Influence Maximization for Informed Agents in Collective Behavior*

Amir Asiaee Taheri, Mohammad Afshar, and Masoud Asadpour

Abstract. Control of collective behavior is an active topic in biology, social, and computer science. In this work we investigate how a minority of informed agents can influence and control the whole society through local interactions. The problem we specifically target is that a minority of people with a bounded budget for initiating new social relations attempt to control the collective behavior of a society and move the crowd toward a specific goal. Assuming that local interactions can only take place between friends, the minority has to initiate some new relations with the majority. The total cost of new relations is limited to a budget. The problem is then finding the optimal links in order to gain maximum impact on the society. We will model the problem as a diffusion process in a social network. The proof of NP-hardness of the problem for Local Interaction Game model of diffusion is presented. Simulations show that the proposed method surpasses the popular strategies based on degree and distance centrality in performance.

1 Introduction

Influencing society and changing the crowd behavior is one of the oldest ambitions of social science. Social and political sciences pursue strong impact on the society to change the attitude of people and prevail a desired behavior in the society. Socio-physics deals with such problems under the Opinion Formation topic [1]. In economical side this phenomena is known as Viral Marketing [2], [3].

The main problem that has been investigated extensively for attaining manipulation of crowd behavior is finding most influential persons of a society whom we

Amir Asiaee Taheri · Mohammad Afshar · Masoud Asadpour
Social Networks Lab, School of Electrical and Computer Engineering,
University of Tehran, Tehran, Iran
e-mail: {aasiaeet,mafshar,asadpour}@ut.ac.ir

Masoud Asadpour
School of Computer Science, Institute for Research in Fundamental Sciences (IPM),
P.O. Box 19395-5746, Tehran, Iran

* This research was in part supported by a grant from IPM (No. CS1388-4-14).

A. Martinoli et al. (Eds.): Distributed Autonomous Robotic Systems, STAR 83, pp. 389–402.
springerlink.com

call initiators from now on. Organizational theory calls such influential persons key players [4] and in political science they are called opinion leaders [5].

The idea is that influencing initiators would lead to the greatest possible diffusion of a behavior in a society. In Viral Marketing initiators are influential customers who are selected for direct marketing. Giving free samples or discounts are examples of marketing strategies for motivating initiators which would lead to stimulation of the others for buying the new product. Two-step flow theory in social and political sciences assumes initiators are well connected opinion leaders who channel media information to the masses [6].

All solutions of "K most influential persons" (K-MIP) problem suppose that changing initiators' opinions or behaviors is possible by costing a budget [7][8]. This simplification is not the case for real world problem, since opinion and behavior of people usually cannot be influenced by paying money.

On the other hand, studies in the field of Swarm Intelligence have more robust solutions to this problem. Couzin et al. [9] showed that among a group of foraging or migrating animals only a small fraction of them have proper information about the location of food source, or about the migration route. But these informed agents can guide the whole group through simple social interactions. The bigger the group is, the smaller the fraction of the required informed agents is. Halloy et al [10] showed in real experiments that informed robots in a mixed-society of animals (cockroaches) and robots can control the aggregation behavior of the mixed-society through microscopic interactions.

The strategy followed in this paper is similar. The minority is regarded as informed agents who want to have control on the opinion of the society. They should do this through social interactions that take place between local neighbors, i.e. agents that have direct friendship ties. So the minority has to have friendship relations with the majority or initiate new ties upon necessity. But in realistic situations, the total number of links that the minority can initiate is limited to a number, due to e.g. time or geographical distance. Now, the minority should choose which links to initiate in order to gain maximum impact on the society.

The rest of this paper organized as follow: in Section 2 related works are discussed. Section 3 presents the selected model of diffusion. In Section 4 we convert our problem to an optimization and solve it in Section 5. Finally Section 6 consists of simulations that compare our method with well-known heuristics.

2 Related Work

K-MIP tries to manipulate crowd behavior by directly targeting initiators. There are many models that describe diffusion phenomena by using methods from different domains. Based on the selected diffusion model, method of finding influential persons can vary. In this section different diffusion models are described and K-MIP solution for each of them is presented.

