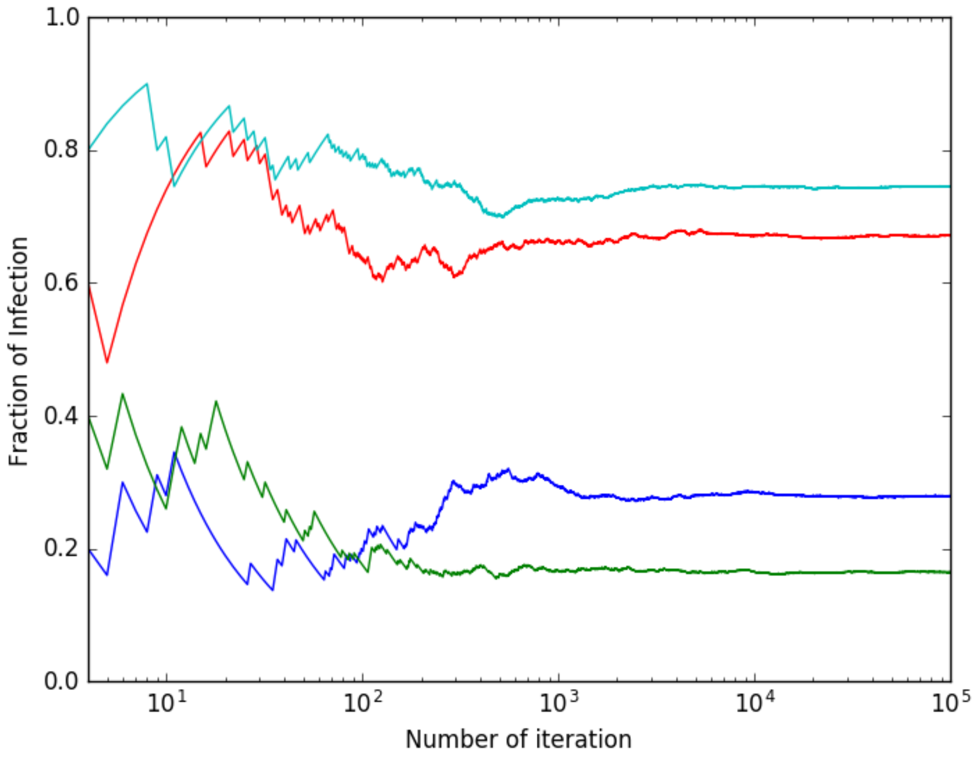
- Cascade Model
- Empirical Results: Testing Network Models
  - Real Data
  - Synthetic Models
    - Configuration Model
    - Stochastic Attachment Model
- Theoretical Results
  - Directed case
  - Undirected case

How would contagion spread on directed PA?

A)



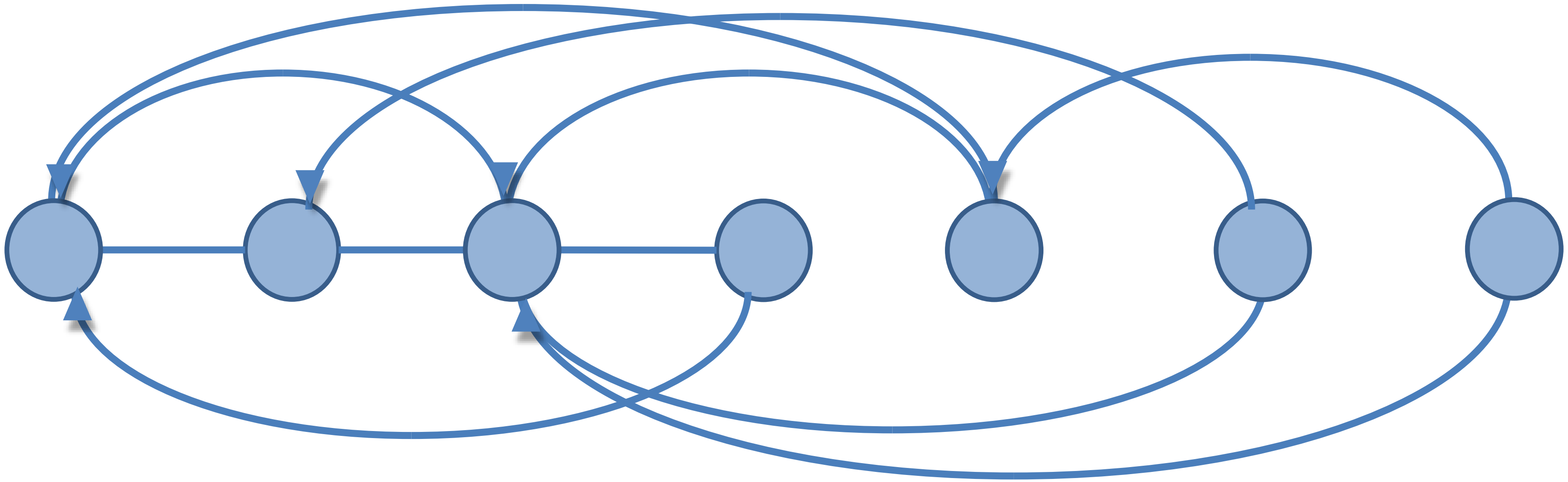
B)



