

Effective Social Network Quarantine with Minimal Isolation Costs

Huanyang Zheng and Jie Wu

Presenter: [Dongyang Zhan](#)

Dept. of Computer and Info. Sciences

Temple University, USA

Introduction

Notion of diseases

- Human communities (e.g., Ebola)
- Online social networks (e.g., rumours)
- Distributed systems (e.g., computer virus)

Susceptible-Infected-Susceptible

- Two states: **susceptible** and **infected**
- Susceptible people can be infected
- Infected people return to susceptible by recovery

Problem Formulation

Social Network $G = (V, E)$

- V is a set of nodes (people)
- E is a set of directed edges
- C_v is the isolation cost of the node v
- Q is the set of isolated nodes (people)

Objective and Constraint

- Minimize the total isolation cost of $\sum_{v \in Q} C_v$
- Epidemic outbreaks are eliminated

Epidemic Model

Parameters

- λ is a constant infection rate
- r is a constant recover rate
- $p(d)$ is the fraction of nodes with in-degree d
- $f_d(t)$ is the fraction of infected nodes with in-degree d at time t , and $[1 - f_d(t)]$ is the fraction of susceptible nodes with in-degree d at time t

The probability that a uniform-randomly selected edge comes from an infected node at the time t is $\Theta(f(t))$

$$\Theta(f(t)) = \frac{\sum_d d \cdot p(d) \cdot f_d(t)}{\sum_d d \cdot p(d)}$$

Epidemic Model

Consider a node with in-degree d

- It has d incoming neighbors
- $d \times \Theta(f(t))$ infected incoming neighbors (expected)
- Each infected neighbor has a infection rate of λ
- The total infection rate is

$$1 - (1 - \lambda)^{d \cdot \Theta(f(t))} \approx \lambda \cdot d \cdot \Theta(f(t))$$

The epidemic state transfer equation is

$$\frac{\partial f_d(t)}{\partial t} = \lambda d \Theta(f(t)) [1 - f_d(t)] - r f_d(t)$$

The 1st part is infection, and the 2nd part is recovery

Epidemic Model

To control epidemic outbreaks

- The new infection must be 0: $\frac{df_d(t)}{dt} = 0$

Further derivation shows

$$f_d(t) = \frac{\lambda d \Theta(f(t))}{r + \lambda d \Theta(f(t))}$$

$$\Theta(f(t)) = \frac{1}{\sum_d dp(d)} \sum_d dp(d) \frac{\lambda d \Theta(f(t))}{r + \lambda d \Theta(f(t))}$$

$$\frac{\partial}{\partial \Theta(f(t))} \left(\Theta(f(t)) - \frac{\sum_d dp(d) \frac{\lambda d \Theta(f(t))}{r + \lambda d \Theta(f(t))}}{\sum_d dp(d)} \right) \geq 0$$

Epidemic Model

The result to control epidemic outbreak:

$$\frac{\lambda \sum_d d^2 p(d)}{r \sum_d d p(d)} \leq 1 \quad \text{or} \quad \frac{\langle d^2 \rangle}{\langle d \rangle} \leq \frac{r}{\lambda}$$

$\langle \cdot \rangle$ denotes the mean value of variable

$$\frac{\langle d^2 \rangle}{\langle d \rangle} = \frac{\langle d^2 \rangle - \langle d \rangle^2}{\langle d \rangle} + \langle d \rangle$$

Larger average degree and larger degree variance bring more network vulnerability to epidemics

Feasibility and Minimality

Let $\Delta(Q)$ denote the degradation of $\frac{\langle d^2 \rangle}{\langle d \rangle}$
(isolation Q can control epidemic outbreaks)

Objective is to minimize $\sum_{v \in Q} C_v$

Constraint is to control epidemic outbreaks:

$$\Delta(Q) \geq \delta \quad \delta = \frac{\langle d^2 \rangle}{\langle d \rangle} - \frac{r}{\lambda}$$

We focus on feasible isolations:

Definition 1: A quarantine strategy, Q , is said to be feasible, if the constraint of $\Delta(Q) \geq \delta$ is satisfied.

Feasibility and Minimality

Key concept of minimality:

Definition 2: A feasible quarantine strategy, Q , is said to be minimal, if $Q \setminus \{v\}$ is not feasible for an arbitrary $v \in Q$.

Key observation:

Theorem 1: A minimal feasible quarantine strategy, Q , satisfies the property that $\delta \leq \Delta(Q) \leq 2\delta$.