In all models, society is modeled by a directed graph $G=(V,E)$ whose vertices V and edges E are representing individuals and social relations respectively. In some models edges are weighted. Weights are usually interpreted as node v's trust on $N(v)$ which is the set of v's neighbors.

The Linear Threshold Model (LTM) that is rooted in mathematical sociology [11] has been widely used in viral marketing [7]. At the beginning, every node v chooses a random threshold $\theta_v \in [0,1]$ that defines its general tendency to adopt a new belief. The link between nodes v and u has weight $b_{u,v}$ where $\sum_{u \in N(v)} b_{uv} \leq 1$. The process begins with a set of active nodes who has adopted the new opinion. In each time step, a node is activated if the weight of its active neighbors exceeds its threshold. The process stops when no new activation is possible.

The Independent Cascade Model (ICM), a well-known model in viral marketing [7] is originated from interacting particle systems [12]. In ICM each active individual has only one chance for activation of its neighbors. The probability of activation of a node v by its neighbor u is equal to the weight of their social connection and is independent of previous attempts of other v 's neighbors. The process starts with a set of initiators and unfolds until all active nodes have used their chance for activation of the others.

Kempe et al. [7] presents an algorithm for K-MIP when diffusions are LTM or ICM. They assume the same costs for activation of each person and showed that the problem is NP-hard. They exploited the submodularity of the problem structure and presented an $(1-1/e)$ - approximation algorithm.

Voter model is another popular model of social influence. In this model at each time step each node picks one of its neighbors at random and adopts its opinion. Even-Dar and Shapira [8] found the exact solution for K-MIP when the underlying interaction model is voter and cost of marketing each person is identical. Also they presented an FPTAS [8] for the case when each person has different costs.

All mentioned methods tackle the problem of maximizing diffusion in social networks by persuading initiators to adopt a product or accept an opinion. But what if there is no way to convince an individual about changing his behavior the way we want? This is the case especially in changing the opinion of a crowd.

This paper investigates the problem of influence maximization from a different view. The problem formulation is changed to a more realistic one. It is assumed that there exist a minority in the society with a different opinion from the majority who tries to propagate its belief by means of making new social relations. Minority has a limited budget and any new link has a cost. So, the problem is converted to finding the best links to be added by minority under the budget constraint.

3 Diffusion Model

As an underlying diffusion model, Local Interaction Game (LIG) [13] is chosen. LIG simultaneously benefits from rigorous game theoretic background and simplicity. In this model each person is under the influence of his neighbors. The person is active if he adopted the minority's opinion and inactive otherwise.

In each relation, participants benefits only if they coordinate and choose the same action. Table 1 summarizes the payoffs of each player in coordination games. For simplicity the zero payoffs is set for incoordination. Person's preferences and tendencies are distinguished by his name index in Table 1.

Table 1 Payoff table of LIG

v \ u	*active*	*inactive*
active	a_v, a_u	$0, 0$
inactive	$0, 0$	b_v, b_u

At the beginning, the society U is inactive except a minority M. The minority intends to activate initiators by adding new links. It can be shown [14] that v become active iff more than $\theta_v = b_v/(a_v + b_v)$ proportion of its neighbors is active. Set A_t is the set of active nodes at time step t.

4 Problem Formulation and Properties

Minority itself can activate a set of nodes A without any cost. Suppose that the effect function $e: 2^M \rightarrow 2^A$ finds the set A which is the union of minority and individuals that are activated by the minority up to the end of diffusion. Thus new links must be added to the members of set $U - e(M)$. Suppose that adding a new link has the constant cost, α, and minority's overall budget is B. Also suppose that for activating each node v, $c(\{v\})$ links must be created from minority to it. Then, activating initiator set S of nodes costs:

$$c(S) = \sum_{v \in S, S \subseteq U - e(M)} \alpha c(\{v\})$$

In this paper, α is considered to be one. For computing *c({v})* values, we take following steps. Each node v has neighbors in M and $U - M$ that are called I and J respectively. Assume that v is inactive, i.e. $|I|/|I \cup J| < \theta_v$. Then set $X \subseteq M$, $X \cap I = \varnothing$ should initiate links to v for its activation. These change the inequality to $\frac{|I \cup X|}{|I \cup J \cup X|} > \theta_v$. Since *c({v})* is the minimum size of set *X* for which the above inequality holds it can be computed as $c(\{v\}) = \left\lfloor \frac{\theta_v |J|}{1 - \theta_v} - |I| \right\rfloor$.

