University of Hormozgan

International Journal of Industrial Engineering & Management Science

journal homepage: www.ijiems.com

A New Marketing Strategy to Identify Influencers Based on Users' Interests in Social Networks

Zeynab Samei [a,*], Mohsen Gharakhani [b]

[a] *Department of Computer Science, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran*
[b] *Department of Finance, Iranian E-Institute of Higher Education, Tehran, Iran*

ARTICLE INFO

Article history:
Received: 2020-12-29
Received in revised form: 2021-09-02
Accepted: 2021-09-14

Keywords:
Social Network
Viral Marketing
Influence maximization
User interest
Influential spreaders
Marketing Strategy

ABSTRACT

Social networks are becoming an easy to use platform for viral marketing that are much more powerful and fast in propagating considered information in different topics. To this end, identification of influential users in social networks plays a crucial role in a successful viral marketing. Most of existing influential maximization methods are based on structural properties of networks. Whereas there are more personal information such as users' interests and friends' common interests that affects the behavior of users in confronting the shared massages. This manuscript proposes a novel method to identify the influential users for marketing in social networks based on their specific interests and power of influencing on their neighbors. We claim that not any hub node can be chosen as the influential spreader in the considered marketing contents and the influential users should be chosen based on their interests topics obtained from their historical activities. We propose a new method to identify the most influential users. In the proposed method, the extent of interest of the influential nodes and their neighbors are considered and the SIR spreading model is used to investigate the spreading process. Experimental results on six real social networks reveal effectiveness of the proposed method as compared to the existing methods based on centrality measures.

1. Introduction

Growing the Internet technology results in rapid spreading of rumors, advertisements and in online social networks that is replacing the traditional communication means rapidly [1] and is converting the social networks to the most applicable advertising platforms[2, 3]. If the influential users can be selected carefully, they would spread the messages widely by spreading them among their neighbors and friendship networks. Obviously, all the users are not important similarly and different features including structural properties and personal information are applied to determine the extent of influentiality of users in social networks. The popular method that has been used in recent years, is to choose the influential nodes is to rank them based on these features and choose the k-most influential users [4, 5] .In most studies, the only available information is the structural properties of the networks which results in many structural-based centrality measure [6-10]. But there is more information behind the connections in social networks. People are mainly joining the communities that have common interests with them and they may become more impressive by viewing their shared contents [11, 12]. The wide range of exchange of information and large number of users interacting in sharing the information in social networks, has converted social networks to a powerful tool to spread advertisement and considered contents. So more practical methods should be employed to select the best users as spreaders to propagate the advertisement widely. In most of studies, advertising and marketing strategies in social networks are mainly based on structural features of users [13-15]. But recent studies have paid much attention to other aspects of social networks behavior such as users interests and the extent of trust between them [16-18]. In this paper we are concentrating on the users interest and its relation with advertisement categories and the power of information propagation based on neighbors common interests. Based on some previous studies, users who are interested in specific topics are more willing to share the information and have more marketing power [19, 20].

* Corresponding author.
E-mail address: z.sameie@gmail.com

We suggest a new approach to determine the initial candidates for identification the influential spreaders.

- We claim that not any hub node can be chosen as the influential spreader in the considered marketing contents and the influential users should be chosen based on their interests topics obtained from their historical activities (post, comments, etc).
- We propose a new method to identify the most influential users. In the proposed method, the extent of interest of the influential nodes and their neighbors are considered and the SIR spreading model is used to investigate the spreading process.

The rest of the paper is organized as follows. In Section 2, the motivations and related works to social networks and user interests are reviewed. In Section **Error! Reference source not found.**, the informal definition of the problem is presented. In Section 4, the problem is formally defined, and the proposed method is explained. The information about the datasets and evaluation metrics and the results and analysis of the parameters is presented in Section 0 .Finally, in Section **Error! Reference source not found.** summery is provided.

2. Related works

Viral marketing has become an effective tool for spreading information about new products and advertising them in social networks. The social networks are growing fast and becoming complex, and besides that identification of user characteristics and interests is becoming more challenging. Thus, identification of real influential nodes considering user characteristics for maximal network coverage in online product marketing has become a key issue which has been known as *influence maximization* in network theory. The influence maximization problem in social networks is to identify a set of k prominent users such that if a message (news, advertisement, information, etc.) is initially reached to this set, it then reaches to the maximum number of other network users via a spreading model. Many topological-based methods have been proposed to determine the power of nodes to be influential. This includes degree, betweenness [7], k-shell [9] , Eigen value, entropy based [21], and the combination of different methods. But there are some considerations while using these methods. Because in most cases there are high overlap between nodes that are covered via each of the selected influential nodes [22]. Recent studies show a great extent of attention to new aspects of effective parameters in identifying the influential spreaders. These parameters are beyond the topological features of nodes that can be reached via network structure. Specifically, in social networks, the personal attributes, characteristics and interest of each user in different shared contents should be considered whilst looking for maximum network coverage. Basically, we can categorized the influential spreaders identification methods to two main category of ordinary methods which are based on network structure and hybrid methods which consider both network structure and nodes characteristics. In most of Ordinary methods that have been introduced and modified many times in recent decades, the topological location of the node defines its power of influentiality in the network. The considered method is computed for all nodes and the top-k nodes with the highest rank are candidate for the most influential nodes. Some of the widely-presented measures that are used in our experiments are as follows:

