Hien T. Nguyen
Vaclav Snasel (Eds.)

LNCS 9795

Computational Social Networks

5th International Conference, CSoNet 2016
Ho Chi Minh City, Vietnam, August 2–4, 2016
Proceedings

Springer

Contents

Shortest Paths on Evolving Graphs 1
Yiming Zou, Gang Zeng, Yuyi Wang, Xingwu Liu, Xiaoming Sun, Jialin Zhang, and Qiang Li

Analysis of a Reciprocal Network Using Google+: Structural Properties and Evolution 14
Braulio Dumba, Golshan Golnari, and Zhi-Li Zhang

Comparison of Random Walk Based Techniques for Estimating Network Averages 27
Konstantin Avrachenkov, Vivek S. Borkar, Arun Kadavankandy, and Jithin K. Sreedharan

Integrating Networks of Equipotent Nodes 39
Anastasia Moskvina and Jiamou Liu

Identify Influential Spreaders in Online Social Networks Based on Social Meta Path and PageRank 51
Vang V. Le, Hien T. Nguyen, Vaclav Snasel, and Tran Trong Dao

Immunization Strategies Based on the Overlapping Nodes in Networks with Community Structure 62
Debayan Chakraborty, Anurag Singh, and Hocine Cherifi

Improving Node Similarity for Discovering Community Structure in Complex Networks 74
Phuong N.H. Pham, Hien T. Nguyen, and Vaclav Snasel

Rumor Propagation Detection System in Social Network Services. 86
Hoonji Yang, Jiaofei Zhong, Dongsoo Ha, and Heekuck Oh

Detecting Overlapping Community in Social Networks Based on Fuzzy Membership Degree 99
Jiajia Rao, Hongwei Du, Xiaoting Yan, and Chuang Liu

Time-Critical Viral Marketing Strategy with the Competition on Online Social Networks 111
Canh V. Pham, My T. Thai, Dung Ha, Dung Q. Ngo, and Huan X. Hoang

Analysis of Viral Advertisement Re-Posting Activity in Social Media 123
Alexander Semenov, Alexander Nikolaev, Alexander Veremyev, Vladimir Boginski, and Eduardo L. Pasiliao

Time-Critical Viral Marketing Strategy with the Competition on Online Social Networks

Canh V. Pham[1,3](✉), My T. Thai[2], Dung Ha[3], Dung Q. Ngo[3], and Huan X. Hoang[1]

[1] University of Technology and Engineering, Vietnam National University, Hanoi, Vietnam
{14025118,huanhx}@vnu.edu.vn

[2] Department of Computer and Information Science and Engineering, University of Florida, Gainesville, USA
mythai@cise.ufl.edu

[3] Faculty of Technology and Information Security, People's Security Academy, Hanoi, Vietnam
cvpham.vnu@gmail.com, dungha.hvan@gmail.com, quocdung.ngo@gmail.com

Abstract. According to the development of the Internet, using social networks has become an efficient way to marketing these days. The problem of Influence Maximization (IM) appeared in marketing diffusion is one of hot subjects. Nevertheless, there are no researches on propagating information whereas limits unwanted users. Moreover, recent researches shows that information spreading seems to dim after some steps. Hence, how to maximize the influence while limits opposite users after a number of steps? The problem has real applications because business companies always mutually compete and extremely potential desire to broad cart their product without the leakage to opponents.

To be motivated by the phenomenon, we proposed a problem called Influence Maximization while unwanted users limited (d-IML) during known propagation hops d. The problem would be proved to be NP-Complete and could not be approximated with the rate $1 - 1/e$ and its objective function was sub modular. Furthermore, we recommended an efficient algorithms to solve the problem. The experiments were handled via the real social networks datasets and the results showed that our algorithm generated better outcome than several other heuristic methods.

Keywords: Influence maximization · Viral marketing · Leakage information

1 Introduction

With the fast development and steady of the Online Social Networks (OSNs), such as Facebook, Twitter, Google+, etc., OSNs have become the most common vehicle for information propagation. OSNs provided a nice platform for information diffusion and fast information exchange among their users.