4.1 Optimization Problem

Previous works focused only on K-MIP which is the identical cost MIP [7]. In our problem, each individual's cost can differ from the others, so the problem can be called N-MIP (Non-identical cost Most Influential Person). We define a set function that takes an initiator set S and maps it to the number of individuals that are going to be activated by the end of the process.

Let $f: 2^{U-e(M)} \rightarrow \mathbb{R}$ map the initiators $S \subseteq U - e(M)$ to the number of active nodes at the end of the process. Then N-MIP problem can be viewed as maximizing f subject to the limited budget. Using above definitions the problem is:

Problem: Find set S that maximize function f subject to the cost constraint:

$$S^* = \arg\max_{S \subseteq U - e(M)} f(S) \qquad s.t. \qquad c(S) \leq B$$

First we show this problem is NP-hard for LIG model. Next we claim $f(S)$ is submodular, so we can exploit maximization algorithms for submodular functions.

4.2 NP Hardness of Efficient Link Addition Problem

Kempe et al. [7] showed that finding K-MIP under LTM is NP-hard. Based on their proof, we show that NP-complete Vertex Cover problem is a special case of K-MIP for LIG model which itself is an especial case of N-MIP. For a graph $G = (V, E)$ and integer k Vertex Cover finds a set $S \subseteq V$ that every edge of G has an endpoint in it. If there is a Vertex Cover S of size k in G then $f(S) = |U - e(M)|$. On the other hand this is the only way that for all settings of thresholds one can deterministically activate all society. So Vertex Cover is an especial case of identical cost most influential person for LIG. Based on definition, K-MIP is an especial case of N-MIP. Since it is proved that Vertex Cover is an especial case of K-MIP, N-MIP is also NP-hard.

4.3 Submodularity of f

f is submodular if it satisfies a natural "diminishing returns" property: adding new element to a subset produce gain which is at least as high as adding that element to a superset [15]. Formally f is submodular iff for every $T \subseteq S$, $f(T \cup \{v\}) - f(T) \geq f(S \cup \{v\}) - f(S)$ holds.

Kempe et al. [7] showed that when diffusion model is LTM, f is a submodular function. We show that LIG is an especial case of LTM so f for LIG is submodular too. In LTM each neighbor u of node v can influence it according to the weight b_{uv} such that $\sum_{u \in N(v)} b_{uv} \leq 1$. Thus v would become active in step $t+1$ if $\sum_{u \in N(v) \wedge u \in A_t} b_{uv} \geq q_v$. If b_{uv} is set to $1/|N(v)|$ the inequality changes to $|A_t|/|N(v)|$ which is the condition of v's activation in LIG. So LIG is a special case of LTM when $b_{uv} = 1/|N(v)|$. Therefore f is submodular for LIG.

5 Optimization Algorithm

Up to this point the link addition problem has been converted to a submodular function optimization problem using game theoretic diffusion models. Submodular function optimization is an active field of research in machine learning. Since submodularity arises in many real world optimization problems many advancements has been made during recent years in this old topic [16]. Nemhauser et al. [15] proved that a simple greedy algorithm is within $(1 - 1/e) \approx 0.63\%$ of maximum when $c(S) = |S|$. For general cost functions [17] showed that for the special case of MAX-COVER problem $(1 - 1/e)/2 \approx 0.31\%$ approximation guarantee is reachable and a $(1 - 1/e)$ guarantee can be achieved using partial enumeration. Recently [18] extended their result to general submodular functions and [19] introduced an online boundary for any algorithm.

