Hindering Influence Diffusion of Community

Jiadong Xie
East China Normal University
xiejiadong0623@gmail.com
Advisor: Fan Zhang (Guangzhou University),
Xuemin Lin (Shanghai Jiaotong University)

Motivation

The prevalence of social network platforms and the global COVID-19 epidemic leads to more research on influence diffusion models.

Inspired by the work *Phase transitions in information spreading on structured*[1]:

- ❖ there is a transition that separates the local and global rumour spread;
- ❖ the transition point is highly related to the interactions between communities.

In social network area, no existing works consider the community structure in diffusion models.

[1] Jessica T. Davis, N. Perra, Q. Zhang, Y. Moreno, and A. Vespignani. 2020. Phase transitions in information spreading on structured populations. Nature Physics. 590-596.

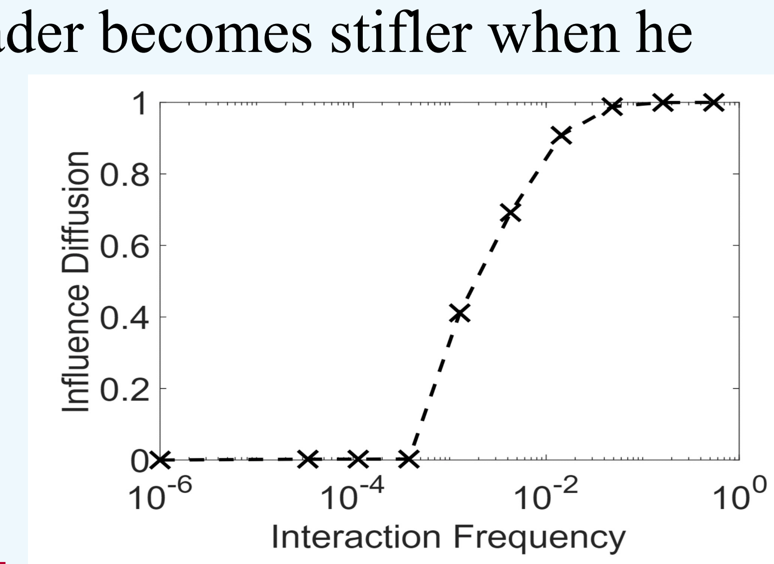
Case Study

Dataset: DBLP network

Community: SIGMOD

Diffusion Model: Maki–Thompson model ($\lambda = 0.1, \alpha = 1.0$).

- ❖ three compartments: ignorants, spreaders, stiflers;
- ❖ λ: the probability that an ignorant become a spreader when it is contacted by a spreader;
- ❖ α: the probability that a spreader becomes stifler when he contacts a stifler or ignorant.



Observations

- ▪ influence diffusion is decreasing with the decrease of IF
- ▪ we can completely prevent rumour spread when IF≈ 0.001.

Approach

We use a binary search to find a optimal solution to LCIF problem, and the search lasts for T rounds.

Binary search ω, each time test if ω' is a feasible IF, i.e., exists a B′ such that $\omega(C\backslash B') \le \omega'$ and $|B'| \le b$.

$$\frac{\sum_{u\in C\backslash B'} d_u^{\notin C}}{\sum_{u\in C\backslash B'} d_u^{+}} \le \omega' \Leftrightarrow \sum_{u\in C\backslash B'} (d_u^{\notin C} - \omega' d_u^{+}) \le 0$$

Test through sort all nodes in C in decreasing $(d_u^{\notin C} - \omega' d_u^{+})$, and then try to add them into B' until $\sum_{u\in C\backslash B'}(d_u^{\notin C} - \omega' d_u^{+}) \le 0$.

Time complexity: $O\left(T \cdot |V(C)| \log|V(C)| + \sum_{u\in V(C)} d_u^{+}\right)$.

Error: $|\omega(C\backslash \mathrm{B}) - \omega(B\backslash \mathrm{B}^*)| \le 2^{-T}$

- ❖ B^*: the optimal solution of LCIF problem;
- ❖ B: our solution.

