eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Optimizing Humanitarian Aids: Formulating Influencer Advertisement in Social Networks

Nastaran Hajiheydari[1], Masoud Salehi[1], Arman Goudarzi[1]

[1] University of Tehran, Faculty of Management, Tehran, Iran
(nhheidari, masuod.salehi, goodarzi_arman) @ut.ac.ir

Abstract. In order to solve problems encountered during natural disasters, in addition to NGOs and relief teams, various individuals intend to help the injured. Although the cooperation of people has remarkable advantages, the disparity between the needs of the injured and the people's donations can cause problems such as trouble for relief teams and wasting the substantial resources. In generic, the influencer selection in the marketing endeavors is mainly aimed to maximize people's awareness and attention, but this research proposes a method for influencer selection, using Social Network Analysis (SNA) and optimization techniques, by which it is possible to establish an adaptation between the public attention and the type of injured necessities. The proposed method is applied to a real sample network of Facebook friends, to evaluate the efficiency and validity of the formulated method.

Keywords: Disaster management, Humanitarian Aid, Advertisement, Influencer selection, Social Network Analysis, Optimization, Influence maximization

1 Introduction

Natural disasters always have been one of the human concerns. The statistics show that between 1980 and 2004 two million people were killed and 5 billion people affected by natural disasters in total [1]. To solve the problems encountered during natural disasters, various people and organizations such as local and International Humanitarian Organizations (IHO), government and military get involved in helping the injured [2, 3]. Meanwhile, the benevolent people are also willing to contribute to solving the problems by providing different kinds of donations. This cooperation alongside its various benefits, can increases the complexity of coordination and implementation of disaster relief programs, especially when donations be unsolicited [2, 4]. This unrequested supplies and donations can cause troubles such as waste relief teams' time, resources, congest the relief chain and create bottlenecks [5]. On the other hand, previous researches have emphasized the role of media attention on increasing these issues [2, 3]. It has been explained that increasing media attention will increase unsolicited donations and consequently, the problems created for relief teams will be intensified [2]. Also, it is completely significant for donors to perceive that their donations are used in a way that has the greatest possible benefit for the injured [2]. In fact, donors are somehow the customers of relief organizations and thereupon, they

must be satisfied to continue their co-operation [4]. Besides, it is proven that the media's attention, especially online social networks, can be used as a powerful platform to attract more donations [4, 6]. For example, After the Haiti earthquake during 48 hours Red Cross received around 8 million dollars of charity through people's texts in social media [7].

The role and importance of social media in attracting help and donations in one side and the obstacles of receiving undemand aids on the other hand, highlights the need for managing and directing the media attention in the event of natural disasters. This research aims to propose a method for influencer selection, using the Social Network Analysis (SNA) and optimization techniques, through which it would be possible to establish an adaptation between the public attention in social media and the type of injured necessities. In this way, current research would come up with a solution to prevent mentioned problems in relief programs. In this paper, Influencer refers to individuals who have some personal or network desirable attributes such as expertise, credibility, and connectivity enabling them to persuade people in their network [8]. Influencers can shape others attitudes by their power of authenticity and reach [9].

2 Social media in disaster relief management

With the development of online social networks, a variety of applications of these networks for disaster relief management also are emerged. Crisis management organizations, are using data extracted from social networks to determine where aid are needed and to broadcast their own needs [10]. As instance, Ushahidi (www.Ushahidi.com) is an open source platform providing crisis map to show the different needs of injured by their locations [7]. A great deal of research has been made in this regard. A survey showed that Twitter activity after a hurricane has a strong correlation with per-capita economic damages, and it suggested that the online social networks could be used for rapid assessment of damage caused by a large-scale natural disaster [11]. Based on Zhu et al. (2011) the understanding of the factors that influence the behavior of Twitter users in retweeting tweets on natural disasters will make it easier for disaster managers to choose optimal messages to propagate in the network. In other words, they focused on how to choose the best content to speed up the information propounding the disaster to help injured [12]. In a comprehensive review, Landwehr et al. (2014) divided the social network analysis and efforts in crisis management into four categories, keyword-based labeling, crowdsourcing-based labeling, sentiment based labeling and network analysis for relevance [10]. Based on literature analysis, the social media have been considered as an influential channel to engage people and donors in crisis management during natural disasters.