H.T. Nguyen and V. Snasel (Eds.): CSoNet 2016, LNCS 9795, pp. 111–122, 2016.
DOI: 10.1007/978-3-319-42345-6_10

The topic of Influence Maximization (IM) has received a lot of research interests in recent years. This problem is firstly proposed by Kempe et al. [5] in two diffusion models which are *Independent Cascade* (IC) model and *Linear Threshold* (LT) model, then rapidly becoming a hot topic in social network field. They also proved that influence maximization problem is NP-hard, and a natural greedy algorithm can obtain $1 - 1/e$. Although extensive related works have been conducted on the IM problem [1–3,12,13], most of them are based on such an assumption that without existence of the unwanted target users whom we do not want information come to. In reality, on OSNs exits the group of users who have opposite viewpoint and benefits with us and they create a negative impact to oppose for information received.

Considering the following example that highlights a basic need for every organization that uses OSNs. There are two mutual competitive companies A and B. The A has been deploying a large advertisement, even via the Internet. They drew a marketing blueprint on several social networks but the A tried to hide everything against every one of the B as long as its possible. Constantly, the advertising information of A can reach to the B after a time. Thus, the A needs a solution help them fast imply the marketing strategy to much many users except unwanted users (from B) to gain the best consumption more quickly than B within t hop.

Motivated by the above phenomenon, this paper proposes a new optimization problem called Maximizing Influence while unwanted target users limited (IML) that finds a set of seeding S to Maximize Influence such that influence to unwanted ones is under a certain threshold after the largest d hop propagations. The total influence is the total activated people. The unwanted ones are those whom we do not want the information come to.

Our contributions in this paper are summarized as follows:

- We first attempt to study the maximizing influence while unwanted target users limited after d hops (d-IML) under LT model and show that the objective function was *submodular*.
- We proved d-IML was NP-Complete and show it can not be approximated in polynomial time with a ratio $1 - 1/e$ unless $NP \subseteq DTIME(n^{O(\log n \log n)})$. We also designed an efficient algorithm for the problem of d-MIL.
- We conducted our experiments on real-world datasets, the results indicated that our algorithm gives better results than several other heuristic methods.

Related Work. The target is to spread the desired information for as many people as possible on OSNs. There are different related works on this topic [1–3,12,13]. Zhang et al. [1] proposed a problem to maximize the positive news in propaganda rather than maximizing the users affected. They said that to maximize positive things in many cases had more beneficial than maximizing the number of people affected. They used the Cascade Opinion (OC) model to solve the problem. On the other hands, Guo et al. [2] recommended to maximize the influence of information to a specific user by finding out the k most influential users and proved that it was NP-hard problem and the function is *submodular*. They also launched an effective approximation algorithm.

Zhuang et al. [3] have studied the *IM* problem in the dynamic social network model over time. In addition, there were several other studies: Chen et al. [12] investigated *IM* problem on a limited time; Gomez-Rodriguez et al. [13] studied *IM* problem for continuous time. Researches on IM with various contexts and various models received many attentions, but the diffusion of information problem, in addition to spreading the positive information still faced with the misinformation. How to spread the positive information while the misinformation limited? To solve it, Budak et al. [4] launched the problem selecting k users to convince them aware of good information so that after the campaign, amount of use influenced by the misinformation was the least. By using Model-Oblivious Independent Campaign Cascade, they proved the problem be NP-hard and the objective function was *submodular*. Nguyen et al. [6] gave the decontamination problem of misinformation by selecting a set of users with sources of misinformation I assumed to have existed on the social network at the rate of $\beta \in [0, 1]$ after T time. They launched the different circumstances of the I and the T, but they only solved the case I was unknown. On preventing infiltration to steal information on OSNs, Pham et al. [11] have built a Safe Community for the purpose of protecting all users in an Organization. In problems of detecting misinformation source on social networks, Nguyen et al. [8] assumed that the exist a set of misinformation sources I, they purposed of finding the largest number of users in I who started to propagate that information. Nevertheless, the predictions were likely confused because they did not know the order of real time start to spread misinformation. Zhang et al. [9] studied the problem of limited resources that often was incorrect information while maximized the positive source of misinformation on OSNs under Competitive Activation model. In this study, they were considered a model of misinformation and good information and presence on the social network, they also proved to be NP-complete problem and could not be approximated with rate $(1 - \frac{1}{e})$ unless $NP \subseteq DTIME(n^{O(\log \log n)})$.

