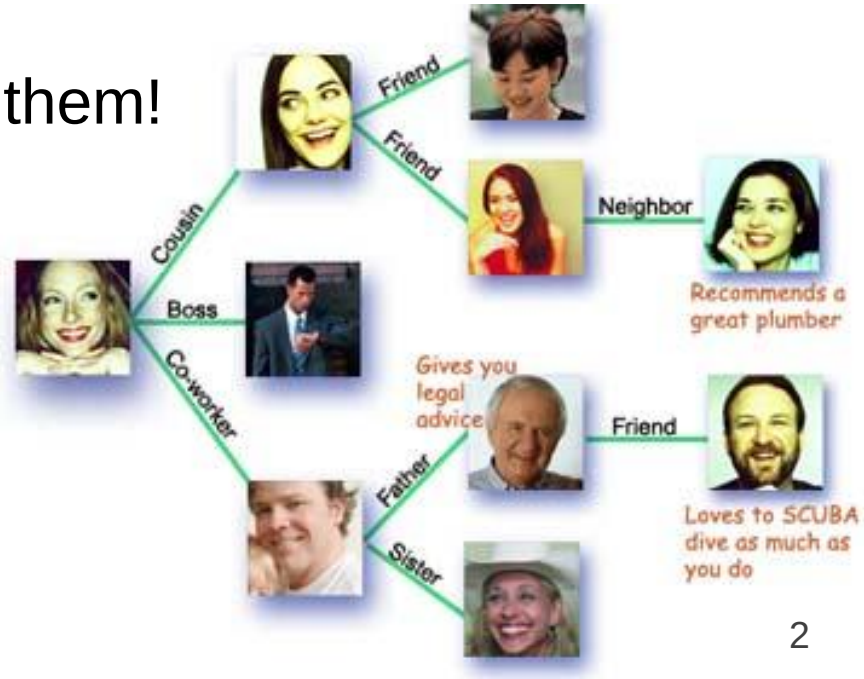
소셜 네트웍에서 영향력 최대화에 대한 연구

KAIST
Applied Algorithm Lab
Kyomin Jung

KAIST
Applied Algorithm Lab
Wooram Heo

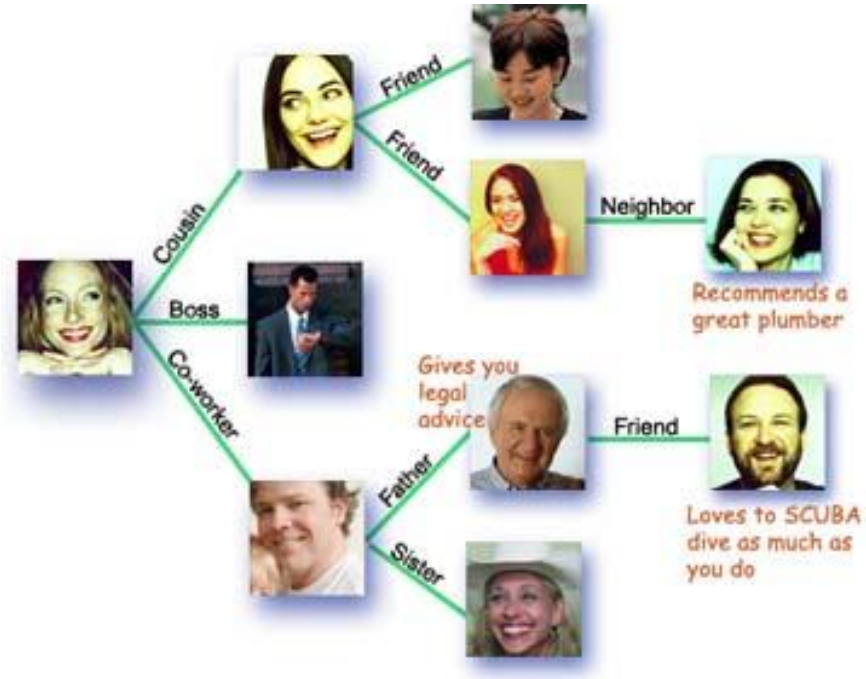
Introduction

- Social network plays a fundamental role as a medium for the spread of INFLUENCE among its members
  - Opinions, ideas, information, innovation…
- Efforts to estimate influence of nodes
- Find important nodes and target them!