3 Method

In order to develop a method for influencer selection, using the Social Network

Analysis (SNA) and optimization techniques to choose the best set of influencers in a complex network is necessary. In this way, an adaption between actual injured needs and public awareness of each kind of the needs through an online social network would be possible. This will be done in three subsections: in the first step, the diffusion model of a message through a social network is discussed. In the second subsection, our subject is differentiated from a typical influence maximization problem and then the problem is formulated to obtain the cost function. And finally in the last section, the optimization algorithm to make the adaptation is applied.

3.1 Diffusion model

As discussed, a social network can be a powerful tool for spreading information among individuals. So far, many scholars have studied the information diffusion model in a social network [13], including researches on the diffusion process in "word of mouth" [14] and "viral marketing"[15]. Influence maximization is also considered as an attractive topic of research in the marketing field, which researchers are trying to maximize the distribution of a message on a network by finding the most influential nodes [16-19]. Meanwhile, David Kempe et al. (2003) formulated this issue by presenting two stochastic diffusion models [17], which have become a basis for many researchers in this field [16]. In this paper, one of Kempe et al. (2003) diffusion models, called Independent Cascade, is used as the basis for our model.

Let's assume $G = (N, E)$ as a directed social network where N represents the set of nodes (individuals) and E represents set of edges (relationships between individuals). In this network, each node is considered as being active or inactive (active nodes have received the message). Now we start with a set of initial active nodes A_0. In the first step $t = 0$ each active node like V has a single chance to activate its inactive neighbor W, with probability P_{VW}. At the end of this step if any of the active nodes succeed to activate any of its neighbors, diffusion process will continue in the step $t = 1$ with the set of newly activated nodes A_1, otherwise the entire diffusion process model will be stopped. It should be noted that in this case, each node could only activate the neighbors who follow him (followers), and can be activated by his neighbors whom he follows (following). Therefore, it is necessary to consider the direction of the edges. For example in Figure 1, node V can be activated by node Y with probability P_{YV} and can activate node W and U respectively with probability P_{VW} and P_{VU}.
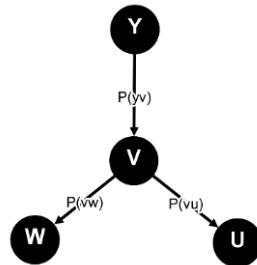

Fig. 1. A sample graph with 4 nodes

In this paper, a real data set of 347 friends on Facebook is utilized to test and run algorithms. This data set is free and available on www.snap.stanford.edu/data.

3.2 Formulating problem

Again, let's assume G = (N, E) as our network where $N = (n_1 \quad \cdots \quad n_i)$ represents set of nodes. In this network, each node like Y can activate all other nodes in t steps with different probabilities. For example in Figure 1, Y can activate W in two steps with probability $P_{YW} = P_{YV}.P_{VW}$. In this way, matrix G_P illustrates the chance of each node to activate any other one.

$$G_P = \begin{matrix} & n_1 & n_2 & \cdots & n_i \\ n_1 \\ n_2 \\ \vdots \\ n_i \end{matrix} \begin{bmatrix} P_{n_1 n_1} & P_{n_1 n_2} & \cdots & P_{n_1 n_i} \\ P_{n_2 n_1} & P_{n_2 n_2} & & P_{n_2 n_i} \\ \vdots & & \ddots & \vdots \\ P_{n_i n_1} & P_{n_i n_2} & \cdots & P_{n_i n_i} \end{bmatrix} \tag{1}$$