In these researches, no one focused on the spread of information with the limiting of information to the set of ones who we did not want the information reach to (called *unwanted users*). While positive information is desired to propagate to more and more users, we also face with the existence of unlike users on OSNs. Because every time they receive the positive information, they can be able to conduct the activities, propagation strategies that opposes to our benefits.

2 Model and Problem Definition

2.1 Network and Information Diffusion Model

We are given a social network modeled as an undirected graph $G = (V, E, w)$ where the vertices in V represent users in the network, the edges in E represent social links between users and the weight represent frequency of interaction between two users. We use n and m to denote the number of vertices and edges. The set of neighbors of a vertex $v \in V$ is denote by $N(v)$. Existing diffusion models can be categorized into two main groups [5]: Threshold model and Independent Cascade model. In this work, we use the Linear Threshold (LT) model which is the one that has been extensively used in studying diffusion models among the generalizations of threshold models.

Linear Threshold (LT) Model. In this model, each node v has a threshold θ_v and for every $u \in N(v)$, edge (u, v) has a nonnegative weight $w(u, v)$ such that $\sum_{v \in N(u)} w(u, v) \leq 1$. Given the thresholds and an initial set of active nodes, the process unfolds deterministically in discrete steps. At hop t, an inactive node v becomes active if

$$\sum_{u \in N^a(v)} w(u, v) \geq \theta_v$$

where $N^a(v)$ denotes the set of active neighbors of v. Every activated node remains active, and the process terminates if no more activations are possible. Kempe et al. [5] prove that influence function $\delta(.)$ is *submodular* function.

2.2 Problem Definition

The paper we are interested in the value of the influence function after d hops. Considering that influence can be propagate at most d hops, we define the influence function $\delta_d(S)$ as total number of nodes have been active by S after d hops. We study maximizing influence while unwanted target users limited after d hops (d-IML) under LT model as follow:

Definition 1 (d-IML problem). *Given an social network represented by a directed graph $G = (V, E, w)$ and an under LT model. Let $T = \{t_1, t_2, .., t_p\}$ be the set of $|T| = p$ unwanted users and d is number of hops limited. Our goal chose the set seed user $S \subseteq V$ at most k-size that maximize $\delta_d(S)$ such that total influence user come to t_i less than threshold for prevent information leakage τ_i i.e.: $\sum_{u_i \in N^a(t_i)} t_i < \tau_i$.*

Lemma 1. *The influence function $\delta_d(.)$ is monotone and submodular for an arbitrary instance of the LT model, given any time $d \geq 1$.*

Proof. LT model is a special case of LT-M model [12] with all parameters $m(u, v) = 1$ and deadline $\tau = d$ and influence function $\delta_d(.)$ is influence function $\delta_\tau(.)$. Due to $\delta_\tau(.)$ is monotone and submodular in LT-M model, thus $\delta_d(.)$ is monotone and submodular in LT model. □

3 Complexity

In this section, we first show the NP-Completeness of IML problem on LT model by reducing it from Maximum Coverage problem. By this result, we further prove the inapproximability of d-MIL which is NP-hard to be approximated within a ratio of $1 - 1/e$ unless $NP \subseteq DTIME(n^{O(\log \log n)})$.

Theorem 1. *d-IML is NP-Complete in LT model.*

Proof. We consider of the decision version of d-MIL problem that asks whether the graph $G = (V, E, w)$ contains a set $k - size$ of seed user $S \subset V$ that number active node at least K, such that $\sum_{u \in N^a(t_i)} w(u, t_i) < \tau_i$ within at most d rounds.

Given $S \subset V$, we can calculate the influence spread of S in polynomial time under LT model. This implies d-MIL is NP. Now we prove a restricted class of d-MIL instance is NP-hard, $d = 1$.

