Friend Recommendation in Online Social Networks: Social Influence Maximization

Huanyang Zheng and Jie Wu

Center for Networked Computing
Dept. of Computer and Info. Sciences
Temple University

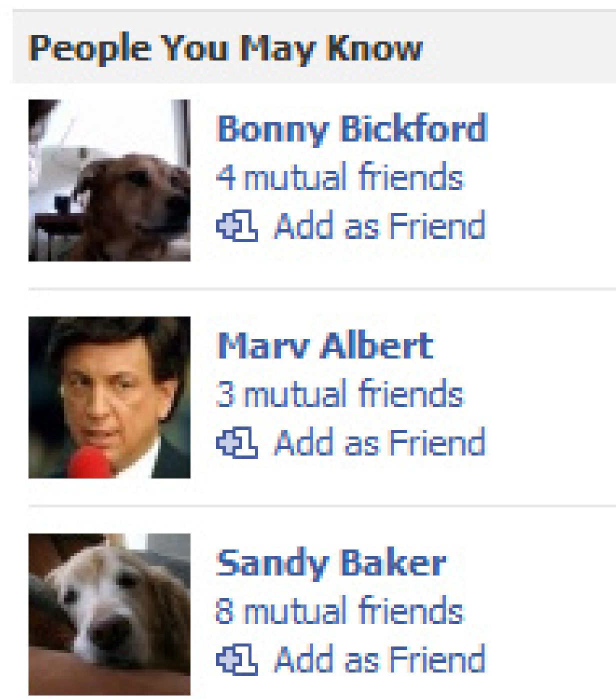
Road Map

- 1. Introduction
- 2. Model and Formulation
- 3. Analysis and Algorithms
- 4. Experiments
- 5. Conclusion

1. Introduction

Online Social Networks (OSNs)

- Social relations among users who share social stories
- 74% of online adults use OSNs
- Examples: Facebook and Twitter



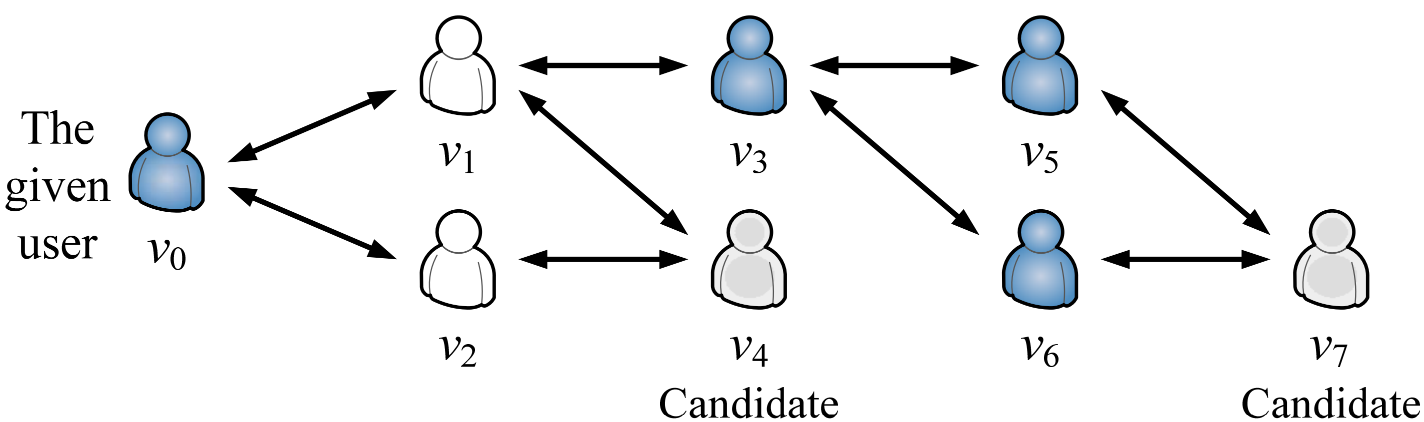
Friend Recommendation System

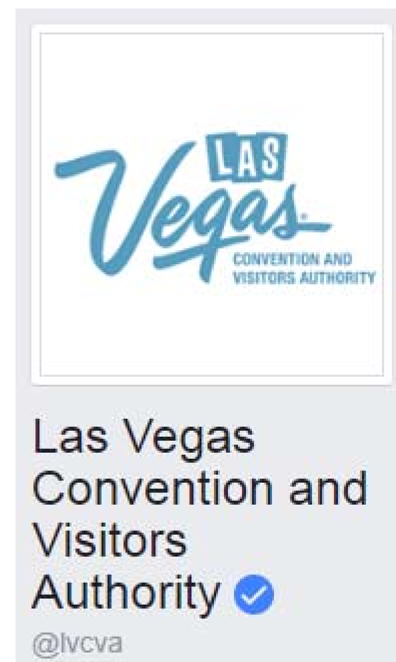
- An essential component in OSNs
- Facebook: friend-of-a-friend