Figure 2 represents an example of the probability of activating other nodes by initially activated node n_{56} in the network of Facebook friends.
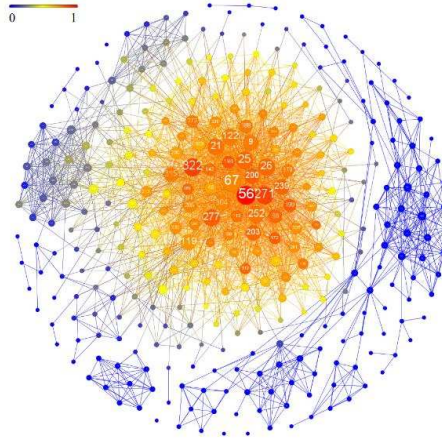
Fig. 2. Influence of sample node "n56"

Now we need to consider set $A = (a_1 \quad \cdots \quad a_j)$ as set of different needs of injured (items), and $V = (v_1 \quad \cdots \quad v_j)$ as set of the volume of each item, which is estimated by relief teams.

The objective is to select a set of influencers S in such a way that the level of awareness of each item through the network be adapted to the volume of actual needs.

$$S = (n_{a_1} \quad n_{a_2} \quad \cdots \quad n_{a_j}) \tag{2}$$

In above set of influencers, each element like n_{a_1}, represents the influencer who is in charge for influence item a_1. As argued, each influencer (each node) can activate all other nodes, so it is considered that there is a possibility for all nodes to donate any item. Matrix G_A represents this issue:

$$G_A = \begin{array}{c} \\ n_{a_1} \\ n_{a_2} \\ \vdots \\ n_{a_j} \end{array} \overset{\begin{array}{cccc} n_1 & n_2 & \cdots & n_i \end{array}}{\begin{bmatrix} P_{n_{a_1}n_1} & P_{n_{a_1}n_2} & \cdots & P_{n_{a_1}n_i} \\ P_{n_{a_2}n_1} & P_{n_{a_2}n_2} & & P_{n_{a_2}n_i} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n_{a_j}n_1} & P_{n_{a_j}n_2} & \cdots & P_{n_{a_j}n_i} \end{bmatrix}} \quad , \ j \le i \tag{3}$$

In above matrix, $P_{n_{a_1}n_1}$ is the probability for n_1 to be activated by n_{a_1} regardless to the presence of other influencers. But n_1 can be activated by other influencers as well. It means that the influence boundaries of influencers may overlap in some nodes. So, the probability that n_1 will be going to donate a_1 in presence of other influencers, is calculated as below:

$$P_{a_1 n_1} = P_{(n_1 \ be \ activated)} \cdot P_{(n_1 \ choose \ a_1)} \tag{4}$$

$$P_{(n_1 \ be \ activated)} = 1 - \prod \left(1 - P_{n_{a_j}n_1}\right) \tag{5}$$

$$P_{(n_1 \ choose \ a_1)} = \frac{P_{n_{a_1}n_1}}{\sum P_{n_{a_j}n_1}} \tag{6}$$

By repeating the above for all nodes, it is possible to calculate the probability for each node who will donate a_1 in presence of other influencers. These calculations can be again repeated for other items to obtain matrix GD.

$$GD = \begin{array}{c} \\ a_1 \\ a_2 \\ \vdots \\ a_j \end{array} \overset{\begin{array}{cccc} n_1 & n_2 & \cdots & n_i \end{array}}{\begin{bmatrix} P_{a_1 n_1} & P_{a_1 n_2} & \cdots & P_{a_1 n_i} \\ P_{a_2 n_1} & P_{a_2 n_2} & & P_{a_2 n_i} \\ \vdots & & \ddots & \vdots \\ P_{a_j n_1} & P_{a_j n_2} & \cdots & P_{a_j n_i} \end{bmatrix}} \quad , \ j \le i \tag{7}$$

$$P_{(Total \ probability \ of \ donating \ a_j)} = \sum_{1}^{i} P_{a_j n_i} \tag{8}$$

To clarify the difference between this matrix and the previous one, it should be noted that in the previous matrix the effect of the presence of other influencers was not taken into account. So any change in influencer set S may cause a change in all elements of the matrix GD and consequently change in total probability for donating all items.