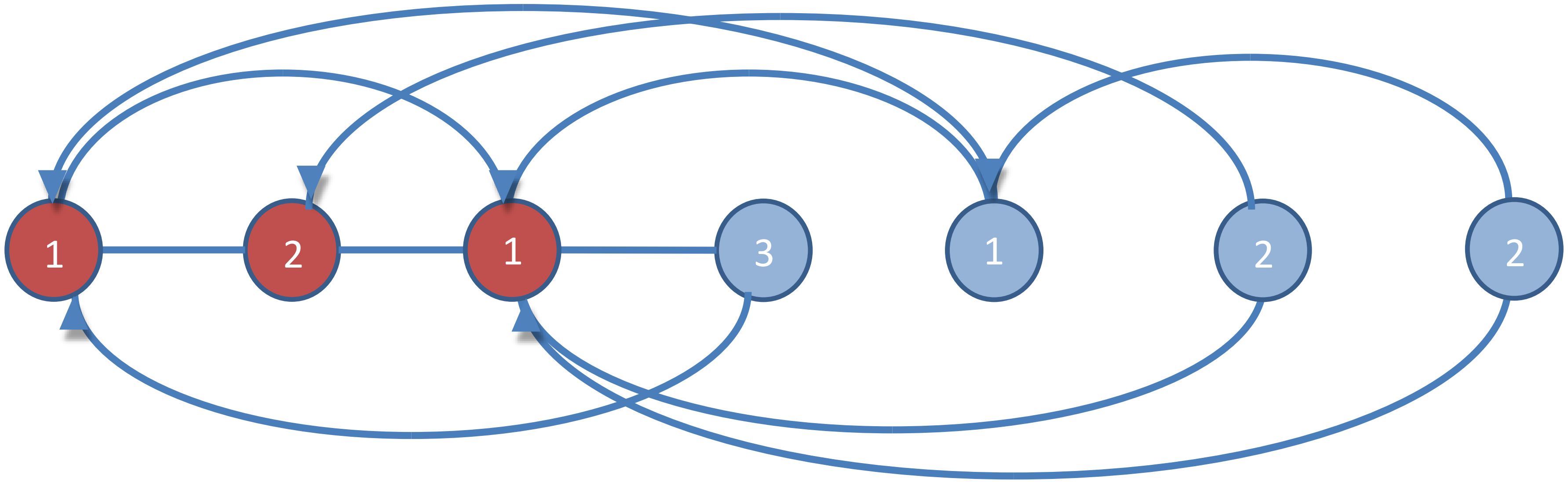
Theorem in Directed Case

- The fraction of infection would converge to the stable fixed points of "feedback function" $f(x)$