Since our problem is reduced to maximizing a submodular function subject to a bounded non-identical cost function, we follow [19] and call our algorithm Efficient Link Addition Strategy (ELAS). The greedy approach proposed by Leskovec et al. [19], iteratively adds nodes to the selected set S by choosing a node v that maximizes $f(S \cup \{v\}) - f(S)/c(\{v\})$. This heuristic is an extension of [15] which uses $f(S \cup \{v\}) - f(S)$ as the selection criteria in each iteration. They showed [19] that choosing best results of one of the mentioned heuristic provides a constant factor approximation. Formally, if NIC be the solution of non-identical cost algorithm that uses $f(S \cup \{v\}) - f(S)/c(\{v\})$ and IC be the solution of identical cost algorithm which uses $f(S \cup \{v\}) - f(S)$, it can be proved that:

$$\max\{f(NIC), f(IC)\} \geq \frac{1}{2}(1-\frac{1}{e}) \underset{S \subseteq U-e(M), c(S) \leq B}{\arg\max} (f(S))$$

6 Experiments and Discussion

We have used heuristics from social science that choose individuals with highest degree and betweenness [21] as the initiators, and compared our method's performance with theirs. To show the advantages of ELAS, three different network models were tested (Table 2). For each of them different parameter settings were tested. For each setting 30 networks were built and diffusion was simulated for 30 randomly chosen thresholds. So for each setting 900 simulations were done.

After building a network, every node chose a random threshold. Then some nodes were randomly selected as the minority and were given opinions opposite to the others. Then different strategies for link addition were used and their impacts were measured. The diffusion continued until no new node could be activated. Society was composed of 400 individuals and minority was 10% of them. The budget limit was 40 (i.e. each minority member could initiate one link on average). For each simulation Social Impact of Minority that is the number of active nodes at the end of the simulation, was recorded.

6.1 Budget Impact on Diffusion and Trends of Diffusion

Fig. 1.a illustrates the effect of budget on success of link addition for ER network whose structural details will be discussed later in this section. The vertical axis shows the number of active individuals at the final time-step of diffusion. As expected, increase in budget would increase the social impact of Minority. It is clear that performance of ELAS is better than degree-based and betweenness-based strategies. Further simulations showed this dominance exists for all mentioned network types and their different parameter settings.

Fig. 1.b compares the methods in different time steps for an ER network. Data is gathered for budget 40. Performances of degree and betweenness heuristics' are close to each other. As it is illustrated, ELAS outperform them in all time steps.

6.2 Effect of Network Structure on Diffusion Process

Experimental results along with analytical demonstrations show the better performance of ELAS in comparison with other link addition strategies. In this part, a closer look is taken to ELAS for finding the structural factors that affects its performance. For this goal, different syntactical networks were built and ELAS performance was tested on them.

Table 2 Network models and their parameter list

Network Type	Parameters
Erd s–Rényi (ER)	P: edge probability
Small World (SW)	A: average degree, R: rewiring probability
Scale Free (SF)	A: average degree, S: initial seed

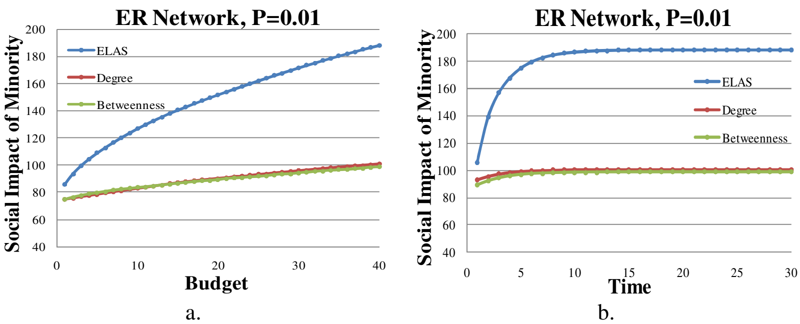

Fig. 1. Comparison of ELAS performance with degree and betweenness based strategy in an instance of ER networks a. At the final time step of simulation. b. During simulation for B = 40.

6.2.1 Erdős–Rényi Network

Erdős–Rényi network model (ER) is the most studied random network model in which the probability of relation between each individual pair is p . Fig. 2 shows the sensitivity of presented method to p . When p is very low the graph is loosely connected which hinders the cascade of influence. As p increases the giant graph component appears and facilitates the diffusion. Since $p = \ln n / n$ is a sharp threshold for the presence of giant component [20] at this point (0.01 for 400 nodes) the effectiveness of the method is maximized. Increasing p creates high degree nodes which are harder to influence. These nodes decrease ELAS influence on the whole network.