Degree centrality (D) [23] : Degree of a node is defined as the total number of connections between that node and the others in the network. Based on this measure, the more degree of a node is, the higher is its influentiality. Which means that if a node with a higher degree is chosen as the information spreader, it is supposed that more nodes can reach that content.

k-shell centrality (KS)[8] : In this measure, the nodes are ranked based on their distance to the core of the network. The idea behind this definition is that anode with high degree that is not near to the core of the network should not be chosen as an influential spreader, since it is not capable to activate large amount of nodes around. So the nodes which are closer to the core of network and have higher k-shell are considered to be more influential.

Neighbors' k-shell centrality (NKS): This measure identify the power of the influentiality of a node based on its neighbors' k-shell. Since there may be many nodes in the same shell that are ranked the same, this measure is proposed to investigate the dispersion power of the nodes based on their neighbors rank. So, this is a hybrid method based on both Degree and k-shell centrality measure.

Neighbors Degree centrality (N): In this measure the sum of the node's neighbors' degree is considered in ranking the nodes and the one with the higher rank is chosen as an influential spreader.

Overlay strategy (O)[13]: This measure is an improved measure based on the degree centrality. The idea behind is that nodes with high degree that are neighbors should not be both chosen as influential spreaders and it is more efficient to disperse the influential nodes set in whole the network. So, the algorithm is iterated k times, and each time the node with the highest degree that is not neighbor with any of the other chosen nodes (in the seed set) is selected and added to the seed set

Nodes and Neighbors Degree centrality (KN): These measurements consider the multiplication of the degree of the node and degree of its neighbors as the ranking criteria and the nodes with the highest score are chosen.

In the context of identification of influential nodes in social networks and viral marketing, Mochalova and Nanopoulos [24] proposed a targeted approach to viral marketing based on local centrality measures. Yang et al. [25] considered Both online and offline interactions of users in identification of influential nodes. Cha et al. [26] investigated the dynamics of user influence based on topics and time via available information of indegree, retweets and mentions in Twitter network. Dave et al. [27] formulated the identification of influencers as a problem of predicting the extent of cascades that any node can trigger. Pei et al. [28] proposed a method to find influential spreaders via considering the real spreading dynamics of networks and found that some widely-used centrality measures such as degree and PageRank fail in ranking users' influence in social networks. Al-garadi et al. [29] presented an improved version of the K-core method for online social networks considering the interaction among users. Berahmand et al. [30] suggested a local ranking method to identify the influence of the nodes based on different similarity measures. They also suggested a new centrality measure based on the negative and positive effects of the clustering coefficient for identifying influential spreaders in complex networks[31] and examined the effect of rich-club on diffusion in complex networks [32]. In another work Zareie et al.

[33]proposed an improved cluster rank approach that considered hierarchy of nodes and their neighborhood set in social networks.

There are many other methods that are defined based on the structural properties of the networks. Besides these measures, recently new approaches have been proposed to combine the personal characteristics of the nodes with their structural features specifically in social networks. In other words, personal information about the users in social network play a crucial role in determining their influentiality. As it was mentioned by Abel et al. [34], the contents that are presented via comments and post by users can finely determine their interests domain, concluding that users' interest in a specific topic results in their interest in products related to that topic. Besides that, users may have different extent of interests in various topics. So, considering the users' characteristics and interest play an important role in choosing the right influential people in social networks. Zhu [35] introduced a method based on user interests and trust between users. Liu et al. [20] proposed a new approach to user interest in different topics based on their historical activities and involved the trust factor between users too. Wu et al. [36] proposed an information-spreading model based on Game theory in multiplex networks in which information spreading between users is based on trust. Zarei et al. [18] suggested a novel criterion to measure the users' interest and proposed a method to obtain the most influential users based on that. Al-Azim et al. [37] introduced two models to obtain users ranking based on influence propagation in social networks via capturing interest groups and a new influence propagation model to rank users in each interest group. The purpose of the proposed UI method is to select a set of influential nodes who have common interest with their neighbors as much as possible and their activity (posts, comments, etc.) in social networks be close to the content (marketing purpose) that is going to be propagated through them.