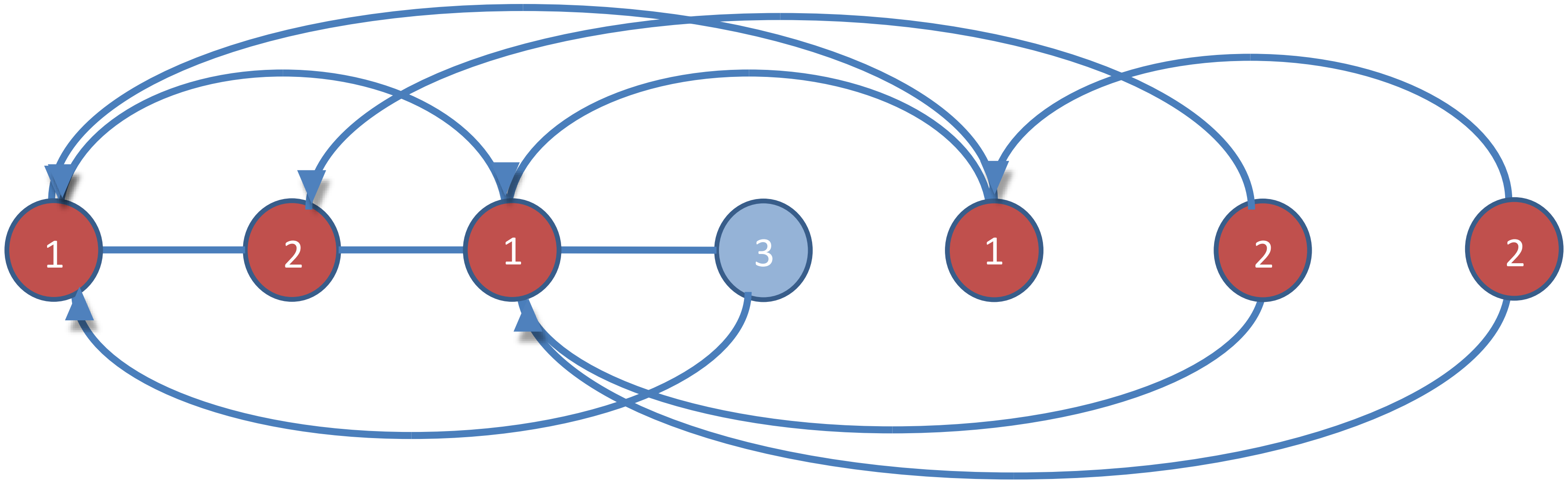
Observations



Observations



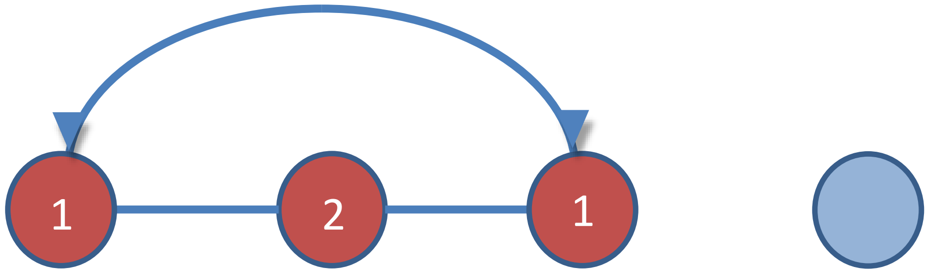
Observations



Observations

- Time evolving property
  - Reveal both the edges and thresholds sequentially

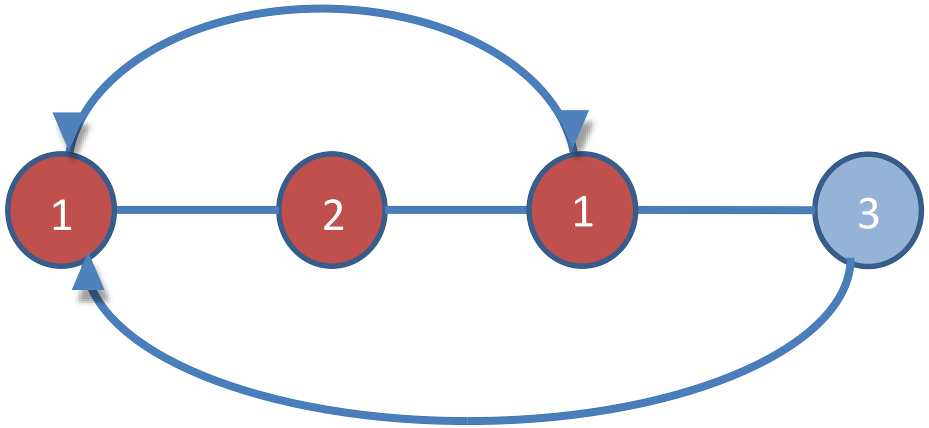
$Y_3 = 1$