Therefore, it is complicated to choose the best set of influencers in such a way that total probability of donating different needs in the network, adapt to the actual volume of each item. This issue necessitates deploying optimization algorithms in the next section.

3.3 Optimization algorithm

The first step to solve an optimization problem is calculating the cost function. The cost function is defined as the distance from the current state to the optimal condition. Therefore, the more this function gets closer to zero, the more favorable the situation is. In this case, the favorable situation is the equality of the actual amount of demand per item with the total probability of donating that item. So, the cost function is as below:

$$CostFunction : \sum_{1}^{j} \left| v_j - \sum_{1}^{i} P_{a_j n_i} \right| \tag{9}$$

In above function, $V = (v_1 \quad \cdots \quad v_j)$ is the actual needed amount of each item, which is estimated by the relief teams. Now by employing an optimization algorithm, we can find the best set of influencers to minimize the difference between the actual need of an item and total probability for donating that item (minimizing cost function).

3.4 Quantum Genetic Algorithm (QGA)

This study deployed an evolutionary Genetic Algorithm, QGA, which is based on quantum computing and quantum theory concepts to minimize the cost function. Researchers have already shown a variety of positive features for this algorithm, including low computing time, powerful search capability and fast convergence [20].

Considering the complexity of the network of Facebook friends and the fact that the ID of the nodes in the dataset does not follow a particular order, which made the network more complicated, this algorithm, QGA, is used for optimization. For the initial random population, after several trials and errors, we have chosen 50 chromosomes and 100 iterations as it leads the results to an acceptable level of convergence and efficiency. Each chromosome contains 5 genes, which in this case represents a set of 5 influencers. So, each gen can get an integer number between 1 and 347 (Population).

4 Results

As explained, in solving influencer maximization problems, researchers are trying to select the most influential set of influencers, whereas, this research intended to solve another type of problem. Therefore, the result of this study is compared with two other methods of selecting the influencers. The first method is simple random selection, which can be considered as the result of decision making without having the network information (Random Selection). In the second method, the criteria for selecting

influencers are the influential power of the individuals in the absence of the other influencers. It means that if we need the awareness of item a1 to be v1, we will simply choose the node with influential power equal to v1 (Fast Selection).

To compare the result of the methods, a set of 5 needs as $V = (10 \quad 30 \quad 14 \quad 23 \quad 8)$ is considered for the population of network = 347. So, 22.7 percent of the network needs to get activated, tailored to the needs. In table 1, the results of the proposed method is compared with two other ones. We repeated QGA selection method 100 times to get better results and also to check the reliability of this method. Additionally, for the random selection method, 100 different sets of influencers have been chosen randomly. But regarding the fast selection method, only one set of influencers, which is the closest set based on the influential power to the set of injured needs V, was picked up.

Table 1. Results

Selection method	Results				
	Min. Cost	Average Cost	St. Dev.	RMSE	Efficiency
QGA Selection	**1.013**	**3.585**	**1.744**	**3.99**	**95.78 %**
Fast Selection	**14.089**	**14.089**	**0**	**14.098**	**83.42 %**
Random Selection	**55.308**	**68.601**	**6.511**	**67.65**	**19.29 %**

In Table 1, the term Efficiency shows how much a method is successful to make the adaption between the awareness in the network and the set of injured necessities and is calculated as bellow:

$$Efficiency : \ 1 - \frac{Cost}{\sum_1^j V_j} \tag{10}$$

Table 1 illustrates a significant 12.3 percent difference between the efficiency of QGA selection method and the Fast Selection method. This difference emphasizes the impact of presence of different influencers in the same network which can cause conflict of information diffusion boundaries.