3. Preliminary information

There are some hidden characteristics in most social networks such as Twitter, Facebook, Instagram, etc. that cannot be reached via their topological structure. This information can be very helpful in solving many problems related to the networks including influential spreader identification. As it was mentioned in the previous section, hybrid methods are trying to consider information beyond the structural properties of the networks. In the context of identifying the influential spreaders, having information about the extent of trust between users, their common interests and other parameters can be very helpful. Besides that, in most proposed method all users are considered as the candidates of being an influential spreader. Meanwhile in real social networks most of the users have small community of relation with their friends and family and have a one side connection (following) with some famous people too. So all these users can be easily ignored from the candidates list.

Based on these facts, we can categorize the users to 2 major group of 'Followers' and 'Leaders. The Leaders are the ones who are actively spending time in social networks, sending several messages every day in different categories such as their lifestyle, news, campaigns, etc. and mostly have many followers. The 'Followers' are the ones that do not make much contents, but follow others' contents. They have common interest with different 'Leaders' and are impressed with their activities. They mainly are involved in a small network of friendship and do not have many followers. Obviously the influential spreaders should be chosen from the first group and spending time on investigating the whole social network users seems useless.

Second, not all users are interested in all topics that are shared in a social network. If we investigate the following list of each user, ignoring the friendships, we get to a set of 'Leaders' who are reputed in a list of topics.

Considering the two above facts, we can have the following assumptions:

The most influential spreaders should be chosen from the 'Leaders' set.

Each leader has got a list of 'Interest' topics and is appropriate to share her specialized contents.

Based on these assumptions, in the proposed method, first we define a list of topics in different categories such as Finance and Business, Health, Politics, Entertainment, Sports, Lifestyles, Education, etc. [38]. The amount of each leader's interest in each topic is specified based on the contents that have been shared by that user and the 'Interested Topic' vector is obtained for each leader. Then, for each ordinary user (from the 'Followers' set) the 'Interest' vector is computed based on the 'Interested Topics' vector of the leaders that user is following. The difference between these two vectors determines the extent that the follower and the leader have common interest.

4. Proposed method

In the proposed approach, it is assumed that not all nodes are candidates for influential spreaders and obviously low degree nodes that are mostly in the set of follower users can be ignored at the beginning and the users with the highest degree should be assigned to the Leaders set and further computations and analysis are done using these nodes. The question is that to what extent the degree of nodes can assign them to the Leaders set. We are using the 20-80 rules to categorize the users to two sets of Leaders and Followers. As it is defined, it can be claimed that in systems that follow power-law distribution, 20% of the members benefit 80% of the resources and this rule can be extended to different applications such as website visits, Interent routers, industry income, etc. So in this context, we can consider that 20% of the users in social networks have 80% of followers.

It will be shown in Table 1 that social networks considered in this study follow power-law distribution and based on the above assumption, the size of the Leader set is assigned to be 20% of the number of users in the social network.

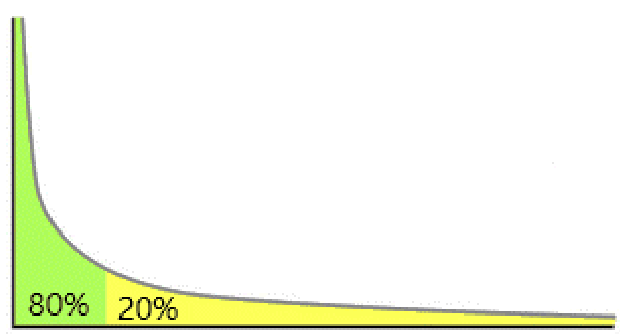

Fig. 1. 20-80 rule in networks with power-law distribution

In order to choose the right users to share the advertisements or proposed contents we need to define a list of interest topics. After that the set of interests of each leader should be defined based on her shared contents. Based on that a vector of interest is defined to summarize the extent of interest of the users in each topic. The interest vector of the users which are categorized in the Followers set is defined based on the interest vector of the leaders. In order to do that, the average of the interest vectors of the leaders who are neighbor with the ordinary user is computed and assigned to it as the interest vector. Continuing the process, the interest vector of the other users who may have no neighbor from the Leader set is computed as the average of the interest vector of her neighbors. This process continues until all users have a non-zero interest vector. Fig. **2** shows a schematic model of the proposed users' hierarchy in social networks.

In the next section, at first the problem the notations are formally stated and then the proposed method in this paper called UI is presented.