Observations

- Time evolving property
  - Reveal both the edges and thresholds sequentially

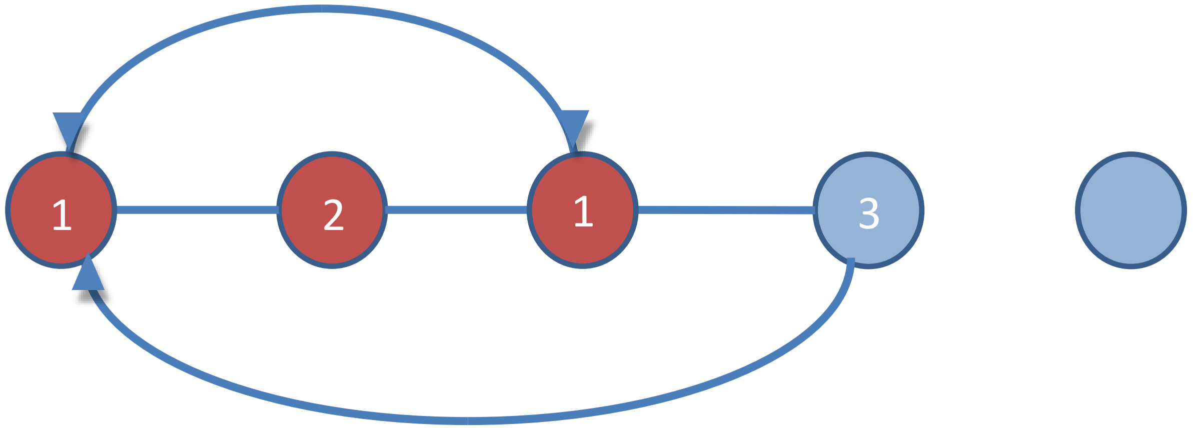
$Y_3 = 1$



Observations

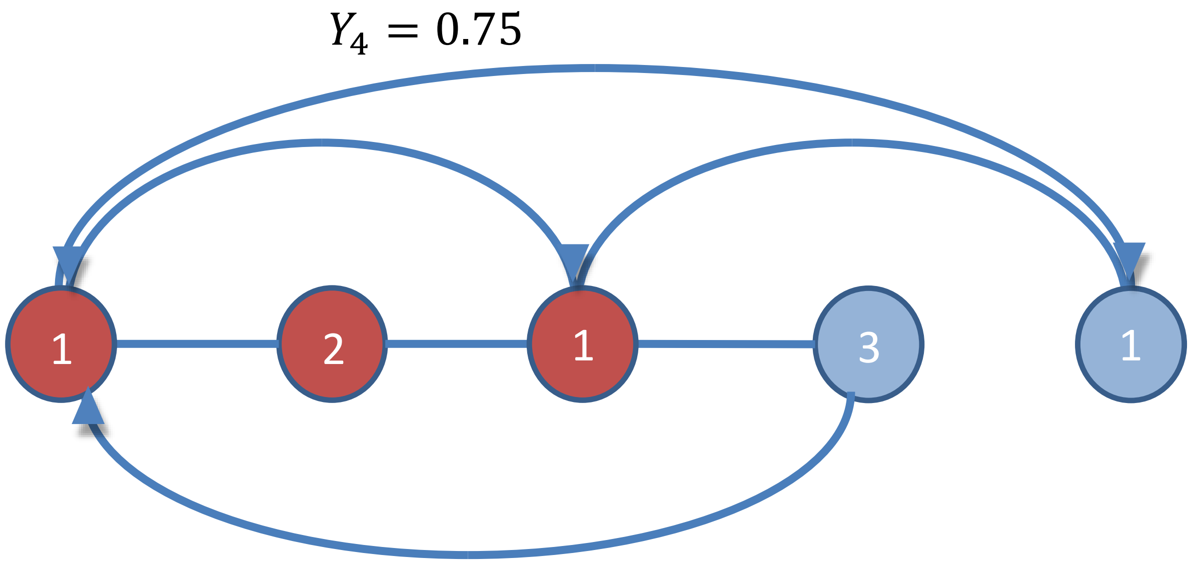
- Time evolving property
  - Reveal both the edges and thresholds sequentially