1. Introduction

Friend Recommendation System

- We aim to recommend friends to maximize influence
  - Friend acceptance probability
  - Influence propagation capability
- Applications: advertisements for salesman
- Example: Facebook business page service





2. Model and Formulation

Social Influence Propagation

- Existing independent cascade model
  - On a weighted directed graph (round-by-round)
- Starts with one seed v_0
- When a node v first becomes influenced, v will influence its neighbors
  - Influence probability depends on edge weight
  - One-time influence from v
- Terminates when no more nodes are influenced

2. Model and Formulation

Social Influence Propagation

Definition: The expected number of nodes eventually-influenced by seed nodes is defined as the influence spread.

- Compute influence spread is NP-hard for given seeds
  - Monte-Carlo simulations (time-consuming)
  - Graph reduction (under-estimate: tree, DAG, our)
  - Iterative exchange (over-estimate: IMRank)

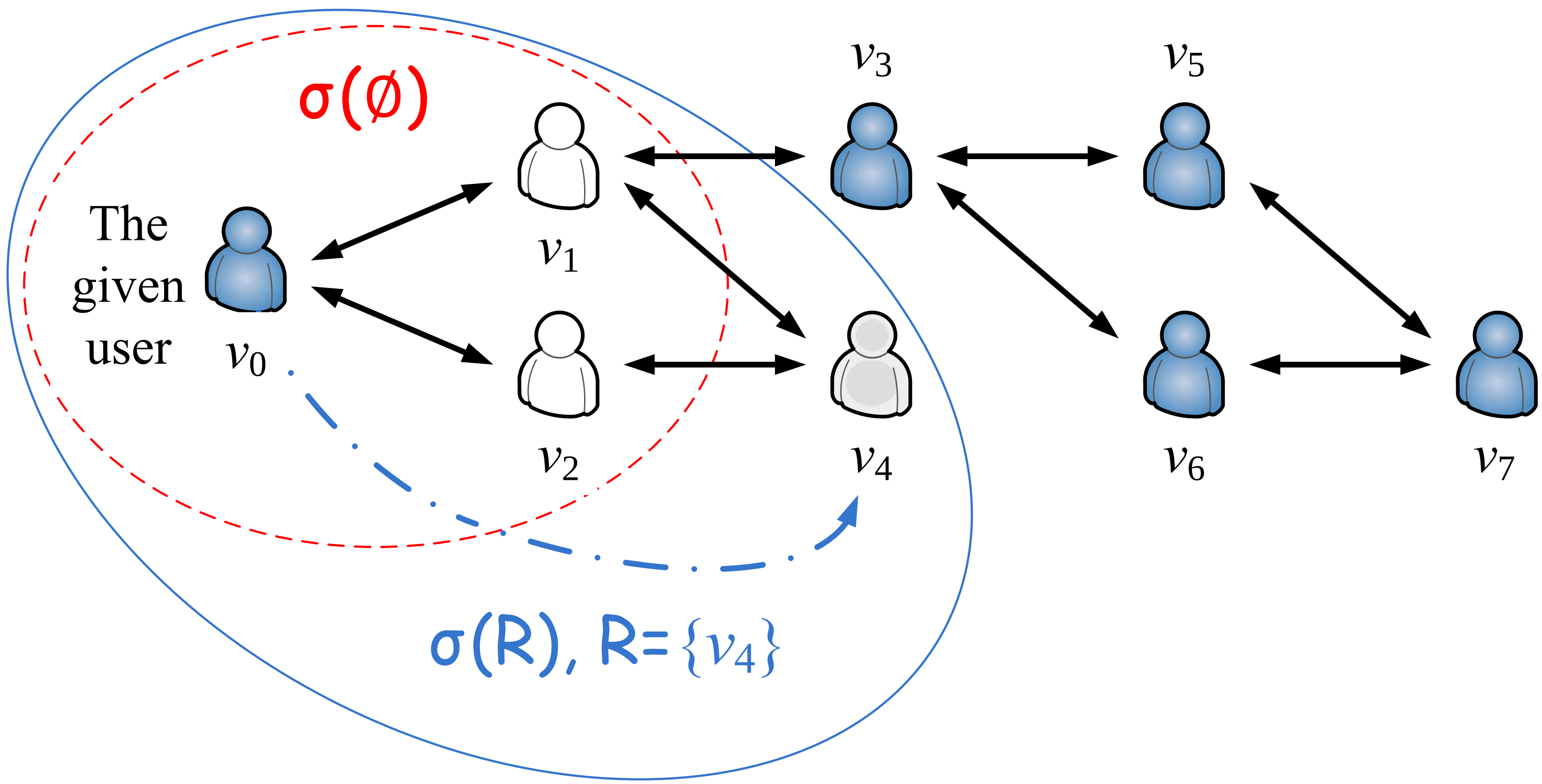
2. Model and Formulation

Problem formulation

- Given knowledge
  - Social network G (weighted & directed)
  - User v_0 and constant k (# of new recommendations)
- Maximize the influence spread gain of v_0
  - Through k new recommended friends
  - Influence spread gain over existing: $\sigma(R) - \sigma(\emptyset)$