Fig 2 also illustrates the difference between ELAS and the better of the two other strategies. As it is clear, the dominance of ELAS decreases as the graph becomes more connected. It can be interpreted as when graph become denser every link addition strategy becomes ineffective.

6.2.2 Scale Free Network

It has been shown [21] that many real world networks are scale free (SF). Based on this, [22] proposed preferential attachment process for generating SF networks. This model has two parameters which are N_0 and k, initial seed of process and average degree of network respectively.

The process starts with N_0 isolated nodes and at every time step a new node is added by making k new links. The probability that a link connects j to node i is linearly proportional to the degree of i [21]: $P(i \rightarrow j) = \deg_i + 1 / \sum_l (\deg_l + 1)$ where deg_i is the degree of node i.

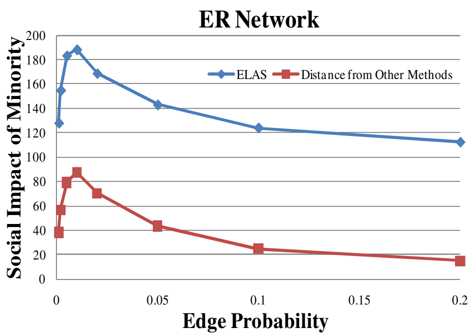

Fig. 2 Effect of network structure on the success of the method for ER network

Fig. 3.a shows the performance of ELAS for different SF networks that have been constructed using different parameters. In this set of experiences 5 value for both N_0 and k have been used. Since for making a SF network we should have

$N_0 \geq k$ no experience has been done for $N_0 < k$ which is shown by zero in Fig 3.a. This reduces the number of different settings for networks from 25 to 15.

Fig 3.b shows the same diagram as Fig 3.a except that it emphasize on the effect of network seed. In each ribbon average degree is constant. In the region of experiences where $N_0 \geq k$, ribbons are flat which shows that N_0 does not have significant impact on ELAS performance. On the other hand ribbons with identical network seed (Fig 4.c) demonstrate that average degree extremely change the number of active individuals at the end of diffusion with inverse relation.

6.2.3 Small World Network

High clustering coefficient (CC) and low average shortest path length (L) are two important characteristics of social relation networks [21]. Small world networks are networks that simultaneously exhibit high CC and low L [23]. Watts and Strogatz model [23] is the most well-known model of small world networks (SW). It has two parameters which are average degree of the network and rewiring probability of edges. The process begins with a ring lattice with n vertices and k edges per vertex. Then edges are rewired with probability p.

Fig. 4.a illustrates the effect of both parameters on ELAS. Fig. 4.b and Fig. 4.c are the same as Fig. 4.a diagram but they illustrate the effect of rewiring probability and average degree respectively. According to flat ribbons of Fig. 4.b, rewiring probability does not have a significant impact on the final result. But it is clear from Fig. 4.c that like previous models, average degree has inverse relation with the final outcome.

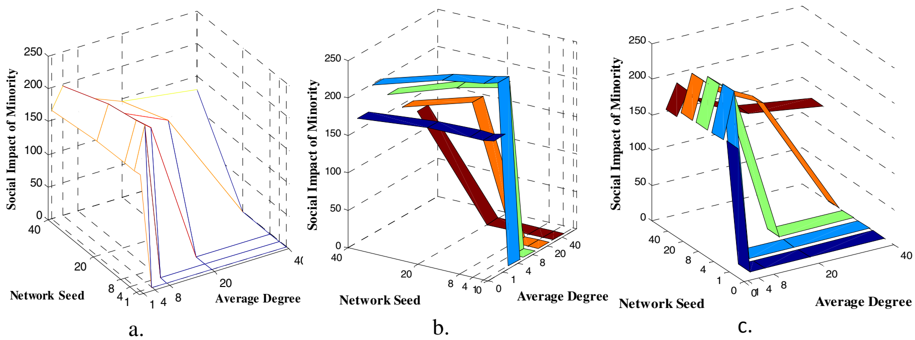

Fig. 3 Effect of network structure parameters on the success of the method for SF networks