Key intuition:

- A minimal feasible quarantine strategy would not lead to unnecessary isolations.
- Unnecessary isolations are saved once the epidemic outbreak is controlled.

Algorithmic Design

Our problem is NP-hard

- A reduction from partial set cover

Unbounded greedy can iteratively choose the lowest marginal cost-to-benefit node

```
1: Initialize  $Q = \emptyset$ .
2: while  $\Delta(Q) < \delta$  do
3:    $v = \arg \min_{v \in V \setminus Q} \frac{C_v}{\Delta(\{v\} \cup Q) - \Delta(Q)}$ .
4:    $Q = Q \cup \{v\}$ .
5: return  $Q$  as the quarantine strategy.
```

Algorithmic Design

Bounded greedy through homogeneous scaling

```
1: if  $\delta < 0$  then
2:   return  $\emptyset$ ;
3:  $v = \arg \min_{u \in V \setminus Q'} \frac{C_u}{\Delta(Q' \cup \{u\}) - \Delta(Q')}$ .
4: Set coefficient  $\epsilon = \frac{C_v}{\Delta(Q' \cup \{v\}) - \Delta(Q')}$ .
5:  $Q' = Q' \cup \{v\}$ .
6: for each  $u \in V \setminus Q'$  do
7:    $C'_u = \epsilon \times \Delta(\{u\})$ . /* split node cost */
8:    $C_u = C_u - C'_u$ . /* residual node cost */
9:  $Q = Q' \cup \text{RECURSIVE}(G, \delta - \Delta(Q'), Q')$ .
10: for each  $u \in Q$  do
11:   if  $Q \setminus \{u\}$  is a feasible quarantine strategy then
12:      $Q = Q \setminus \{u\}$ .
13: return  $Q$  as the quarantine strategy.
```

Algorithmic Design

Approximation ratio is 2

- At most double the optimal isolation cost
- Insight is that a minimal feasible isolation strategy is close to the optimal isolation strategy

Time complexity is $O(V^2)$

Experiments

Epidemics in online social networks

- Epinions is a general consumer review site
- Wikipedia is a free encyclopedia

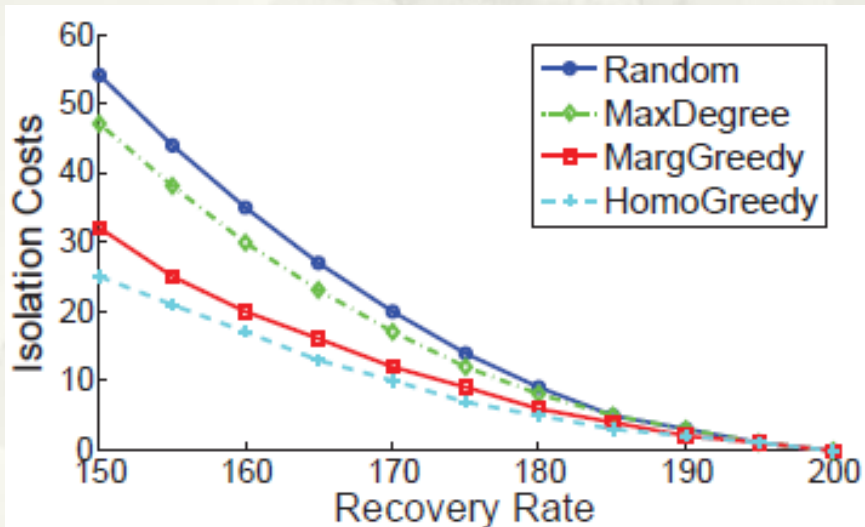
DATASET STATISTICS

	Epinions	Wikipedia
Number of nodes	18,098	7,115
Number of edges	355,754	103,689
Average degree	19.6	14.6
In-degree Variance	3615.8	1006.9
Network Diameter	11	7
Global clustering coefficient	0.138	0.141
Average edge weight	0.0285	0.0076