To prove that 1-MIL is NP-hard, we reduce it from the decision version of Maximum Coverage problem defined as follows.

Maximum Coverage. Given a positive integer k, a set of m element $\mathcal{U} = \{e_1, e_2, \ldots, e_m\}$ and a collection of set $\mathcal{S} = \{S_1, S_2, \ldots, S_n\}$. The sets may have common elements. The *Maximum Coverage* problem asks to find a subset $S' \subset S$, such that $| \cup_{S_i \in S'} S_i|$ is maximized with $|S'| \leq k$. The decision of this problem asks whatever the input instance contains a subset S' of size k which can cover at least t elements where t is a positive integer.

Reduction. Given an instance $I = \{\mathcal{U}, \mathcal{S}, k, t\}$ of the maximum coverage, we construct an instance $G = (V, E, w, \theta)$ of 1-IML problem as follows:

- *The set of vertices:* add one vertex u_i for each subset $S_i \in \mathcal{S}$, once vertex v_j for each $e_j \in \mathcal{U}$, and a vertex x is a unwanted users.
- *The set of edges:* add an edge (v_i, u_j) for each $e_j \in S_i$ and connect x to each vertex v_j.
- *Thresholds and weights:* assign all vertices the same threshold $\theta = \dfrac{1}{m}$ and each edges (u_i, v_j) has weight $w_{u_i v_j} = \dfrac{1}{m}$. In addition, for all edges (v_j, x), we assign their weight $\dfrac{1}{m}$.
- *Threshold for prevent leakage information:* we assign threshold for prevent leakage information for vertex x is $\tau_x = \dfrac{1}{m}$.

The reduction is illustrated in Fig. 1. Finally, set $d = 1, K = t$.

Suppose that S^* is a solution to the maximum coverage instance, thus $|S^*| \leq k$ and it can cover at least t elements in $\mathcal{U}$. By our construction, we select all nodes u_i corresponding to subset $S_i \in S^*$ as seeding set S. Thus, $|S| = k$. Since S^* cover at least t elements e_j in $\mathcal{U}$ so S influence at least t vertices v_j corresponding to those e_j. Additionally, for each v_j total influence incoming based on LT model at least $w_{u_i v_j} = \theta_{u_i v_j} = w_{u_i v_j}$. Hence, there are at least $t = K$ nodes in the 1-IML has been active.

Conversely, suppose there is seeding $S, |S| = k$ such that the number of active node at least K. We see that $v_j \notin S, j = 1, 2, .., m$ because total influence incoming x at least $w_{v_j x} = \tau_x = \dfrac{1}{m}$. Thus $S \subseteq \{u_1, u_2, \ldots, u_n\}$. Then S^* can be collection of subset S_i corresponding to those $u_i \in S$. Hence the number of elements which it can cover is at least $K = t$. □

Based on above reduction, we further show that inapproximation of IML in the following theorem.

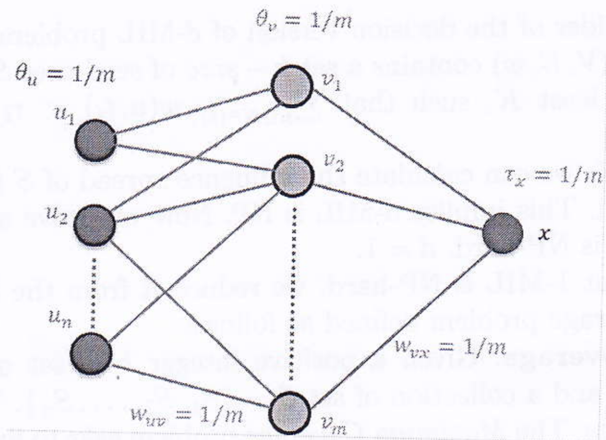


Fig. 1. Reduction from MC to 1-IML

Theorem 2. *The d-IML problem can not be approximated in polynomial time within a ratio of* $\frac{e}{1-e}$ *unless* $NP \subseteq DTIME(n^{O(\log \log n)})$.