4.1 Decision-making framework

Based on the aforementioned, it is assumed that each nodes could potentially be an influencer, therefore accessing and communication with them is not a serious problem. Therefore, in the results section, the algorithm only shows the best results. But in a more realistic approach in a network like G, some nodes may be not available or might

be not willing to spread the information on the network. Therefore, the results of the current optimization may be inadequate for decision makers to choose the best set of influencers. In such circumstances, decision makers can check all network nodes in term of availability before running the algorithm. As checking the availability of all the nodes in a big network can take a long time, it is not recommended especially in crisis situations, Another solution is that decision makers be able to pass on to the next best answer in the case of unavailability of any of influencers in the first best set.

On the other hand, in disaster management, "time" is always a key factor. The relief teams estimation of the volume of different items is mainly time related. Therefore, their schedule to expand awareness of the necessities through the network has some time boundaries. To formulate our problem, it is considered that each node's diffusion process is reached to its final state and cannot extend more. But it could take several days for a node as an influencer to reach to its final state (see figure 3).
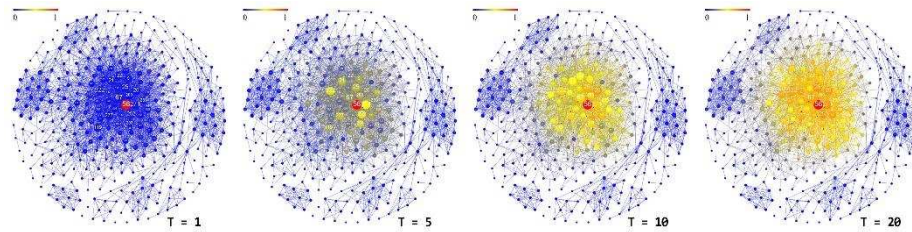

Fig. 3. Information diffusion over time

Hence, adding the issue of time to the problem and the availability of nodes are also important for the optimal result. It means that best set of influencers may be different from a time boundary to another one, and the best set of influencers in a time boundary may be not available. Therefore, the results should be made available to decision-makers in a way so that they can make better decisions according to different circumstances. In figure 4, a chart that helps decision makers make flexible decisions in specific circumstances is proposed, where each circle represents a distinct set of influencers. Decision makers can choose the best set of influencers according to time constraints, and if any of influencers was not available, they will easily pass on to the next best set.
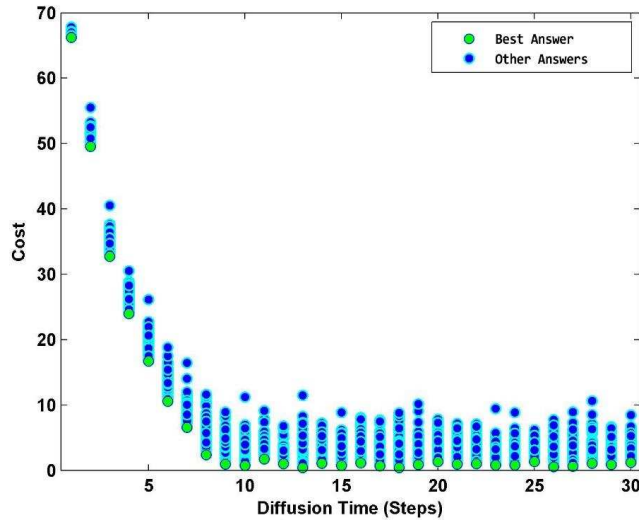
Fig. 4. Decision making framework

5 Conclusion

In this paper, the importance of managing people's awareness of the injured necessities was presented as one of the challenges of disaster management. In the case of inappropriate information diffusion of needs, philanthropists with inappropriate donations, in addition to wasting their donations, might congest the relief chain. Formulating message propagation on social networks in order to create adaptive awareness is the main contribution of this paper. Time constraints in crisis management and influencers' willingness to collaborate on spreading of the messages, were considered as two important issues to find the best answer, which well answered using a decision making framework.