$Y_4 = 0.75$



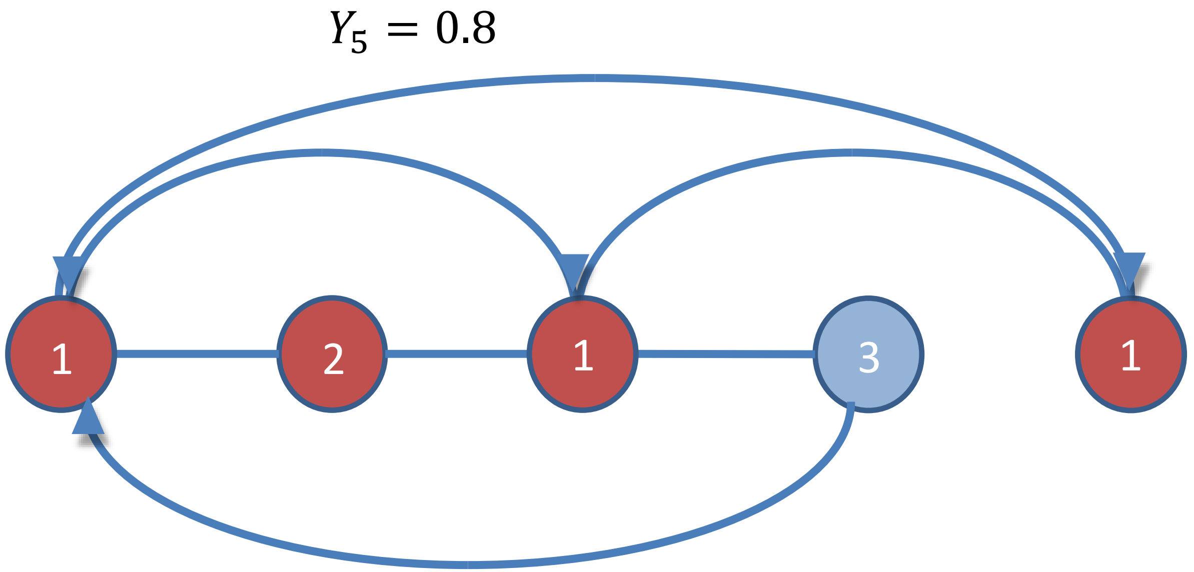
Observations

- Time evolving property
  - Reveal both the edges and thresholds sequentially



Observations

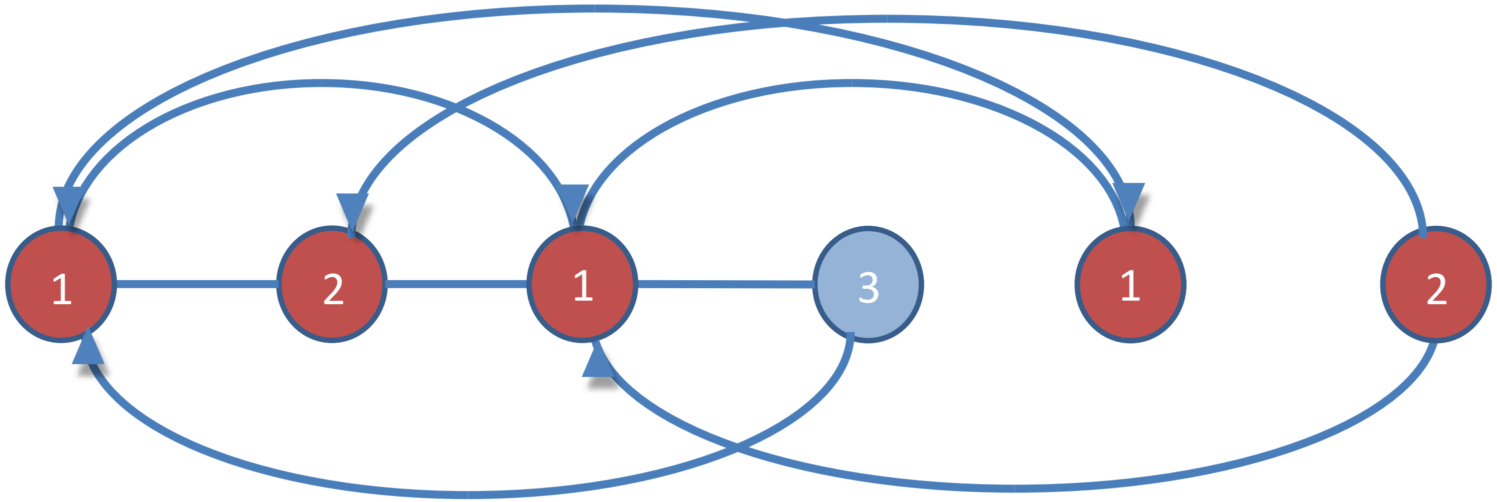