4.1. Problem formulation

In this paper, a social network is modeled as a simple directed and unweighted single-layer network denoted by $G = (V, E)$, where $V = \{v_1, v_2, \dots, v_N\}$ and $E = e_1, e_2, \dots, e_m\}$ are the set of nodes and links, respectively that $|V| = N$. Node v_i follows node v_j f there exist a directed link from v_i to v_j in the network. Γ_i is defined as the neighbors set of node i. The Leaders set is defined as U that $|U| = 0.2 \times |V|$ and the Followers set is $V - U$. The Interest topic list is $T = \{T_1, T_2, \dots, T_{|T|}\}$ and we define an interest set of size $t \leq T$ for each leader. $VI_u = \{I_{1u}, I_{2u}, \dots, I_{Tu}\}$ is the interest vector of leader u, such that t members of the vector are non-zero and the rest of them are assigned to zero (since it is supposed that each leader has limited interests) and:

$$\sum_{j=1}^{T} I_j = 1 \qquad (1)$$

The interest vector of users in the Follower set is defined the same, but the constraint on t non-zero members is not considered for them.

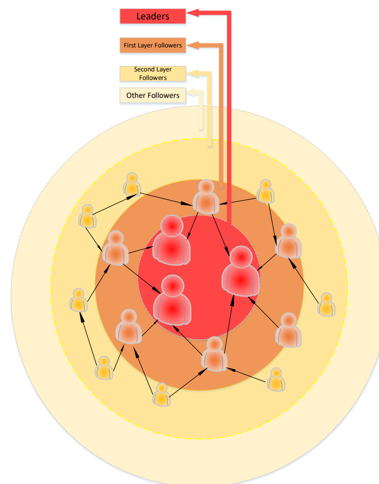

Fig. 2. Proposed users hierarchy in social network

4.2. Method

In social networks, people mainly follow leaders with similar interests to them. So, we assume that not all contents or advertisement can be assigned to all leaders and the interest of each leader in different topics has an important role in accepting the posted content via her neighbors. Thus, assuming content C relating to topic T_s in T, we find the leaders with $I_{su} > 0$ and choose the appropriate leaders from this set as the influential users in that topic:

$$\forall T_s \in T: L_s = \{u \in U | I_{su} > 0\} \qquad (2)$$

and L_s is computed for all topics in T.

In the proposed algorithm (UI), a weight W is set to each leader node to compute her similarity with her neighbors:

$$M(u) = \frac{1}{\sum_{v \in \Gamma_u} \sum_{i=1}^{|T|} |I_{iu} - I_{iv}|} \qquad (3)$$

$$W(u) = I_{su} \times M(u) \qquad (4)$$

I_{su} shows the extent of interest of the leader u in topic s and $M(u)$compute the amount of similarity of u to her neighbors.

Since we do not have access to node interests in considered social networks, the interest vectors (VI) of the leaders are generated randomly and the VI for other users is computed based on them. The following pseudo-code is presented to explain the UI method.

In lines 5-6 the 20% top degree nodes are considered as Leaders set U and the user interest vector (VI) for T topic and t specific interest for each leader is randomly generated in lines 7-14. The interest vector of other users are repeatedly generated in lines 15-27 and the similarity weight is assigned to leaders in lines 28-30. For the rest, the top users are ranked and chose in each topic and the evaluation metrics have been used to measure the performance of the algorithm.

Algorithm UI:

Input: Directed, Unweight Graph $G < V, E >$ and Topics set T
and Topic Interests size t
Output: set of Covered Nodes CN and interested Covered Nodes ICN
Begin Algorithm
Set D as the Node degree vector and sort it descending
Set U as the top 20% of D as the influential users candidates
For u ∈ U
%find t integer random number in range [1,T]:
r1=Rand_Integer(T,t)
% find t random number in range [0,1]:
r2=Rand(1,t)
R(r1)=r2/sum(r2)
VI(u)=R
End For
%First Layer:
For $i \in V - U$
Set $C\{i\}$ as the common members of Γ_i and U
Set $VI(i)$ as the average of $VI(C\{i\})$
Add i to V'
End For
%Second Layer and Others:
While $\exists i: VI(i) = 0$
For $i \in V - \{U \cup V'\}$
Set $C\{i\}$ as the common members of Γ_i and V'
Set $VI(i)$ as the average of $VI(C\{i\})$
End For
End While
For $u \in U$ & $T_s \in T$
$W(u) = I_{su} \times \frac{1}{\sum_{v \in \Gamma_u} |VI(u) - VI(v)|}$
End For
For toc= 1 to $|S|$
For $k = 1$ to $|T|$
Find the top most users $u \in U$ that $I_{su} > 0$
%Use SIR model with p=0.1 to compute the approximated Cover Set
and Interested Cover Set $ICS\{k\}$
End For
CN=Average(CS)
ICN=Average(ICS)
End For
Return CN, ICN
End Algorithm

5. Experimental Results and analysis

5.1. Datasets

In order to validate the performance of the proposed method, we perform experiments on six real social networks. In the following we provide explanation of these networks.