Determining the actual set of nodes and edges that shapes network structure, till now, is the substantial limitation for this type of research. For further research, the proposed model could be validated in a real disaster context. Also, it is suggested that the impact of the presence of influencers in other social networks and effect of traditional media be considered in formulating this problem.

References

1. Strömberg, D., Natural disasters, economic development, and humanitarian aid. Journal of Economic perspectives, 2007. 21(3): p. 199-222.

2. Besiou, M., O. Stapleton, and L.N. Van Wassenhove, System dynamics for humanitarian operations. Journal of Humanitarian Logistics and Supply Chain Management, 2011. 1(1): p. 78-103.
3. Van Wassenhove, L.N., Humanitarian aid logistics: supply chain management in high gear. Journal of the Operational research Society, 2006. 57(5): p. 475-489.
4. Balcik, B., et al., Coordination in humanitarian relief chains: Practices, challenges and opportunities. International Journal of production economics, 2010. 126(1): p. 22-34.
5. Russell, T.E., The humanitarian relief supply chain: analysis of the 2004 South East Asia earthquake and tsunami. 2005, Massachusetts Institute of Technology.
6. Martin, J.A., Disasters and donations: The conditional effects of news attention on charitable giving. International Journal of Public Opinion Research, 2013. 25(4): p. 547-560.
7. Gao, H., G. Barbier, and R. Goolsby, Harnessing the crowdsourcing power of social media for disaster relief. IEEE Intelligent Systems, 2011. 26(3): p. 10-14.
8. Bakshy, E., et al. Everyone's an influencer: quantifying influence on twitter. in Proceedings of the fourth ACM international conference on Web search and data mining. 2011. ACM.
9. Freberg, K., et al., Who are the social media influencers? A study of public perceptions of personality. Public Relations Review, 2011. 37(1): p. 90-92.
10. Landwehr, P.M. and K.M. Carley, Social media in disaster relief, in Data mining and knowledge discovery for big data. 2014, Springer. p. 225-257.
11. Kryvasheyeu, Y., et al., Rapid assessment of disaster damage using social media activity. Science advances, 2016. 2(3): p. e1500779.
12. Zhu, J., et al. Statistically modeling the effectiveness of disaster information in social media. in Global Humanitarian Technology Conference (GHTC), 2011 IEEE. 2011. IEEE.
13. Goldenberg, J., B. Libai, and E. Muller, Using complex systems analysis to advance marketing theory development: Modeling heterogeneity effects on new product growth through stochastic cellular automata. Academy of Marketing Science Review, 2001. 2001: p. 1.
14. Goldenberg, J., B. Libai, and E. Muller, Talk of the network: A complex systems look at the underlying process of word-of-mouth. Marketing letters, 2001. 12(3): p. 211-223.
15. Richardson, M. and P. Domingos. Mining knowledge-sharing sites for viral marketing. in Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining. 2002. ACM.
16. Mohammadi, A. and M. Saraee, Finding influential users for different time bounds in social networks using multi-objective optimization. Swarm and Evolutionary Computation, 2018.
17. Kempe, D., J. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. in Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. 2003. ACM.
18. Barbieri, N. and F. Bonchi. Influence maximization with viral product design. in Proceedings of the 2014 SIAM International Conference on Data Mining. 2014. SIAM.
19. Liu, B., et al., Influence spreading path and its application to the time constrained social influence maximization problem and beyond. IEEE Transactions on Knowledge and Data Engineering, 2014. 26(8): p. 1904-1917.
20. Zhang, G.-x., et al., Novel quantum genetic algorithm and its applications. Frontiers of Electrical and Electronic Engineering in China, 2006. 1(1): p. 31-36.