- Time evolving property
  - Reveal both the edges and thresholds sequentially



Observations

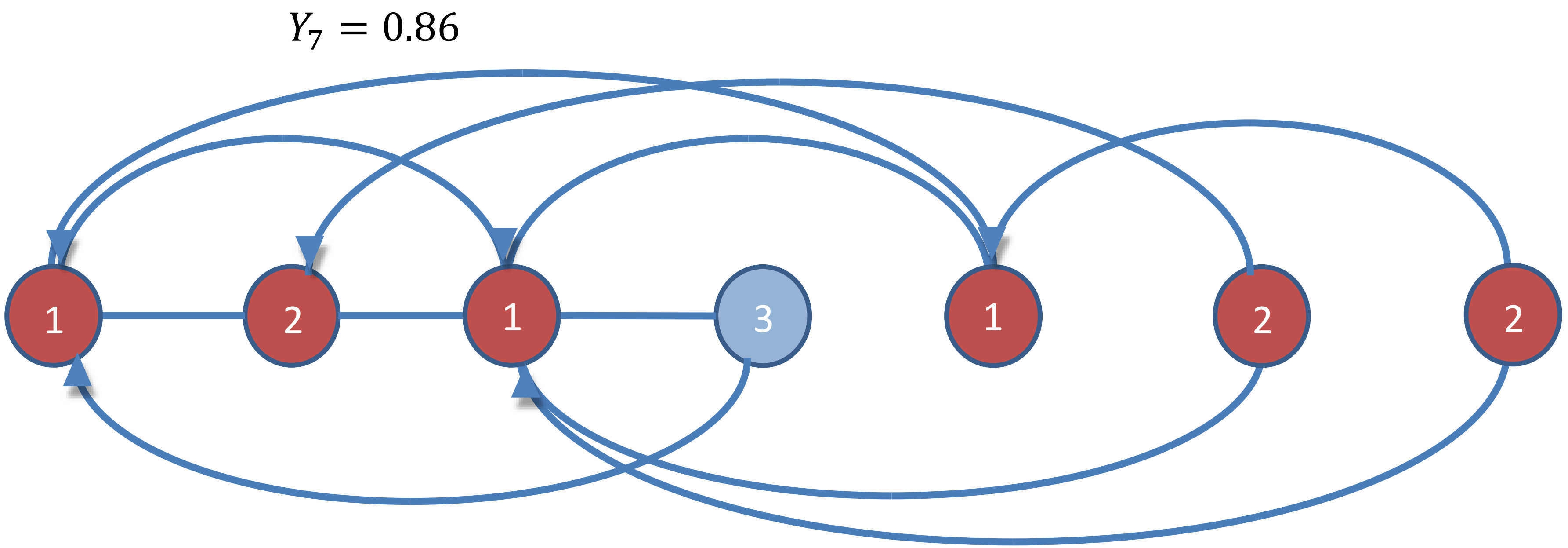
- Time evolving property
  - Reveal both the edges and thresholds sequentially