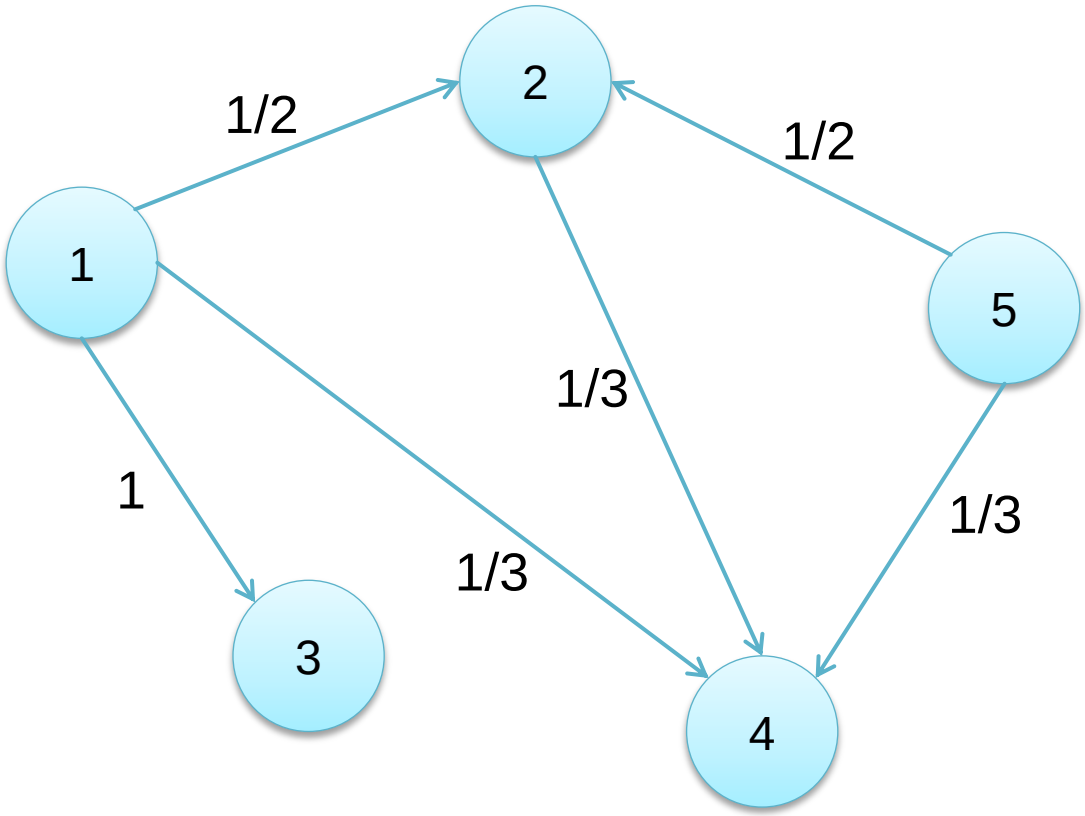


Introduction

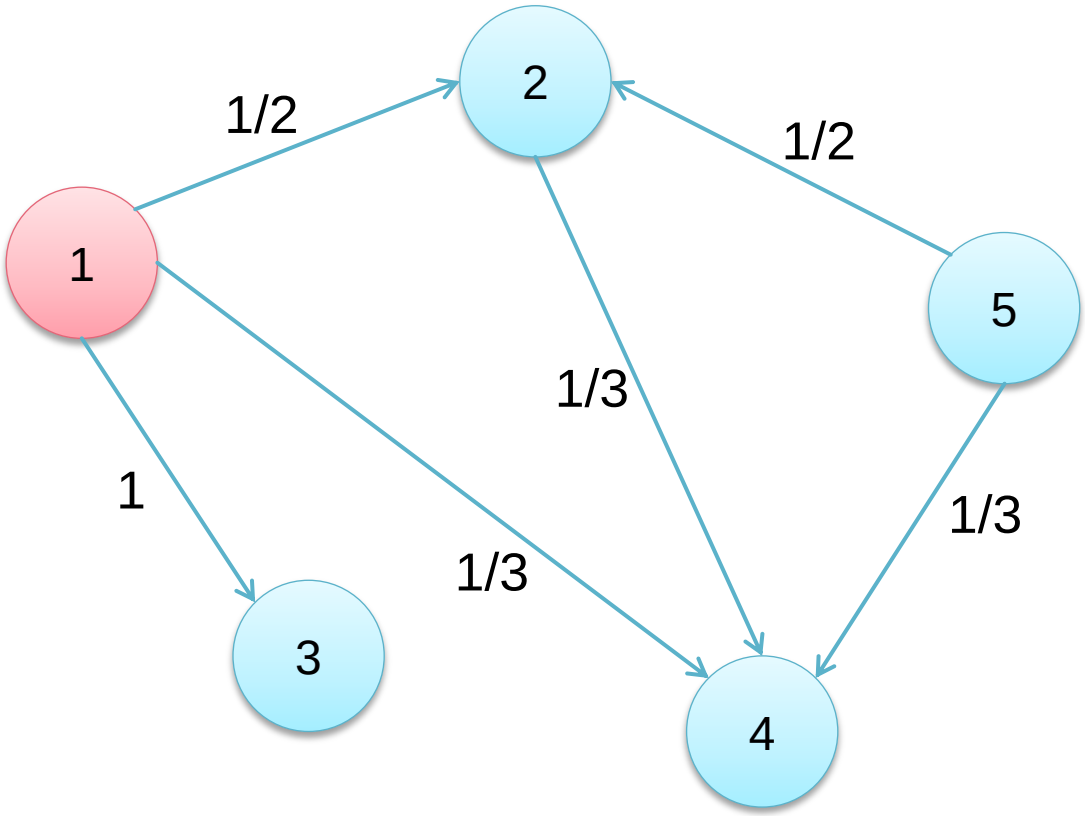
- Can be applied in
  - Marketing
  - Advertisement
  - Information/opinion spreading