- NP-hardness
  - Reduction from maximum coverage problem

2. Model and Formulation

Problem formulation



3. Analysis and Algorithms

Friend Acceptance Probability

- Justifications for probability prediction
- Let f_{vu} be predicted friend acceptance probability

$$f_{vu} = \alpha_u \times \frac{|N(v) \cap N'(u)|}{|N(v) \cup N'(u)|} + \beta_u$$

- N and N' are sets of outgoing & incoming neighbors
- α_u and β_u are constant coefficients for u

3. Analysis and Algorithms

Submodular Property

- Diminishing return effect of recommendation
  - R is set of recommended friends, $\sigma(R)$ is influence spread gain

Theorem: $\sigma(R)$ is submodular with respect to R,
i.e., $\sigma(R \cup \{v\}) - \sigma(R) \geq \sigma(R' \cup \{v\}) - \sigma(R')$ for $R \subseteq R'$.

  - Kempe et al. "Maximizing the spread of influence through a social network." *ACM SIGKDD*, 2003

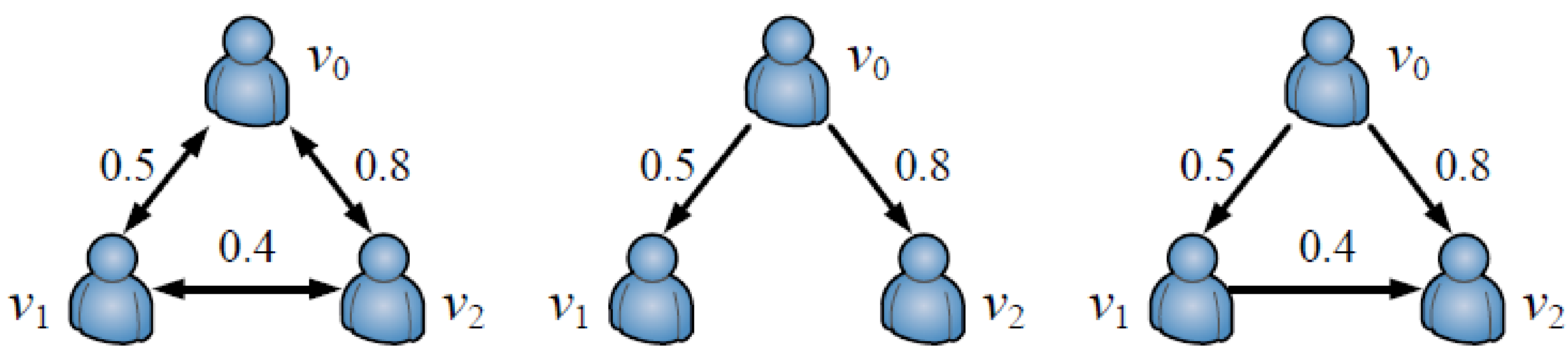
Alg 1: greedy approximation (ratio: 1-1/e)

- Iteratively recommend the friend that maximizes the marginal gain of v's influence spread

3. Analysis and Algorithms

Influence Spread Computation

- NP-hard to compute influence spread for given seeds
- Graph reduction approach
  - Under-estimate influence spread
  - Chen et al. "Scalable influence maximization for prevalent viral marketing in large-scale social networks." ACM SIGKDD, 2010.