1. **Advogato**: This is a directed network of trust relationships bbetween users on Advogato which is an online community of open source software developers.
2. **Trust**: This is a dataset which was collected in a 5-week crawl in 2003 from the Epinions.com web site.
3. **BrightKite**: This is a directed, location-based social service provider where users shared their locations and the friendship network was collected using their public API.
4. **Epinion**: This is a directed network presenting the who-trust-whom relationship in online social network of the site Epinions.com.
5. **Douban**: This is the social network of a Chinese online recommendation site.

Gowalla: This is another location-based social network where users share their locations and the friendship network was collected using their public API.

Table 1. Information on the six networks used in the experiments. <k> represents the average degree and λ is the power-law coefficient.

Network	N	E	<k>	λ
Advogato	6541	51,127	15.63	3.0410
Trust	49288	487183	19.77	3.4901
BrightKite	58,228	214,078	7.35	2.4810
Epinion	75,879	508,837	13.41	2.0258
Douban	154,908	327,162	4.22	2.0810
Gowalla	196,591	950,327	9.67	2.6510

5.2. Evaluation Metrics

The performance of the existing and proposed method are compared in terms of a number of evaluation metrics. One of the evaluation metrics is the percentage of covered nodes (users) respect to changes of size of S, which is defined as below:

$$CN = \frac{|AV|}{|V|} \tag{5}$$

Which $|AV|$ is sum of the activated node via SIR model and $|V|$ is the total number of nodes.

Another evaluation metric is used to measure the percentage of interested user in the specific topic that are activated respect to changes of size of S, which is defined as below:

$$ICN = \frac{|AIV|}{|IV|} \tag{6}$$

Which $|AIV|$ is sum of the activated users via SIR model who are interested in the considered topic and $|IV|$ is the total number of users who are interested in the considered topic.

5.3. Spreading model

The susceptible–infected–recovered (SIR) spreading model is employed [39] as the influence analysis model. At each time step, each node can be in one of the three possible states:

Susceptible (S): The node is vulnerable to become infected.

Infectious (I): The node is infected and tries to infect its susceptible neighbors.

Recovered (R): The node has recovered and cannot become infected anymore or infect others.

In a network if node v is infected it can infect her neighbors with a certain probability [40]. At the next time step the infected node v becomes recovered and would not be able to infect the others any more. When the algorithm identifies the influential spreaders, they all are set to be infected and all other nodes are set to be susceptible, then the spreading process start by infecting the susceptible nodes with probability of λ and they become recovered with probability of γ. This process ends when there would be no more infected node in the network. A schematic view of the process of the SIR model is shown in Fig. **3** and the differential equations are as below:

$$\frac{dS}{dt} = -\lambda SI \qquad (7)$$

$$\frac{dS}{dt} = -\lambda SI - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

In this study we assume $\gamma = 1$ and $\lambda = 0.1$.

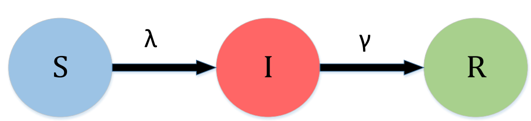

Fig. 3. A schematic view of the SIR model

5.4. Results and analysis

In order to validate the performance of the proposed method UI, it is compared with a number of existing methods including degree centrality (D), k-shell centrality (KS), Neighbors' k-shell centrality (NKS), Neighbors Degree centrality (N), Overlay strategy (O) and Nodes and Neighbors Degree centrality (KN). In the UI method, the value of covered nodes (CN) and interested covered nodes (ICN) is computed for each topic and different fraction of influential spreaders $F_s = \frac{|S|}{|V|}$, and the average is considered for each network over 20 independent experiments. The SIR spreading model with the parameters $\gamma = 1$ and $\lambda = 0.1$ is considered.

Fig. **4** shows the average number of covered nodes (CN) based on different fraction of seed set F_s. As it can be seen, none of the methods has the best performance in all social networks and the behavior of them is quite divergent and dependent to the structure of the networks. But in most cases the main compatitors are UI method and Overlay Strategy (O).

It is worth noting that a constraint has been set to identify the influential nodes based on UI methods. Since we believe that a content (post, advertisement, etc.) that is related to a topic should be passed to a user who has a common interest with that topic, So not all users are appropriate to be chosen as influential nodes. This is the shortcoming of most structural-based methods that do not consider extra information about the nodes and their functionality. Based on this assumption, in this method the influential nodes are selected separately based on the topics, i.e. the members of the candidate list should all be interested in the proposed topic. This constraint is considered for all topics and the final result is the average of covered nodes examining all topics. But there is no constraint for the other compared methods. Our experiments shows that employing this constraint for other compared nodes results in decreasing their performance and the UI method would outperform all the other methods in case of evaluation of covered nodes.

