On the Efficiency of the Information Networks in Social Media

Mahmoudreza Babaie, Przemyslaw A. Grabowicz
Isabel Valera, Krishna P. Gummadi, Manuel Gomez-Rodriguez

1. Information Network

- Twitter
- Facebook
- Google Scholar
- Google +

2. Motivation

Social media users follow multiple other users to receive relevant information:
- 30% of Twitter users follow 50+ other users
- 50% of Twitter users receive 500+ tweets/day

Problems:
1. Difficulties in managing large number of followers/subscriptions
2. Overload with redundant information
3. Delays in arrival of information

3. Key research question

- How efficient are users at selecting their information sources?
- How can we improve the efficiency of users?

4. Definitions of efficiency

To compute an efficiency we compare the original set of followees with a corresponding optimal set that provides the same pieces of information.

1. Link-optimal set contains the smallest number of users.

2. Inflow-optimal set provides the least amount of tweets per time unit.

3. Delay-optimal set provides the contagions as early as possible.

5. Efficiency of users in real world

- Users acquire information sub-optimally (efficiencies<1)
- Users tend to be less efficient at acquiring popular pieces of information

We introduce a fast heuristic method that cross-optimizes both inflow and delay efficiencies at once, using a greedy algorithm solving weighted set cover problem.

Improved by 80%!
Improved by 40%!

6. Structure of optimized users’ networks

The optimized information networks cannot be discovered via triangle closure (creating links to friends of friends)

7. Future work

- Could we creating a relink recommender system for real information networks that improves various efficiencies of users?
- Does user efficiency change over time? What is its relation with user activity?
- Compare user efficiencies across different existing information networks