(a) Original graph. (b) Tree reduction. (c) DAG reduction.

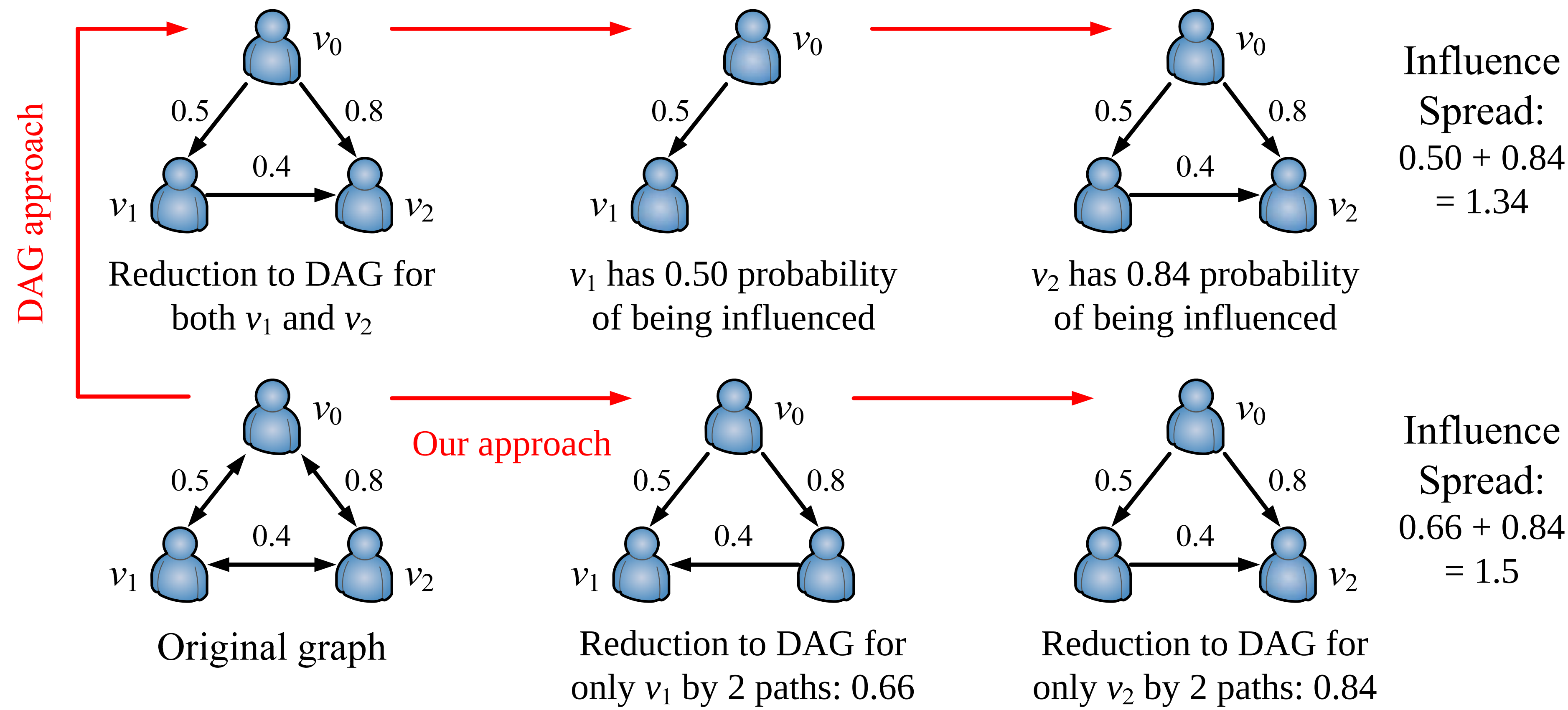
3. Analysis and Algorithms

Alg 2: influence spread computation

- Key idea: consider more paths in polynomial time
  - Estimate a number of needed paths for v, based on graph structure
  - Use Yen's algorithm to obtain a set of loop-free shortest paths from v_0 to v
  - Construct a directed acyclic graph (based on the set of loop-free shortest path) to compute v's probability of being influenced by v_0

3. Analysis and Algorithms

Example: DAG reduction v.s. our approach



Note: 1-(1-0.5)(1-0.4x0.8) = 0.66

4. Experiments

Real data-driven experiments

- Facebook, Epinions, and Wiki

	Facebook	Epinions	Wiki
Number of nodes	63,731	18,098	7,115
Number of edges	817,035	355,754	103,689
Average degree	25.6	19.6	14.6
Network Diameter	15	11	7
Global clustering coefficient	0.148	0.138	0.141
Average edge weight	0.0271	0.0285	0.0076