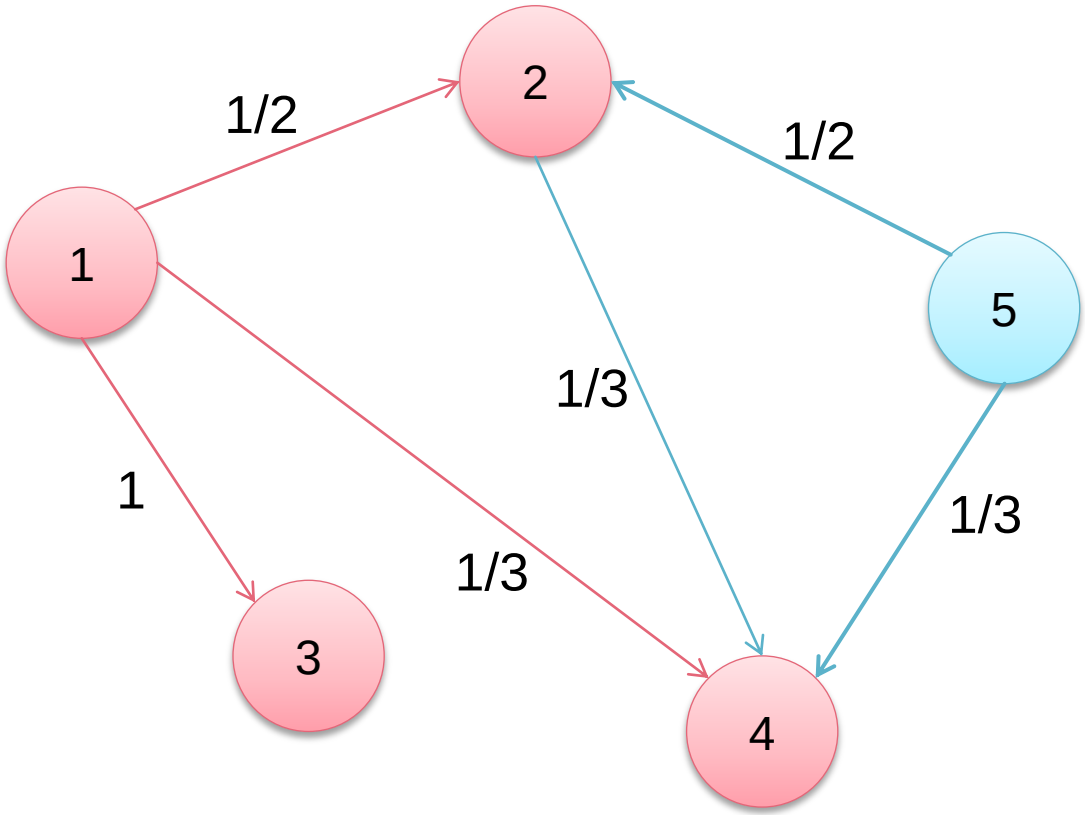
Diffusion process



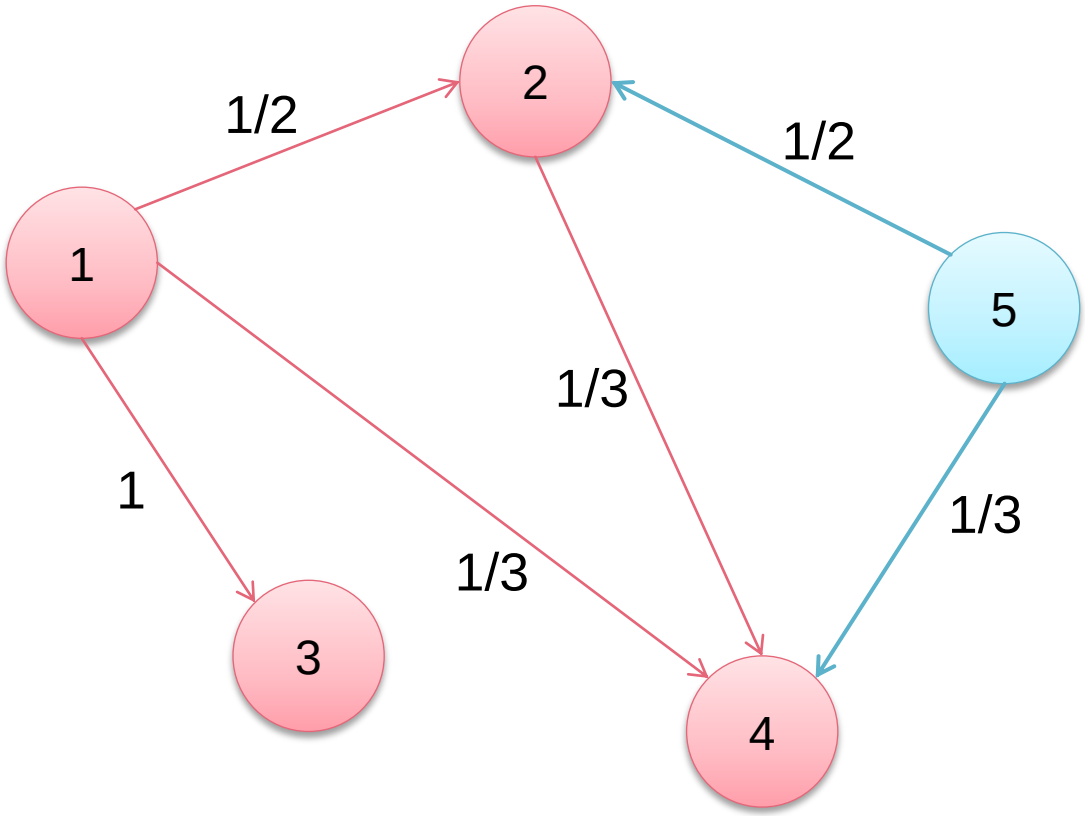
Diffusion process



Diffusion process



Diffusion process



Influence Maximization Problem

- Influence of node set S: f(S)
  - expected number of active nodes at the end, if set S is the initial active set
- Problem:
  - Given a parameter *k* (budget), find a *k*-node set *S* to maximize f(S)
  - Constrained optimization problem with f(S) as the objective function