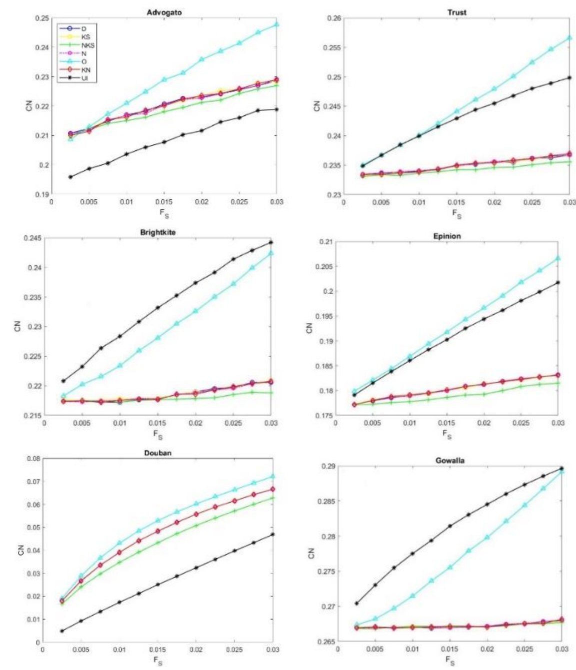

Fig. 4. Average number of covered nodes (CN) based on fraction of influential nodes F_s using different centrality measures in real social networks.

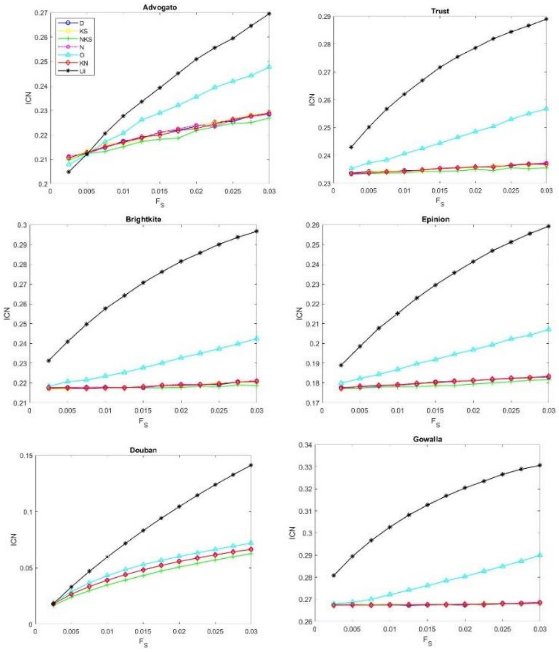

Fig. 5. Average number of interested covered nodes (ICN) based on fraction of influential nodes F_S using different centrality measures in real social networks.

Fig. **5** shows the average number of interested covered nodes (ICN) based on different fraction of seed set F_S. As it can be seen, the proposed method UI outperform the other compared methods and the number of covered nodes who are interested in the shared content, grows significantly.

Besides that, a constraint can be defined for the number of topics that each user can decide to be a member of the corresponding list. The users may be authorized to select a limited list of interest topics which are quite expert in them or plenty of different interest topics which are not that much professional in them. Obviously the more limited the number of the selected topics, the more specialist the users are in that topic. In this section, the impact of parameters T and t is examined. T is the list of the interests topics and t is the number of topics that each candidates in U is interested in it. The impact of variety of parameter T and limitation of t is examined using the proposed UI method. The SIR spreading model with the same parameters $\gamma = 1$ and $\lambda = 0.1$ is considered. For each network, the results show the mean value of interested covered nodes over 20 independent experiments.

Firstly, the amount of T is changed in the range of [10,20,30,40] and the number of selected topics t is fixed to 4. As it is shown in Fig. **6**, increasing the variety of the topics, results in increasing the percentage of interested nodes covered by the proposed UI method. So, specifying the topics with more details and categorizing them to more groups improves the performance of the method.

In the next phase, the amount of T is fixed to 20 and the number of topics t that can be selected by each candidate in U is changed in the range of [4,6,8,10]. As it is shown in Fig. **7**, increasing the variety of the topics, results in increasing the percentage of interested nodes covered by the proposed UI method. So, specifying the topics with more details and categorizing them to more groups improves the performance of the method.