- Parameter settings
  - f_{vu} is based on $\alpha_u = 0.9$ and $\beta_u = 0.1$
  - Given user is randomly selected (averaged 1,000 times)

4. Experiments

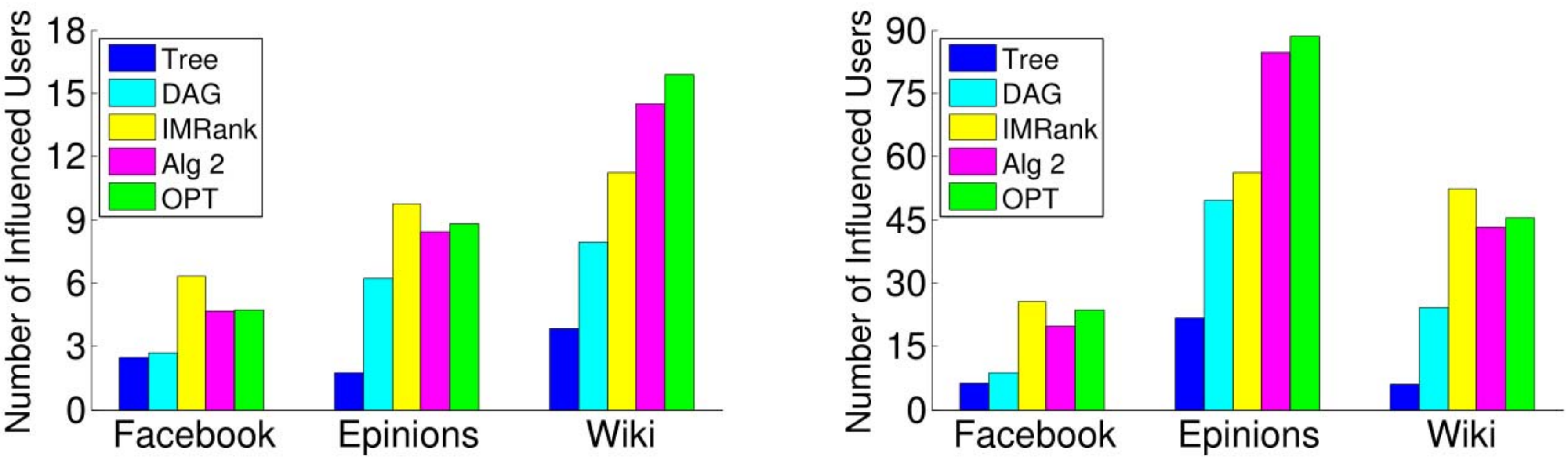
Comparison algorithms (influence spread)

- OPT: the influence spread through time-consuming Monte-Carlo simulations
- Alg 2: our influence spread computation through multiple influence propagation paths
- Tree: graph reduction to a tree to compute the influence spread
- DAG: graph reduction to directed acyclic graph to compute the influence spread
- IMRank: influence spread through iterative neighborhood exchanges

4. Experiments

Results (influence spread)

- Popular user: more than 100 friend (otherwise normal users)



(a) Normal users.

(b) Popular users.

- Tree and DAG underestimate the influence spread
- IMRank may overestimate the influence spread

4. Experiments

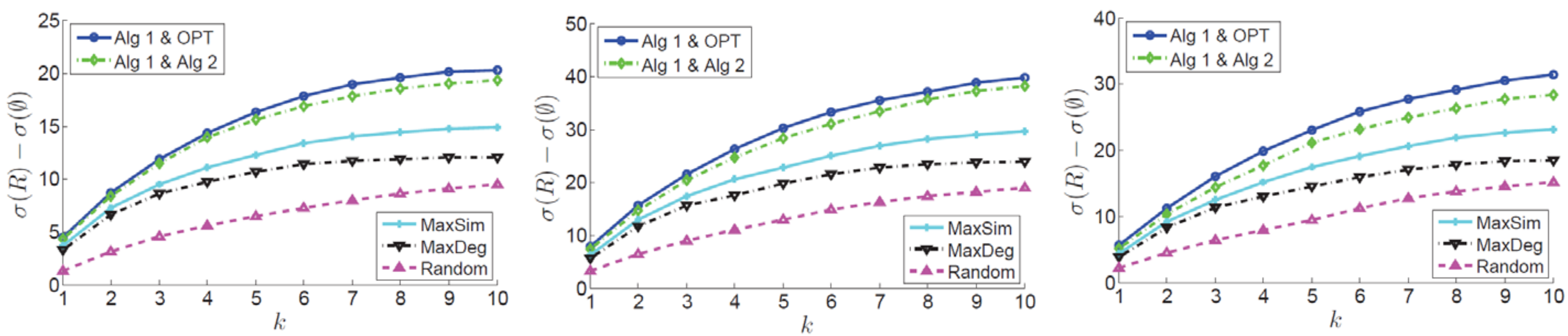
Comparison algorithms (friend recommendation)