Related Work

- D. Kempe, J. M. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In KDD ’ 03
  - They show the optimization problem is NP-hard
  - Greedy algorithm using monotone submodular function
  - (1 – 1/e) – approximation
  - Not scalable

- W. Chen, C. Wang, and Y. Wang. Scalable influence maximization for prevalent viral marketing in large-scale social networks, 2010
  - Heuristic algorithm
  - Not as good as greedy algorithm

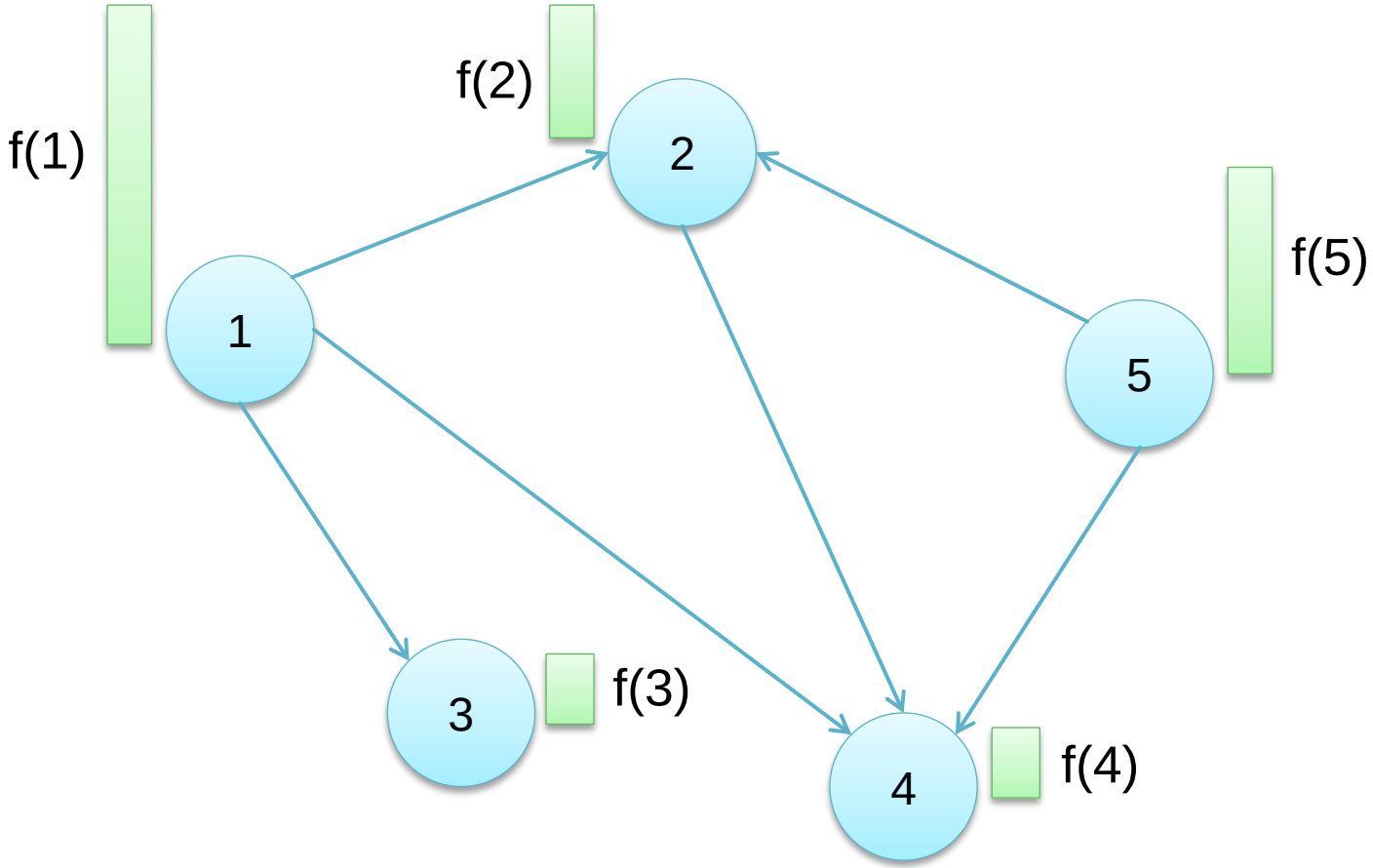
Basic greedy algorithm

Algorithm Greedy(k, f)

1: initialize $S = \emptyset$
2: **for** $i = 1$ to k **do**
3: select $u = \arg\max_{w \in V \setminus S}(f(S \cup \{w\}) - f(S))$
4: $S = S \cup \{u\}$
5: **end for**
6: output S