Proof. Supposed that there exits a $\frac{e}{1-e}$-approximation algorithm $\mathcal{A}$ for d-MIL problem. We use the above reduction in proof of Theorem 1 then $\mathcal{A}$ can return the number of active nodes in G with seeding size equal to k. By our constructed instance in Theorem 1, we obtain the Maximum Coverage with size t if the number of active nodes in optimal solution given by $\mathcal{A}$ is K. Thus algorithm $\mathcal{A}$ can be applied to solve the Maximum Coverage problem in polynomial time. This contradict to the NP-hardness of Maximum Coverage problem in [10].

4 Methodology

4.1 ILP Formulation

One advantage of our discrete diffusion model over probabilities is that the exact solution can found be using mathematical programming. Thus, we formulate d-IML problem as an $0-1$ Integer Linear Programming (ILP) problem below.

The objective function (1) of the ILP is to find the number of node is active. The constraint (2) is number of set seed is bounded by k; the constraints (3) capture the propagation model; the constraint (4) limit leakage information income to unwanted user by threshold τ_i; and the constraint is simply to keep vertices active once they are activated. The number of variables and constraints of ILP are nd.

$$\texttt{maximize} \sum_{v \in V \setminus T} x_v^d \tag{1}$$

$$\texttt{st:} \sum_{v \in V \setminus T} x_v^0 \leq k \tag{2}$$

$$\sum_{v \in N(u)} x_u^{i-1}.w(u,v) + \theta_v.x_v^{i-1} \geq \theta_v.x_v^i,$$
$$\forall v \in V, i = 1..d \tag{3}$$
$$\sum_{v \in N(t_i)} x_v^d.w(v,t_i) < \tau_i \tag{4}$$
$$x_v^i \geq x_w^{i-1}, \forall v \in V, i = 1..d \tag{5}$$

$$\text{where } x_v^i = \begin{cases} 1 & \text{if } x \text{ is active at round (hop) } i \\ 0 & \text{otherwise} \end{cases}$$

4.2 Meta Heuristic Algorithm

A commonly heuristic used for *IM* problem in the simple case is greedy algorithm, the idea of which choses the node is maximize *influence marginal gain* in each step:

$$\delta_d(S,v) = \delta_d(S + \{v\}) - \delta_d(S) \tag{6}$$

Although the objective function is *submodular*, propagation of influence is constraint by the leak. Hence we can not give an algorithm for approximately with the ratio $1 - 1/e$ as Kemp et al. [5]. To avoid this issue, we designed the algorithm combine the influence marginal gain and evaluation information leakage based on idea of IGC [7]. Accordingly, we used a heuristic function $f(v)$ to evaluate the fitness of user v which defined as follows:

$$f(v) = \frac{\delta_d(S,v)}{1 + \frac{1}{p}\sum_{t_i \in T} l_{t_i}(v)} \tag{7}$$

Where $L_{t_i}(S)$ is the total influence to t_i respect to seeding sets S after d hop *i.e.*, $L_{t_i}(S) = \sum_{v \in N^a(t_i)} w(v,t_i)$, $l_{t_i}(v) = \frac{L_{t_i}(S+\{v\})}{\tau_i}$ is the normalized leakage level at t_i after adding v to seed set S. The numerator of $f(v)$ is selected to be influence marginal gain $\delta_d(S,v)$ so that the algorithm will favor users have maximizing influence, denominator of $f(v)$ will favor users with lower information leakage.

The Meta-heuristic (MH) algorithm as shown in Algorithm 1. In each iteration, firstly, we update the set of candidate users C, those whose addition to seeding set S still guarantees that the information leakage to each unwanted t_i does not exceed the threshold τ_i. The algorithms also adds one user v of candidate set C into S which has $f(v)$ is maximize until size of S no exceed k.

5 Experiment

In this section, we do a lot of experiment on three real-world datasets, and compare our algorithm with algorithms: Random method, Max degree, Greedy and ILP method.