- Alg 1 & OPT: our greedy recommendation with optimal influence spread computation (Monte-Carlo)
- Alg 1 & Alg 2: our greedy recommendation with our multi-path influence spread computation
- MaxSim: greedy algorithm that iteratively recommends a user with a maximum common neighbor similarity
- MaxDeg: greedy algorithm that iteratively recommends a user with a maximum outgoing degree
- Random: a baseline algorithm that recommends friends uniform-randomly

4. Experiments

Results (friend recommendation)

- Tune k (number of recommended friend)



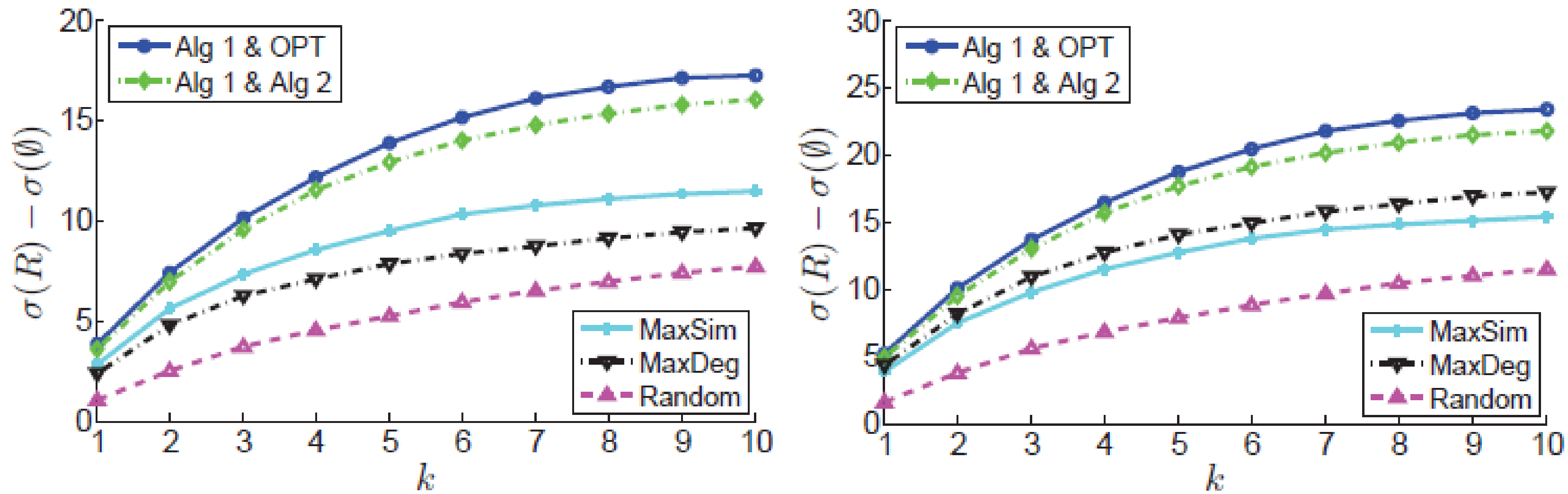
(a) Facebook. (b) Epinions. (c) Wiki.

- All algorithms suffers from diminishing return
- Alg 1 outperforms others, Alg 2 is close to OPT

4. Experiments

Results (friend recommendation)

- Popular user: outgoing degree > 100 (otherwise normal users)



(a) Facebook (normal users). (d) Facebook (popular users).

- Popular users have higher gain on the influence spread by friend recommendations

5. Conclusion

Friend recommendation

- Aim to maximize the influence spread of the user
- Friend acceptance probability v.s. influence capability
- Greedy recommendation has a ratio of 1-1/e

Influence spread computation

- Consider multiple influence propagation paths
- Closely approximate the optimal solution