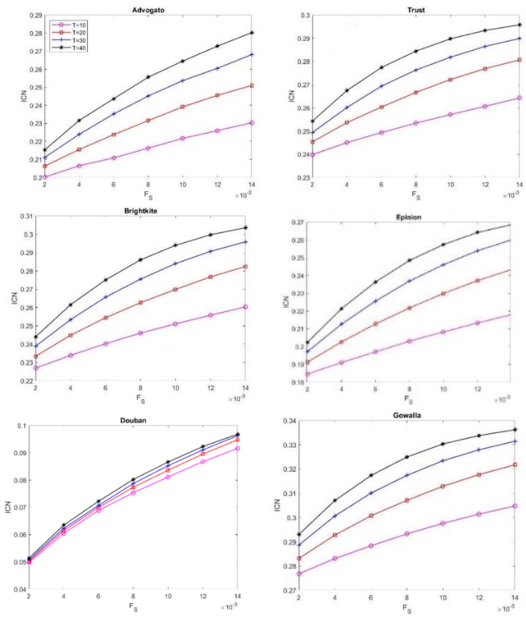

Fig. 6. Impact of number of topics T on average number of interested covered nodes based on fraction of influential nodes F_S using UI method in real social networks.

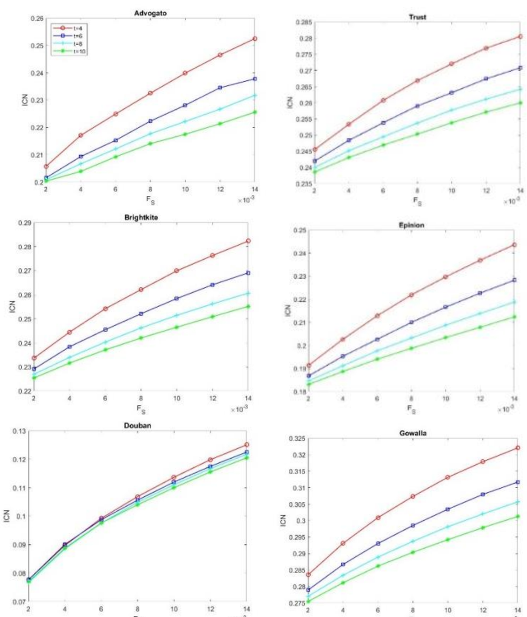

Fig. 7. Impact of number of selected topics t on average number of interested covered nodes based on fraction of influential nodes F_S using UI method in real social networks

6. Conclusion

Identification of the most influential spreaders in social network have an important role in the result of viral marketing strategies. In many of the existing structural-based methods, all users in social network are regarded interested in any kind of marketing contents belonging to different categories, whereas this is not true in real world. In this paper, a method was proposed to identify the most influential users on social networks based on their interests. We claimed that not all users can be the initial candidates and ignored 80% of the network's users. A list of interest topics was assigned to the candidates and their common interests with their neighbors was computed and regarded as their influence power. Our experiments on six real social networks revealed effectiveness of the proposed method in covering the interested users in all cases and ordinary users in most cases.

References

[1] Chen, W., Wang, C. & Wang. Y. (2010). 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.

[2] Cheung, M.Y., et al., (2009). Credibility of electronic word-of-mouth: Informational and normative determinants of on-line consumer recommendations. International journal of electronic commerce, 13(4), 9-38.

[3] Bond, R.M., et al., (2012). A 61-million-person experiment in social influence and political mobilization. Nature, 489 (7415), 295-298.

[4] Bae, J., & Kim, S. (2014). Identifying and ranking influential spreaders in complex networks by neighborhood coreness. Physica A: Statistical Mechanics and its Applications, 395: p. 549-559.

[5] Wang, Z., et al. (2017). Ranking influential nodes in social networks based on node position and neighborhood. Neurocomputing, 260, 466-477.

[6] Sabidussi, G. (1966). The centrality index of a graph. Psychometrika, 31(4), 581-603.

[7] Freeman, L.C. (1977). A set of measures of centrality based on betweenness. Sociometry, 35-41.

[8] Kitsak, M., et al., (2010). Identification of influential spreaders in complex networks. Nature physics, 6(11), 888-893.

[9] Zareie, A., & Sheikhahmadi, A. (2018). A hierarchical approach for influential node ranking in complex social networks. Expert Systems with Applications, 93, 200-211.

[10] Zareie, A., Sheikhahmadi, A. & Jalili, M. (2019). Influential node ranking in social networks based on neighborhood diversity. Future Generation Computer Systems, 94, 120-129.

[11] Wasserman, S., & Faust, K. (1994). Social network analysis: Methods and applications. Vol. 8, Cambridge university press.

[12] Henri, F., & Pudelko, B. (2003). Understanding and analysing activity and learning in virtual communities. Journal of Computer Assisted Learning, 19(4), 474-487.