Algorithm 1. Meta-heuristic Algorithm

Data: $G = (V, E, w, \theta, \tau)$, set users $U = \{t_1, t_2, \ldots, t_p\}$, p, k;
Result: Seeding S;
$S \leftarrow \emptyset$;
$C \leftarrow V$;
for $i = 1$ *to* k **do**
  foreach $v \in C$ **do**
    if $\exists j : L_{t_j}(S + \{v\}) \geq \tau_j$ **then**
      $C \leftarrow C - \{v\}$;
    end
  end
  if $C = \emptyset$ **then**
    Return S;
  end
  Find $v \in C$ that maximizes $f(v)$;
  $S \leftarrow S + \{v\}$;
  $C \leftarrow C - \{v\}$;
end
Return S;

5.1 Datasets

BlogCatalog. BlogCatalog is a social blog directory website. This contains the friendship network crawled and group memberships. Nodes represent users and edges are the friendship network among the users. Since the network is symmetric, each edge is represented only once [14].
ArXiv-Collaboration. The data covers papers in the period from January 1993 to April 2003 (124 months). It begins within a few months of the inception of the arXiv, and thus represents essentially the complete history of its GR-QC section. If an author i co-authored a paper with author j, the graph contains a undirected edge from i to j. If the paper is co-authored by k authors this generates a completely connected (sub)graph on k nodes [15].
Gnutella. A sequence of snapshots of the Gnutella peer-to-peer file sharing network from August 2002. There are total of 9 snapshots of Gnutella network collected in August 2002. Nodes represent hosts in the Gnutella network topology and edges represent connections between the Gnutella hosts [15].

Table 1. Basic information of Network Datasets

Network	Nodes	Edges	Type	Avg. Degree
BlogCatalog	10,312	333,983	Undirect	32.39
ArXiv-Collaboration	5,242	28,980	Direct	5.53
Gnutella	6,301	20,777	Direct	3.30

In each graph, we used the method in [5] to assign the diffusion weight to each edge and then normalize the weights of all incoming edges of a node v to let it satisfy that $\sum_{u \in N^{in}(v)} w(u,v) \leq 1$ (Table 1).

5.2 Comparision Algorithm

In this part, we describe tow comparison algorithms: random and Max degree.

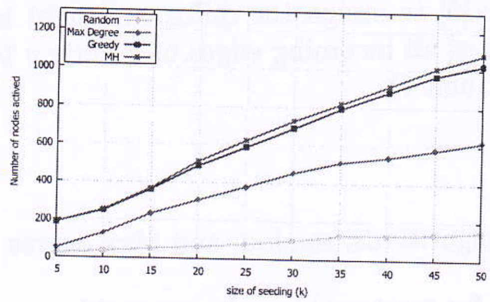
1. *Random:* This is a general method used for most problem. In our problem, we chose the seeding node randomly when the information leaked to unwanted users less than the threshold leakage.
2. *Max Degree method:* The greedy algorithm chose the vertex v that had maximum degree when the information leaked to unwanted users less than the threshold leakage.
3. *Greedy algorithm (GA) method:* The method based on the idea that chose the node maximize information diffusion gain when information leaked to unwanted users less than the threshold leakage.
4. *Meta heuristic (MH) algorithm:* Here the algorithm in Sect. 4 collectively called our algorithms.
5. *ILP method:* Solve the ILP problem to compare with optimal seeding.

We solved the ILP problem on Gnutella network [15], with $d = 4$, The ILP was solve with CPLEX version 12.6 on Intel Xeon 3.6 Ghz, 16G memories and setting time limit for the solver to be 48 h. For $k = 5, 10, 15$ and 20 the solver return the optimal solution. However, for $k = 25, 30, 35, 40, 45$ and 50, the solver can not find the optimal solution within time limit and return sub-optimal solutions.