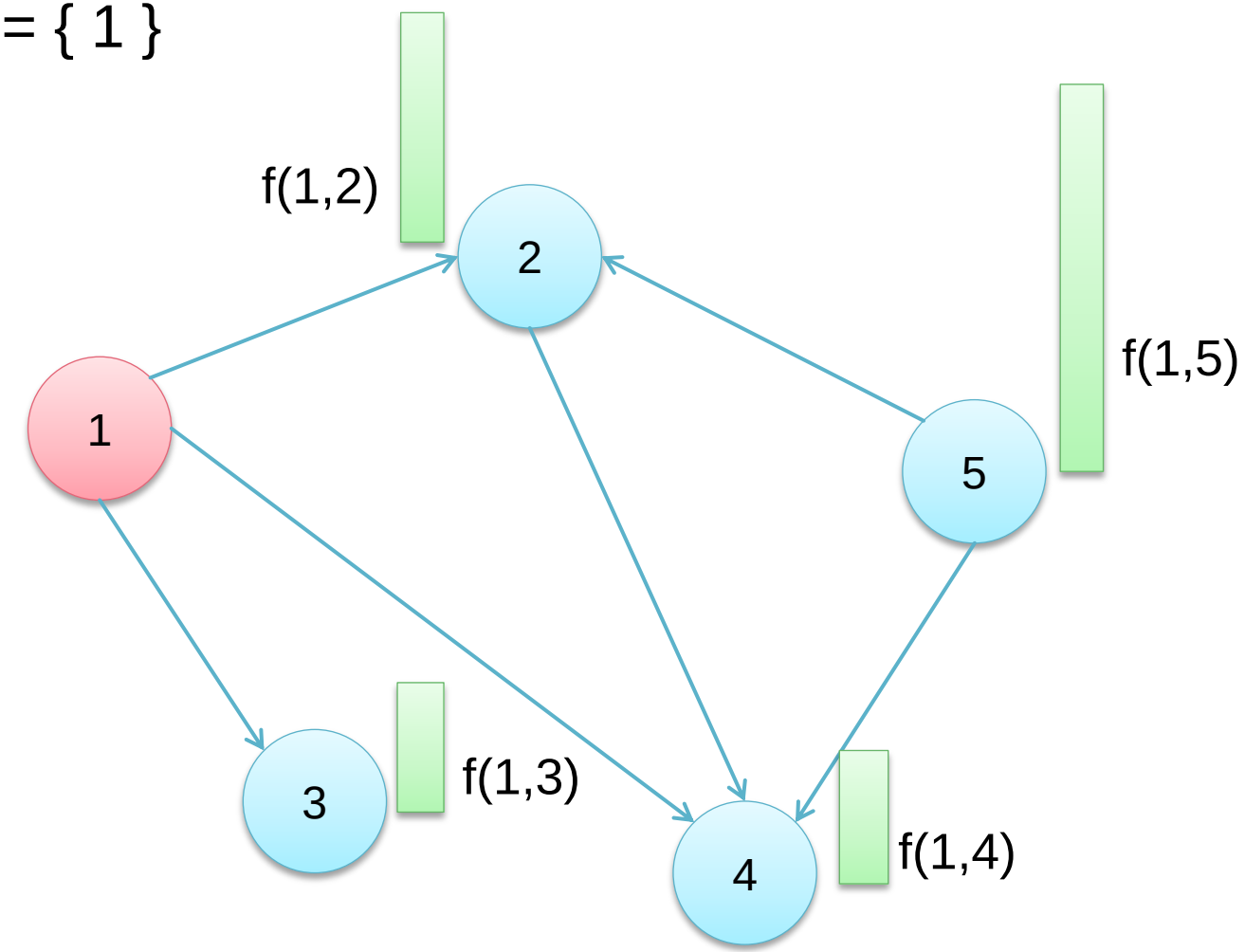
Basic greedy algorithm

- K = 3, S = { }



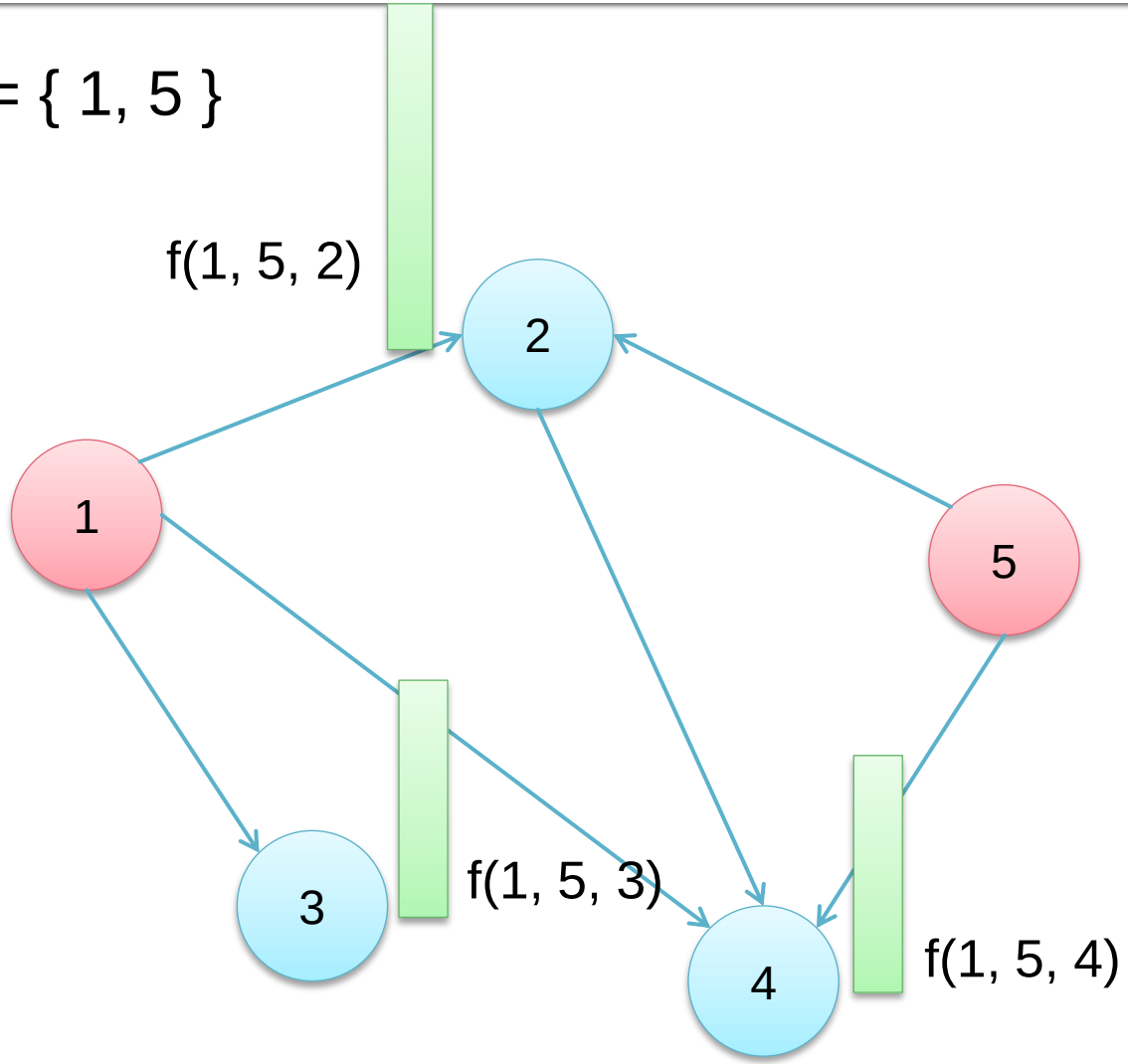
Basic greedy algorithm

- K = 3, S = { 1 }



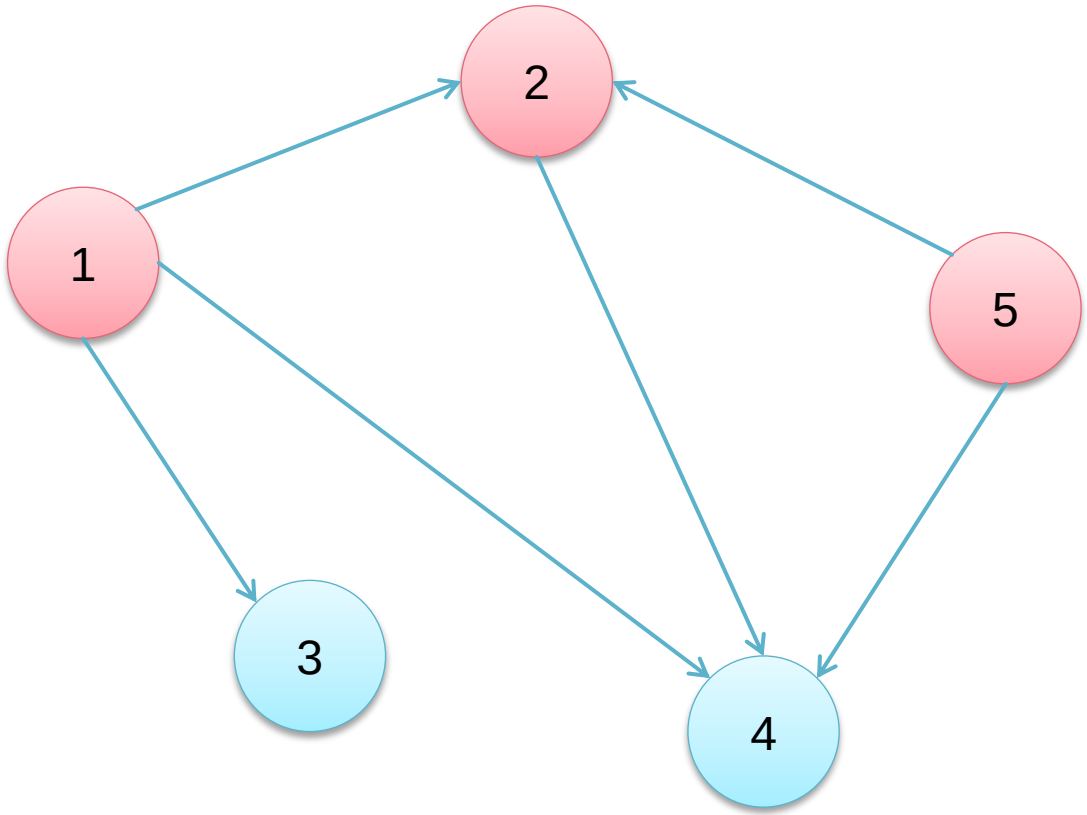
Basic greedy algorithm

- K = 3, S = { 1, 5 }



Basic greedy algorithm

- K = 3, S = { 1, 2, 5 }



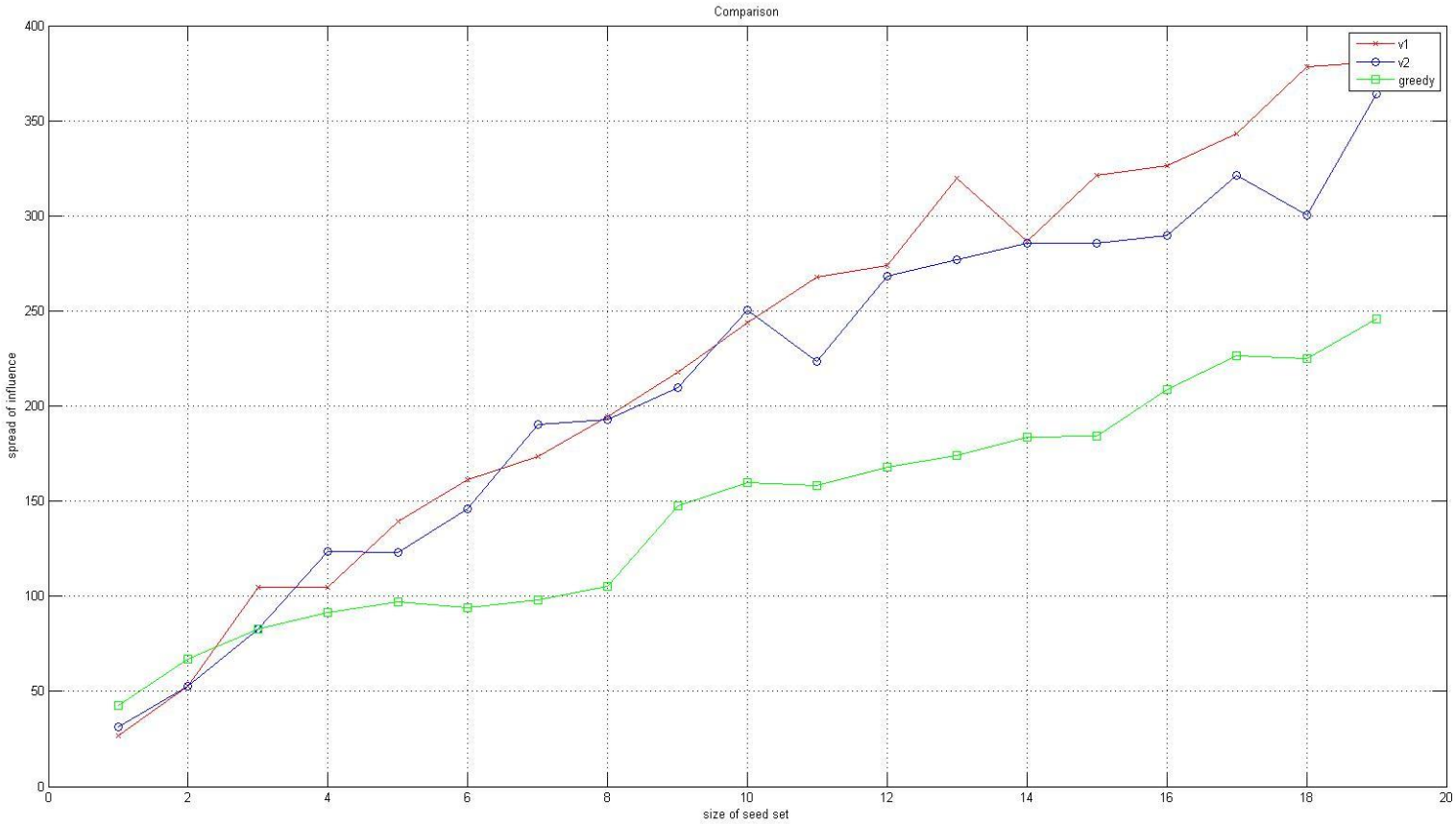
Our approach

- We propose a novel recursive method computing influence