6.2.4 Other Structural Factors

At the first glance, result of ELAS in all network structure has inverse relation with the average degree of the nodes. But Fig. 5 shows that for the same average degree, ELAS performs significantly better on scale free network than other networks. This cause to conclude that average degree is not the only structural factor which has impact on ELAS. Further investigations on structural properties of studied networks led to interesting results. Fig. 6.a shows average degree distribution of 30 networks that have been used in Fig. 5. As expected [21], ER network has Poisson like degree distribution with the mean around 4 (Fig. 6.a) and SF network presents power law distribution (Fig. 6.b). Since the process of building SW network begins with lattice of degree 4 and rewiring probability is low (0.01) the SW network has impulse like degree distribution around 4 (Fig. 6.c).

Fig. 7 shows the degree distribution of the sets that have been activated by the minority using ELAS in different network structure of Fig. 5. Degree distribution of activated set for ER network is like Poisson distribution with mean 4 (Fig. 7.a).The activated set in SF network (Fig. 7.b) has high density in lower degrees in contrast with impulse like function of SW network (Fig. 7.c).

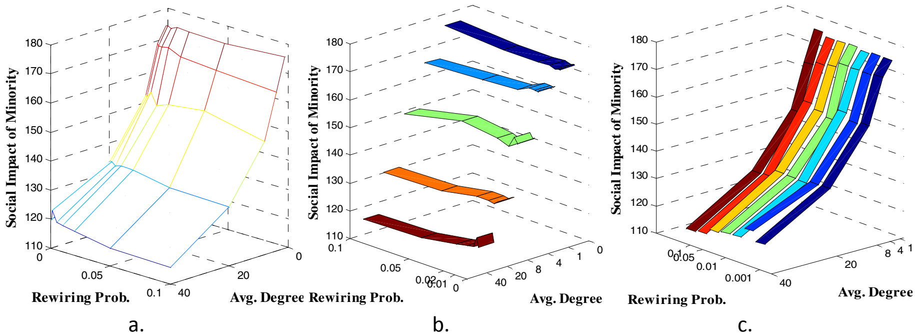

Fig. 4 Effect of network structure parameters on the success of the method for SW networks

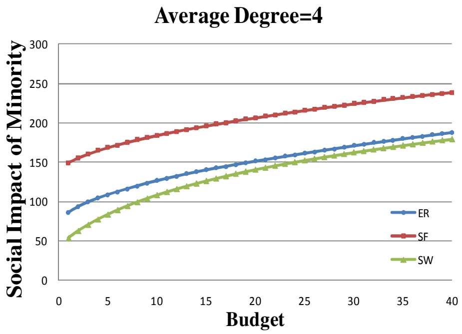

Fig. 5 Effect of network structure for same average degree

From these distributions, it can be concluded that the power of ELAS in SF networks originates from highly available individuals that can be influenced easily not the power of special persons or hubs. This is confirmed by the fact that SW network is in last place in Fig. 5, because SW mostly consists of nodes with degree of 4 which is higher than degree of available nodes in SF and ER networks.