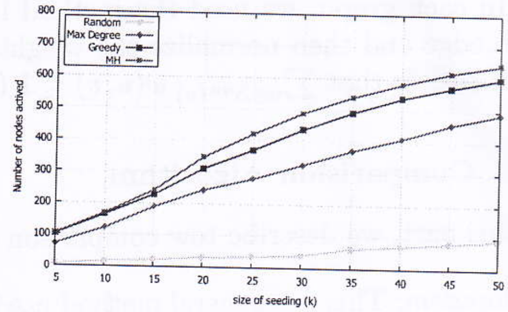
5.3 Experiment Results

Solution Quality. The number of active users changed when the number of steps d changed and fixed $k = 50$ shown in Fig. 3. The algorithm really resulted better than Max Degree and GA. The bigger k was, the better the result was than Max Degree. For example with the social network BlogCatalog, the MH generated more active users with 1.71 times ($k = 50$, $d = 4$). According to GA, when k was small, MH and GA issued the same outcome. When k was larger, the gap between MH and GA became clearer. MH produced better than GA 7.3 % with $k = 50$, using Guntella network.

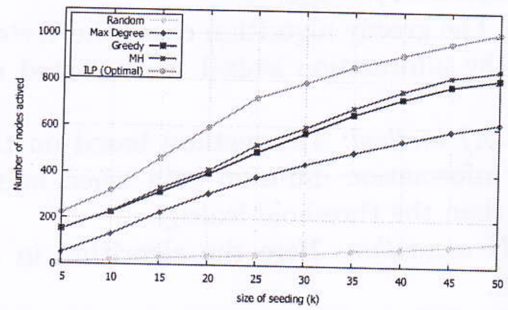
Number of Activated Users. We compared the performance of MH with the others when k changed and $d = 4$. The number of active users was detailed via Figs. 1, 2 and 3. Definitely, MH generated better than Max. Even MH worked better than Maxdegree 1.7 times when $k = 10$ via the network BlogCatalog. GA generated approximately to MH when k was small ($k = 5$, 10). When k was larger, MH worked better than GA. In case of maximum of k ($k = 50$) Gnutella activated users using MH more 56 people than GA whereas in BlogCatalog, activated users of MH and GA were 56 and 44, respectively.



(a) BlogCatalog

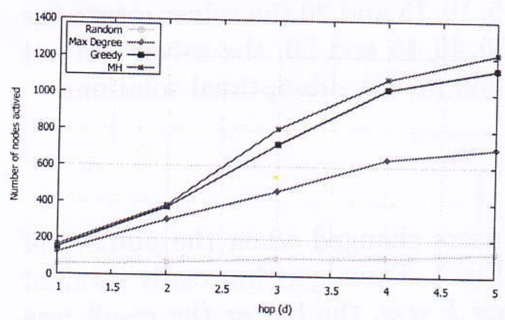


(b) Gnutella

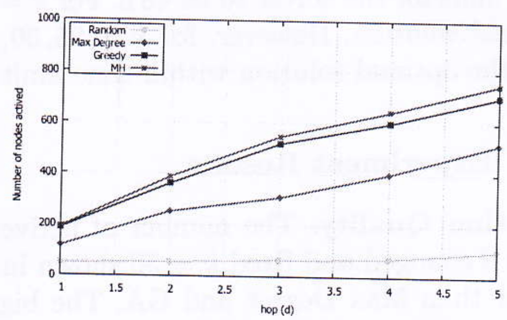


(c) arXiv-Collaboration

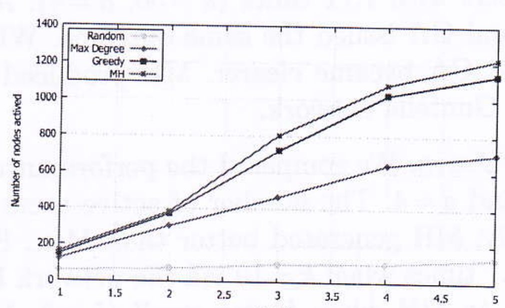
Fig. 2. The actived nodes when the size of seeding set varies ($d = 4$).



(a) BlogCatalog



(b) Gnutella



(c) arXiv-Collaboration

Fig. 3. The actived nodes when number of propagation varies ($k = 50$).

On the whole, the estimation of the function $f(.)$ resulted better than the one of the maximization of Influence gain. Nevertheless, when k was small these ways are the same. The comparison with ILP in Arxiv Collaboration, the proportion of the solution of MH at least marked at 68 % when $k = 20$ with activate users was less than 185. The rating was smallest when k = 50, pointed at 80 % of sub-optimal. Note that ILP did not generate optimized solution in this case.