LCIF Problem

Input: a directed graph $G = (V, E)$, a community $C \subseteq V$, and a budget $b (b < |C|)$;

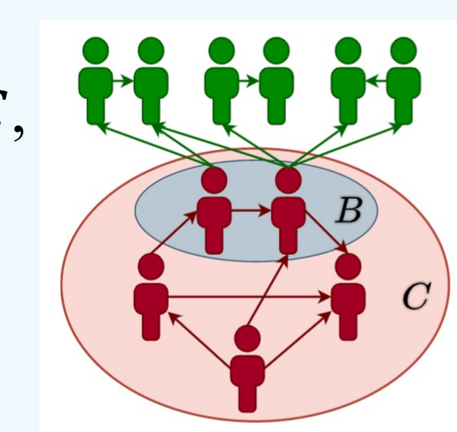
Output: A set B of at most b nodes such that remove it will get minimal interaction frequency.

Interaction frequency (IF): $\omega(C) = \frac{\sum_{u\in C} d_u^{\notin C}}{\sum_{u\in C} d_u^{+}}$.

- ❖ C: The community;
- ❖ $d_u^{\notin C}$: the cross-community out-degree of u, i.e., $d_u^{\notin C} = |\{\langle u, v\rangle \in E | u \in C \land v \notin C\}|$;
- ❖ d_u^{+}: The out-degree of a node u, i.e., $d_u^{+} = |\{\langle u, v\rangle \in E\}|$.

Example

A rumour is diffusing in the community C, and it may affect ignorant people (green) via directed edges. We propose to remove B to hinder the rumour propagation.



Related Work

Rumour Diffusion Model: Different influence diffusion models to simulate the rumour propagation.

Influence Minimization Problem: The problem aims to minimize the expected probability of influence diffusion though remoing nodes or edges.

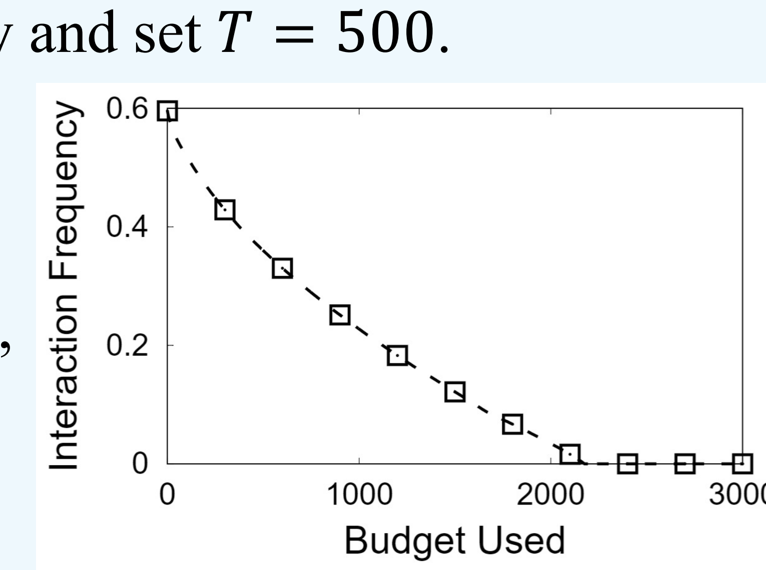
Influence Maximization Problem: The problem searches for a seed set of fixed size that can maximize the expected probability of influence diffusion.

Experiment

Dataset: DBLP network from SNAP. 317,080 nodes, 1,049,866 edges, and 8,734 groud-truth community.

We use SIGMOD community and set $T = 500$.

- ❖ **Initial IF** of SIGMOD:
  - ❖ 0.596.
- ❖ When budget $\mathrm{b} = 2177$, 42.1% nodes in SIGMOD, **IF drop to 0**.
- ❖ **Average running time**:
  - ❖ 2.95ms.



Future Work

- ♦ Combine the methods from the algorithms of the influence minimization and maximization problem, and try to hinder the rumour influence diffusion under transition by fewer nodes.