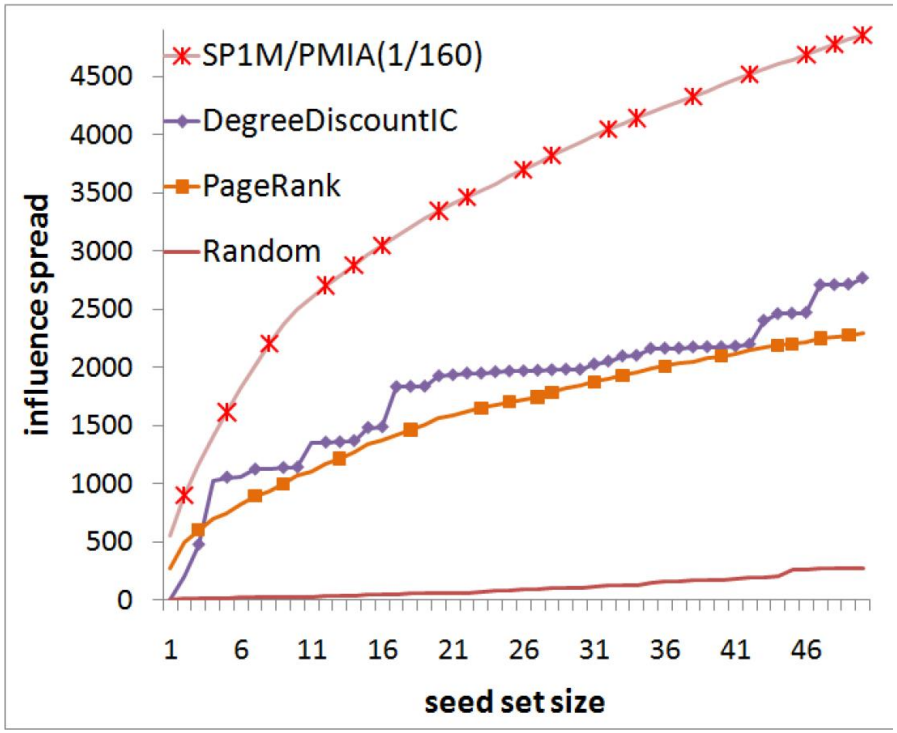
$$f(v) = \sum_{w \in N(v)} \alpha \cdot p_{vw} \cdot f(w)$$

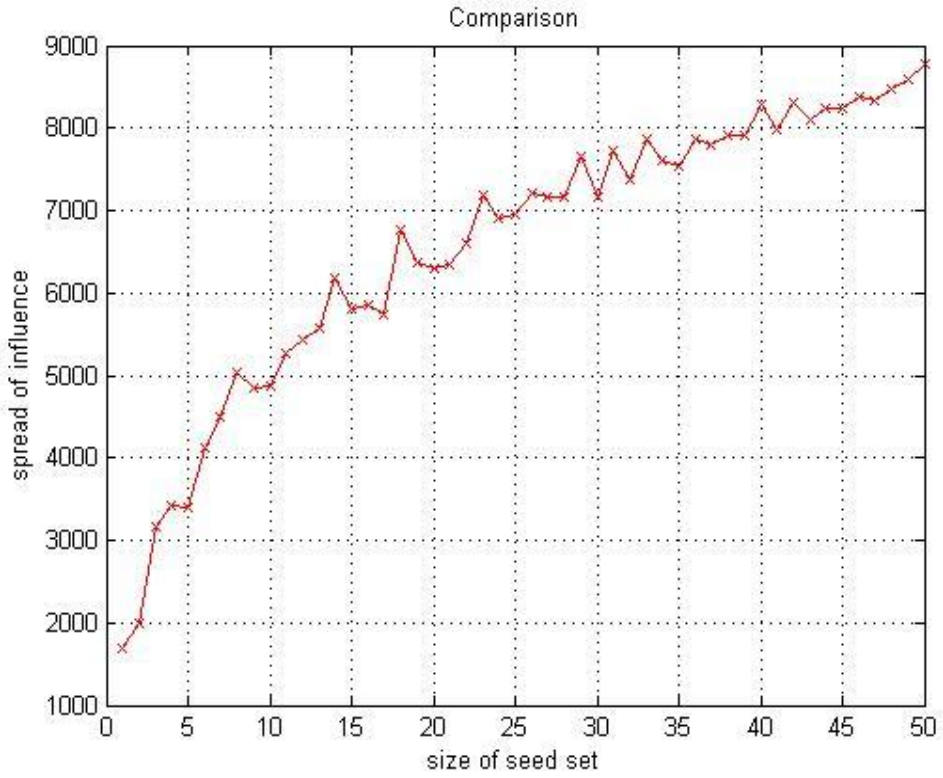
- It works empirically very well

Experiment


Comparison
v1
v2
greedy
spread of influence
size of seed set
0
50
100
150
200
250
300
350
400
0
2
4
6
8
10
12
14
16
18
20

Experiment


SP1M/PMIA(1/160)
DegreeDiscountIC
PageRank
Random
influence spread
seed set size
0
500
1000
1500
2000
2500
3000
3500
4000
4500
1
6
11
16
21
26
31
36
41
46


Comparison
spread of influence
size of seed set
1000
2000
3000
4000
5000
6000
7000
8000
9000
0
5
10
15
20
25
30
35
40
45
50

Reference

- [1] D. Kempe, J. M. Kleinberg, and E. Tardos. Maximizing the spread of influence through a social network. In KDD ’10, 2003
- [2] W. Chen, C. Wang, and Y. Wang. Scalable influence maximization for prevalent viral marketing in large-scale social networks. In KDD ’10, 2010
- [3] P. Domingos, M. Richardson. Mining the Network Value of Customers. Seventh International Conference on Knowledge Discovery and Data Mining, 2001
- [4] M. Richardson, P. Domingos. Mining Knowledge-Sharing Sites for Viral Marketing. *Eighth Intl. Conf. on Knowledge Discovery and Data Mining*, 2002

Q & A