Number of Hops. When d was small, MH and GA issued the same outcome. When d was large, MH resulted moderately better than MD and quite better than GA. It can be seen via BlogCatalog, when $d = 5$, the largest distance between GA and MH was 86 nodes. It proved that the larger d was, the better estimation the function $f()$ had to choose the optimization values.

6 Conclusions

In order to propose a viral marketing solution while exists the competition between organizations that have benefit collisions, we built the problem of maximization influence to users whereas limits the information reach to unwanted ones in constrained time. We proved it be an NP-complete and not be approximated with $1 - 1/e$ rating number. We also recommended an efficient solution MH to solve the problem. The experiment via social networks data showed that our algorithm got the better result of object function than some ones and got desired rate in optimized solution.

References

1. Zhang, H., Dinh, T.N., Thai, M.T.: Maximizing the spread of positive influence in online social networks. In: Proceedings of the IEEE International Conference on Distributed Computing Systems (ICDCS) (2013)
2. Guo, J., Zhang, P., Zhou, C., Cao, Y., Guo, L.: Personalized influence maximization on social networks. In: Proceedings of the 22nd ACM International Conference on Conference on Information and Knowledge Management (2011)
3. Zhuang, H., Sun, Y., Tang, J., Zhang, J., Sun, X.: Influence maximization in dynamic social networks. In: Proceedings of IEEE International Conference on Data Mining (ICDM) (2013)
4. Budak, C., Agrawal, D., El Abbadi, A.: Limiting the spread of misinformation in social networks. In: Proceedings of the 20th International Conference on World Wide Web (WWW 2011), pp. 665–674. ACM, New York, NY, USA (2011)
5. Kempe, D., Kleinberg, J., Tardos, E.: Maximizing the spread of influence through a social network. In: Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2003, New York, NY, USA, pp. 137–146 (2003)
6. Nguyen, N.P., Yan, G., Thai, M.T., Eidenbenz, S.: Containment of misinformation spread in online social networks. In: Proceedings of ACM Web Science (WebSci) (2012)
7. Dinh, T.N., Shen, Y., Thai, M.T.: The walls have ears: optimize sharing for visibility and privacy in online social networks. In: Proceedings of ACM International Conference on Information and Knowledge Management (CIKM) (2012)

8. Nguyen, D.T., Nguyen, N.P., Thai, M.T.: Sources of misinformation in online social networks: who to suspect? In: Proceedings of the IEEE Military Communications Conference (MILCOM) (2012)
9. Zhang, H., Li, X., Thai, M.: Limiting the spread of misinformation while effectively raising awareness in social networks. In: Proceedings of the 4th International Conference on Computational Social Networks (CSoNet) (2015)
10. Feige, U.: A threshold of ln n for approximating set cover. J. ACM (JACM) **45**(4), 634–652 (1998)
11. Pham, C.V., Hoang, H.X., Vu, M.M.: Preventing and detecting infiltration on online social networks. In: Thai, M.T., Nguyen, N.T., Shen, H. (eds.) CSoNet 2015. LNCS, vol. 9197, pp. 60–73. Springer, Heidelberg (2015)
12. Chen, W., Wei, L., Zhang, N.: Time-critical influence maximization in social networks with time-delayed diffusion process. http://arxiv.org/abs/1204.3074
13. Gomez-Rodriguez, M., Song, L., Nan, D., Zha, H., Scholkopf, B.: Influence estimation and maximization in continuous-time diffusion networks. ACM Trans. Inf. Syst. **34**, 2 (2016). doi:10.1145/2824253
14. Tang, L., Liu, H.: Relational learning via latent social dimensions. In: Proceedings of the 15th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2009), pp. 817–826 (2009)
15. Leskovec, J., Kleinberg, J., Faloutsos, C.: Graph evolution: densification and shrinking diameters. ACM Trans. Knowl. Disc. Data (ACM TKDD) **1**(1), 2 (2007)