$Y_6 = 0.83$



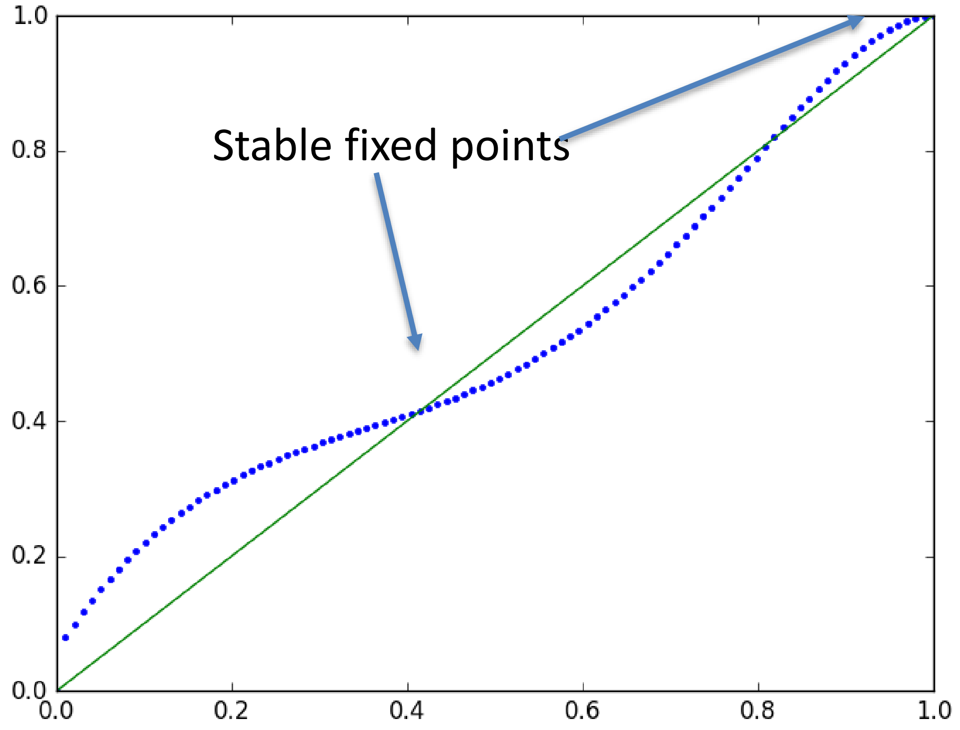
Observations

- Time evolving property
  – Reveal both the edges and thresholds sequentially

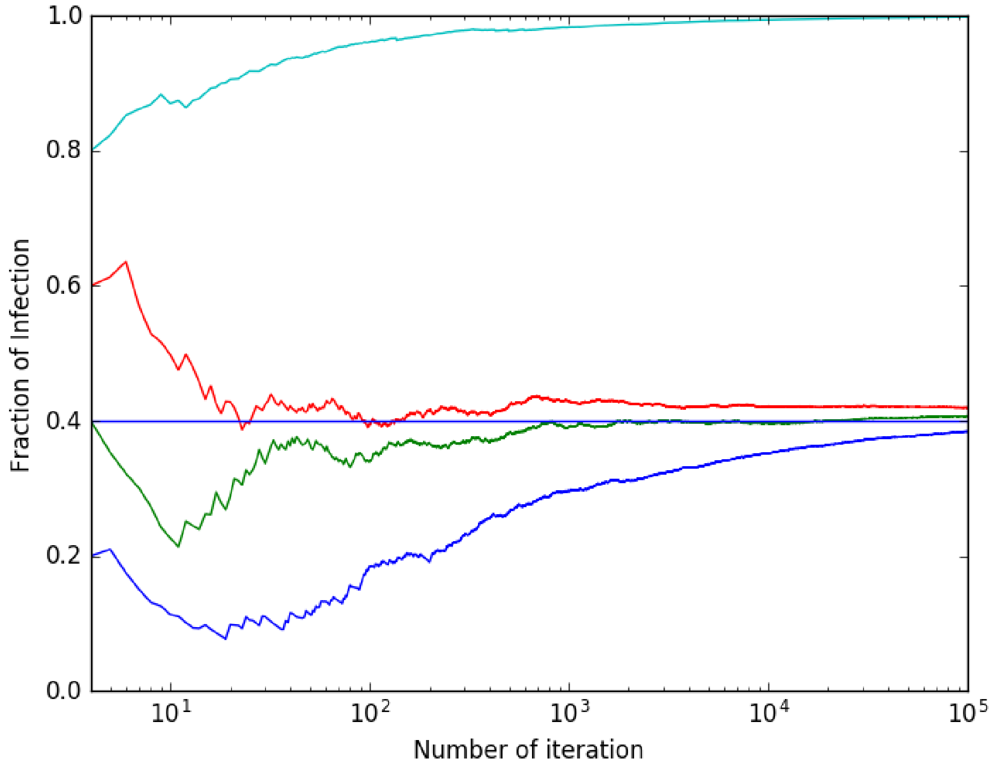


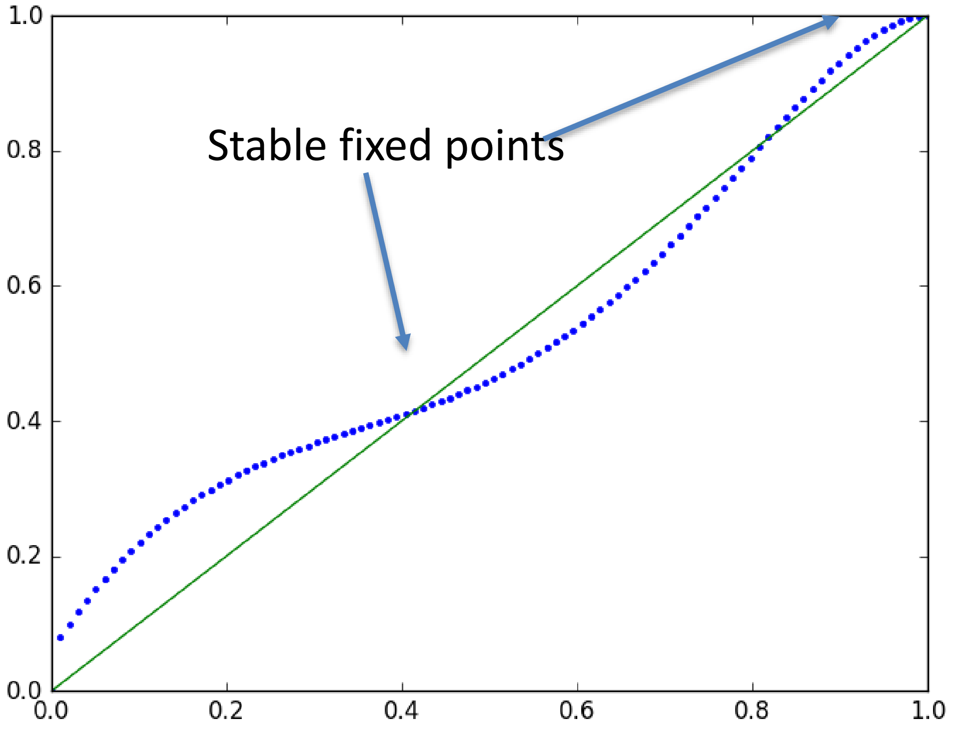
Feedback Function

- The probability of a newcomer get infected
  - Distribution of threshold
  - M out-links



Feedback Function





Outline

- Background and Motivation
- Model and Experimental Results
  - General Threshold Contagion
  - Experiment on Real Network
  - Stochastic Attachment Network
- Theoretical results
  - Directed cases
  - **Undirected cases (please see the paper)**

Future Work

- Better graph models to approximate contagions on real networks
- Unclear when the contagions can die out in undirected case with 0 as a fixed point