[13] Chen, W., Wang, Y., & Yang, S. (2009). Efficient influence maximization in social networks. In Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining.

[14] Guo, L., et al., (2016). Identifying multiple influential spreaders in term of the distance-based coloring. Physics Letters A, 380(7-8), 837-842.

[15] Wang, X., et al., (2016).Effective identification of multiple influential spreaders by DegreePunishment. Physica A: Statistical Mechanics and its Applications, 461, 238-247.

[16] Mohamadi-Baghmolaei, R., Mozafari, N., & Hamzeh, A. (2015).Trust based latency aware influence maximization in social networks. Engineering Applications of Artificial Intelligence, 41, 195-206.

[17] Nguyen, H.T., Dinh, T.N., & Thai. M.T. (2016). Cost-aware targeted viral marketing in billion-scale networks. in IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications. IEEE.

[18] Zareie, A., Sheikhahmadi, A., & Jalili, M. (2019). Identification of influential users in social networks based on users' interest. Information Sciences, 493, 217-231.

[19] Chevalier, J.A., & Mayzlin, D. (2006).The effect of word of mouth on sales: Online book reviews. Journal of marketing research, 43(3), 345-354.

[20] Liu, S., et al., (2015). Identifying effective influencers based on trust for electronic word-of-mouth marketing: A domain-aware approach. Information sciences, 306, 34-52.

[21] Zareie, A., Sheikhahmadi, A. & Fatemi, A. (2017). Influential nodes ranking in complex networks: An entropy-based approach. Chaos, Solitons & Fractals, 104, 485-494.

[22] Bao, Z.-K., Liu, J.-G. & Zhang, H.-F. (2017). Identifying multiple influential spreaders by a heuristic clustering algorithm. Physics Letters A, 381(11), 976-983.

[23] Freeman, L.C., Centrality in social networks conceptual clarification. Social networks, 1978. 1(3), 215-239.

[24] Mochalova, A. & Nanopoulos, A. (2014). A targeted approach to viral marketing. Electronic Commerce Research and Applications, 13(4), 283-294.

[25] Yang, Y., et al., (2018). Exploring influence maximization in online and offline double-layer propagation scheme. Information Sciences, 450, 182-199.

[26] Cha, M., et al. (2010). Measuring user influence in twitter: The million follower fallacy. in fourth international AAAI conference on weblogs and social media.

[27] Dave, K., Bhatt, R., & Varma, V. (2011). Identifying influencers in social networks. in Proceedings of the 5th International Conference on Weblogs and Social Media.

[28] Pei, S., et al., (2014). Searching for superspreaders of information in real-world social media. Scientific reports, 4, 5547.

[29] Al-garadi, M.A., Varathan, K.D., & Ravana, S.D. (2017). Identification of influential spreaders in online social networks using interaction weighted K-core decomposition method. Physica A: Statistical Mechanics and its Applications, 468, 278-288.

[30] Berahmand, K., Bouyer, A., & Samadi, N. (2019). A new local and multidimensional ranking measure to detect spreaders in social networks. Computing, 101(11): p. 1711-1733.

[31] Berahmand, K., Bouyer, A., & Samadi, N. (2018). A new centrality measure based on the negative and positive effects of clustering coefficient for identifying influential spreaders in complex networks. Chaos, Solitons & Fractals, 110, 41-54.

[32] Berahmand, K., Samadi, N., & Sheikholeslami, S.M. (2018). Effect of rich-club on diffusion in complex networks. International Journal of Modern Physics B, 32(12), 1850142.
[33] Zareie, A., et al., (2020). Finding influential nodes in social networks based on neighborhood correlation coefficient. Knowledge-Based Systems, 105580.
[34] Abel, F., et al. (2011). Semantic enrichment of twitter posts for user profile construction on the social web. in Extended semantic web conference. Springer.
[35] Zhu, Z., (2013). Discovering the influential users oriented to viral marketing based on online social networks. Physica A: Statistical Mechanics and its Applications, 392(16), 3459-3469.
[36] Wu, H., Arenas, A., & Gómez, S. (2017). Influence of trust in the spreading of information. Physical Review E, 95(1), 012301.
[37] Abd Al-Azim, N.A.R., et al., (2020). Influence propagation: Interest groups and node ranking models. Physica A: Statistical Mechanics and its Applications, 124247.
[38] Tao, K., et al. (2011). Tums: twitter-based user modeling service. in Extended Semantic Web Conference. Springer.
[39] Hethcote, H.W. (2000). The mathematics of infectious diseases. SIAM review, 42(4), 599-653.
[40] Guan-Rong, C., Xiao-Fan, W., & Xiang, L. (2012). Introduction to Complex Networks: Models, Structures and Dynamics.