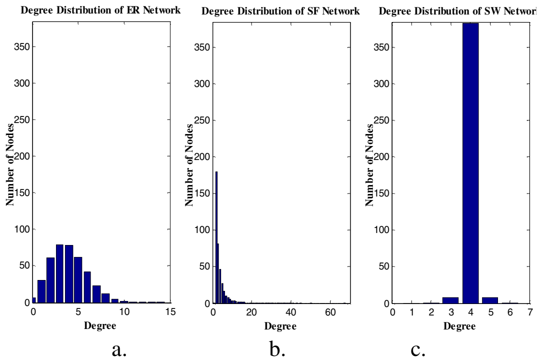

Fig. 6 Degree distribution of different network structure

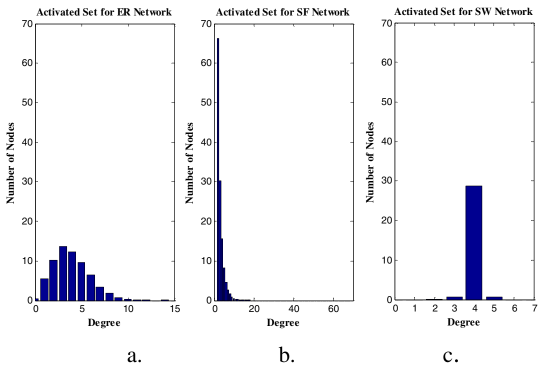

Fig. 7 Degree distribution of activated set for different network structure

7 Discussion and Future Directions

The drastic success of ELAS over other heuristics seems suspicious at the first look. So it should be mentioned that this performance is gained by spending more time for finding the initiators in ELAS. This fact becomes critical when we want to run several experiments for understanding the effect of structure on ELAS performance. In fact our experiments are infeasible for even 500 individuals.

Structure of the greedy algorithm seems neat and efficient but when it comes to the implementation part the main question is how to find $f(S)$. As we stated earlier (IV.a.) $f(S)$ is the expected value of the number of active individuals at the end of the diffusion process. So the very primitive approach to estimate $f(S)$ is to run the diffusion process for many times starting with initiator set S and take the average number of active nodes as the value of $f(S)$. This was the method that used by Kempe et al. [7]. However they mentioned that finding $f(S)$ is the open problem for future research. After this first solution for $f(S)$ estimation, almost all efforts focused to explore the other aspects of cascading and researchers preferred to suppose that $f(S)$ is known by an oracle.

Using simple averaging as the method of computing $f(S)$ will takes hours on a modern server to select 50 seeds in a moderate sized graph (15K nodes and 31K edges) while it becomes infeasible for larger graphs [24]. Even these numbers are too large for running several experiments for exploring structural effects.

Some recent works have developed algorithms for speeding up $f(S)$ calculation with several approaches. The one that is used here is Lazy Forward Evaluation which is introduced in [19] which actually avoids $f(S)$ computing. Leskovec et al. [19] has reported the 700 times speed up in experiments. Very recent methods [24], [25] are developed separately for LTM and ICM. Both of these methods have viewed the influence propagation locally and tried to estimate $f(S)$ as the aggregation of these local cascades. Simulation results show that the final outcomes of greedy algorithm based on these methods for $f(S)$ estimation are always among best results [24], [25].

Another important extension of the naïve influence propagation is the setting in which multiple minorities exist in the society and compete with each other for adding new links and change the crowd behavior. This domain is very novel even in the context of finding K-MIP which as mentioned in IV.a is simpler than link addition problem. They are some recent works which addressed competitive setting for K-MIP problem [26], [27].

8 Conclusion

Changing belief of the majority of individuals by means of a minority was the main focus of this work. Each individual's belief is emerged from his neighbors by a simple rule that has selfishness in its nature. Based on this rule belief change propagates through the society. Minority want to change the belief of majority by making new relation with them. We leave the competitive scheme in which there exist many minorities competing on influence maximization, for future works.

A greedy algorithm was presented for finding the best new relation and its performance was compared with different relation initiation strategies. Our method, ELAS, outperformed them along with its rigorous mathematical background that shows its performance is within the specified distance of the optimal solution.

Also the effect of structure on ELAS was measured and it was found out that degree of each individual is the main parameter that impact ELAS with an inverse relation. In addition it was shown that in a population with the same average degree, number of available easily-influenced individuals is more important than influencers for the success of diffusion.