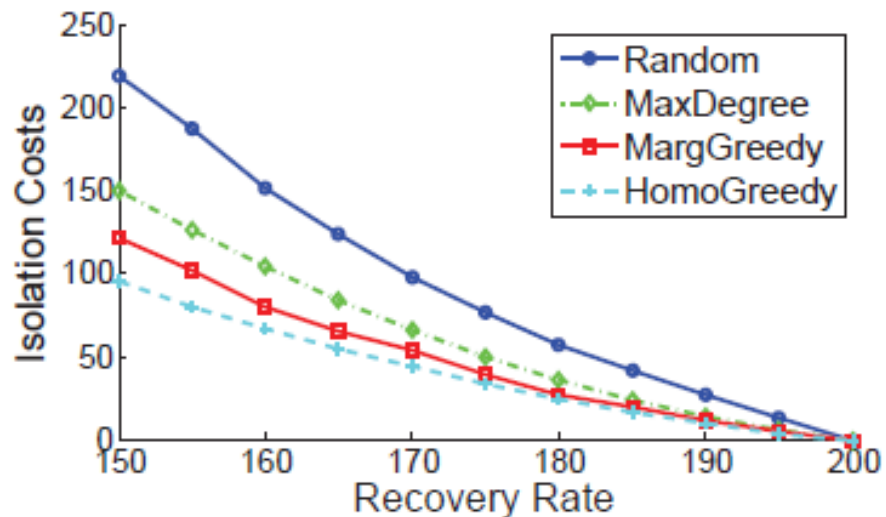
Experiments

Results on Epinions

Isolation cost depends on node degree



(a) Epinions with constant cost.

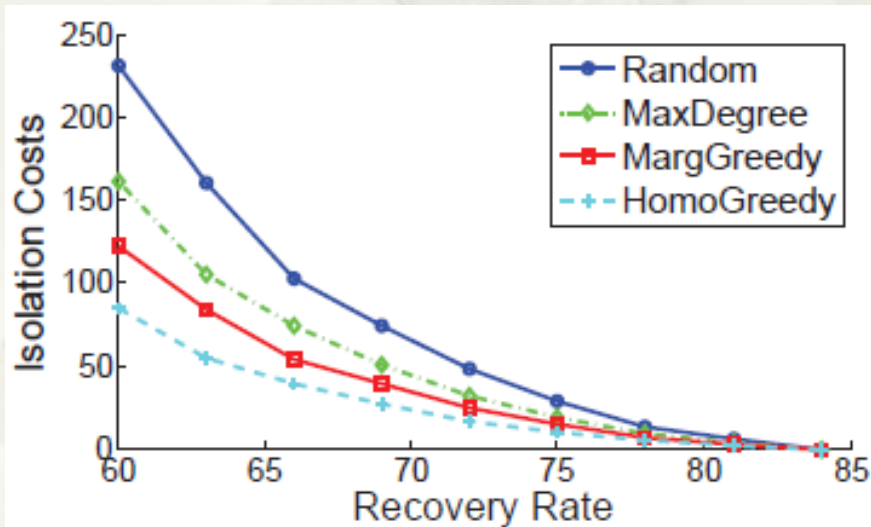


(b) Epinions with logarithmic cost.

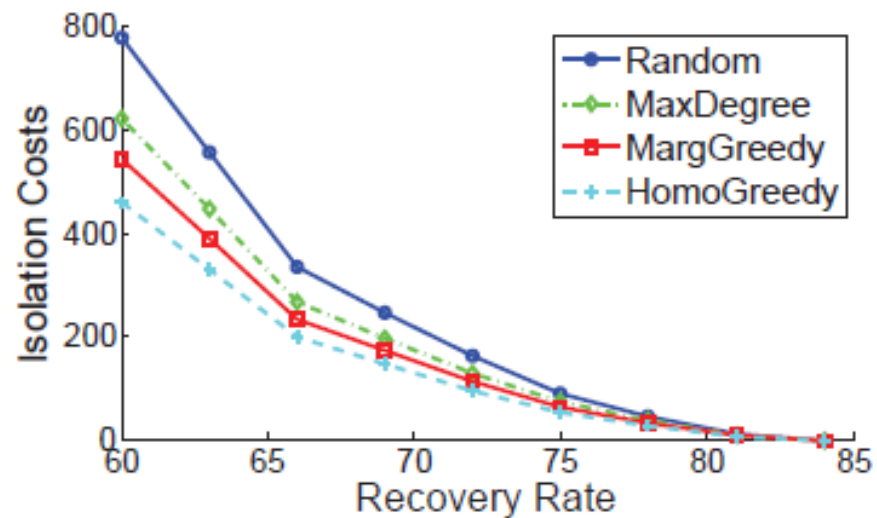
Experiments

Results on Wikipedia

Isolation cost depends on node degree



(d) Wikipedia with constant cost.



(e) Wikipedia with logarithmic cost.

Conclusion

Epidemic outbreak depends on both the average node degree and the degree variance

A minimal feasible quarantine can avoid unnecessary isolations, and lead to a bounded algorithm

Thank You