References

[1] Sobkowicz, P.: Modeling Opinion Formation with Physics Tools: Call for Closer Link with Reality. In: JASSS (2009)

[2] Jurvetson, S.: What Exactly Is Viral Marketing? Red Herring, 110–111 (2000)

[3] Leskovec, J., Adamic, L.A., Huberman, B.A.: The Dynamics of Viral Marketing. In: Proc. of the 7th ACM Conf. on Electronic Commerce, pp. 228–237. ACM, Ann Arbor (2006)

[4] Borgatti, S.P.: Identifying Sets of Key Players in a Social Network. Comput. Math. Organ. Theory 12, 21–34 (2006)

[5] Valente, T.W., Davis, R.L.: Accelerating the Diffusion of Innovations Using Opinion Leaders. The Annals of the American Academy of Political and Social Science 566, 55–67 (1999)

[6] Lazarsfeld, P.F., Berelson, B., Gaudet, H.: The People's Choice: How the Voter Makes up His Mind in a Presidential Campaign. Columbia University Press (1944)

[7] Kempe, D., Kleinberg, J., Tardos, É.: Maximizing the Spread of Influence through a Social Network. In: Proc. of the 9th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, pp. 137–146. ACM, Washington, D.C (2003)

[8] Even-Dar, E., Shapira, A.: A Note on Maximizing the Spread of Influence in Social Networks. In: Deng, X., Graham, F.C. (eds.) WINE 2007. LNCS, vol. 4858, pp. 281–286. Springer, Heidelberg (2007)

[9] Couzin, I.D., Krause, J., Franks, N.R., Levin, S.A.: Effective Leadership and Decision-making in Animal Groups on the Move. Nature 433, 513–516 (2005)

[10] Caprari, G., Colot, A., Halloy, J., Deneubourg, J.L.: Building Mixed Societies of Animals and Robots. IEEE Robotics & Automation Magazine 12, 58–65 (2005)

[11] Granovetter, M.: Threshold Models of Collective Behavior. The American Journal of Sociology 83, 1420–1443 (1978)

[12] Durrett, R.: Lecture Notes on Particle Systems and Percolation. Brooks/Cole Pub. Co. (1988)

[13] Morris, S.: Contagion. Review of Economic Studies 67, 57–78 (2000)

[14] Easley, D., Kleinberg, J.: Networks, Crowds, and Markets: Reasoning About a Highly Connected World. Cambridge University Press (2010)

[15] Wolsey, L.A., Fisher, M.L., Nemhauser, G.: An Analysis of Approximations for Maximizing Submodular Set Functions-I. Mathematical Programming 14, 265–294 (1978)

[16] Krause, A.: Near-optimal Observation Selection Using Submodular Functions. In: AAAI NECTAR (2007)

[17] Khuller, S., Moss, A., Naor, J.: The Budgeted Maximum Coverage Problem. Inf. Process. Lett. 70, 39–45 (1999)

[18] Krause, A., Guestrin, C.: A Note on the Budgeted Maximization of Submodular Functions. pp. 05–103 (2005)

[19] Leskovec, J., Krause, A., Guestrin, C., Faloutsos, C., VanBriesen, J., Glance, N.: Cost-effective Outbreak Detection in Networks. In: Proc. of the 13th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, pp. 420–429. ACM, San Jose (2007)

[20] Erdos, P., Renyi, A.: On the Evolution of Random Graphs. Publ. Math. Inst. Hung. Acad. Sci. 5, 17–61 (1960)

[21] Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.: Complex Networks: Structure and Dynamics. Physics Reports 424, 175–308 (2006)

[22] Barabási, A., Albert, R.: Emergence of Scaling in Random Networks. Science 286, 509–512 (1999)

[23] Watts, D.J., Strogatz, S.H.: Collective Dynamics of "Small-World" Networks. Nature 393, 440–442 (1998)

[24] Chen, W., Yuan, Y., Zhang, L.: Scalable Influence Maximization in Social Networks under the Linear Threshold Model. In: IEEE International Conference on Data Mining, ICDM (2010)

[25] Chen, W., Wang, C., Wang, Y.: Scalable Influence Maximization for Prevalent Viral Marketing in Large-scale Social Networks. In: Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD 2010, Washington, DC, USA, p. 1029 (2010)

[26] Bharathi, S., Kempe, D., Salek, M.: Competitive Influence Maximization in Social Networks. In: Deng, X., Graham, F.C. (eds.) WINE 2007. LNCS, vol. 4858, pp. 306–311. Springer, Heidelberg (2007)

[27] Pathak, N., Banerjee, A., Srivastava, J.: A Generalized Linear Threshold Model for Multiple Cascades. In: IEEE International Conference on Data Mining, ICDM (2010)