Journal of Theoretical and Applied Information Technology
10th March 2015. Vol.73 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

RANKING THE INFLUENCE USERS IN A SOCIAL NETWORKING SITE USING AN IMPROVED TOPSIS METHOD

[1]A. MURUGANANTHAM, [2]DR. G. MEERA GANDHI

[1]Research Scholar, Sathyabama University, Rajiv Gandhi Road, Chennai-600 119, Tamil Nadu, India.

[2]Professor, Department of Computer Science and Engineering, Sathyabama University, Rajiv Gandhi Road, Chennai-600 119, Tamil Nadu, India.

E-mail: [1]amuruganantham@gmail.com , [2]drmeeragandhii@gmail.com

ABSTRACT

With the wide spread popularity of social networking sites (SNS), enterprise organizations have started to explore the business opportunities in SNS such as Facebook to conduct targeted marketing and reputation management. Customers or users tend to trust the opinion of other customers or users, especially those with prior experience of a product or service, rather than company marketing. One of the important challenges to these enterprises is to conduct cost-effective marketing and reputation management on SNS through influencing users. When it comes to marketing, the users' influence is associated with a certain topic or field on which people have different levels of preference and expertise is called homophily. In order to identify and predict influential users in a specific topic/subject more effectively, this paper introduces a new method to effectively identify the most influence users, who can generate the maximum of total benefit with respect to specific topics or business situations. This method uses homophily characteristics along with the technique for order preference by similarity to ideal solution (TOPSIS). Using this improved TOPSIS method, influencing users are identified and ranked on a Facebook dataset and compared against with TOPSIS method with no homophily. The experimental results show that how well the proposed technique precociously identify and rank influential users based on a certain topic or business situation.

Keywords: *Influence Users, Betweenness Centrality, Closeness Centrality, Eigenvector Centrality, Homophily, TOPSIS*

1. INTRODUCITON

Social networking sites (SNS) become very popular in recent years because of the increasing proliferation and affordability of internet enabled devices such as mobile devices and tablets. Social network site like Facebook is a huge virtual space where to express and share individual opinions, emotions, influencing one's view or opinions. Many such social networks are extremely rich in content, and they typically contain a tremendous amount of content and linkage data which can be leveraged for analysis.

Reviews and quality comments by important members, who is part of one's social network like Facebook plays significant information for others while they do evaluation of products and services offered by business organizations. In some business sector like e-retailer, it is even becoming a fundamental variable in the "purchase" decision. A recent Forrester study

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

showed that more than 30% of Internet users have evaluated products or services online.

Consumers tend to trust the opinion of other consumers, especially those with prior experience of a product or service, rather than company marketing. An important customer's influence is called social influence influences other customers' preferences by shaping their attitudes and behaviours. Monitoring, identifying and engaging the Influence Users who are most relevant to the brand, product or campaign is become important now. In this way, business enterprises could retain efforts aimed at sustaining the activity of influential users, who take minimal effort and resources to improve product sales and enhance their reputations to improve the business enterprise.

Structural Network Metrics such as Betweenness centrality and Closeness centrality can better identify influential nodes, but they are incapable preciously identifying the influence users as the number of nodes increases, the graph starts to become too complicated to identify influence users. A new method is presented with a case-study to understand and identify influence users in a Facebook account using homophily characteristics along with the technique for order preference by similarity to ideal solution (TOPSIS).

The remaining parts of the paper are organized as follows. The first section describes previous literature of related research. The second section describes various influence related Structural Network Metrics and TOPSIS method whereas third section explain the proposed method and the fourth section presents how to identify influence users using homophily characteristics along with the technique for order preference by similarity to ideal solution (TOPSIS) in a Facebook account dataset.

2. RELATED WORKS

In the literature, there are a number of works related to influential user identification in social networking sites.

Models were proposed to compute social influence probabilities from real social network data, which uses action logs of web data, and next define propagation of actions and propagation graph [1]. The algorithms first learn model parameters, and then test the learned models to make predictions. In the process, influence probabilities were calculated based on actions in action logs. Thus each action can appear completely either in training or test dataset. Another technique [2] proposed by King and Chan summarize tasks and techniques in social computing mainly include but not limited to: social network theory, modelling, and analysis; ranking; query log processing; web spam detection; graph/link analysis and mining; collaborative filtering; sentiment analysis and opinion mining. Bharathi & Tang proposed methods [3][4] to focus on influence maximization problem, and Pang [5] and Turney [6] found polarity detection from web text, but few of them attempted to analyze deeply and find how a user sentimentally influences or is influenced by another in social networks. TwitterRank was proposed by Weng [7] to identify influential users in Twitter. As an extension of PageRank algorithm, it measures the influence taking both the topical similarity between users and the link structure into account. They truly process the tweets published in Twitter, and present their results to validate their solution on influence maximization problem. Duanbing Chen[8] used the Susceptible–Infected–Recovered (SIR) model to examine the spreading influence of the nodes ranked by different centrality measures. Kaiquan [9] identified influencers using joint influence powers through Influence network, which took long time to build. Zhigu [10] uses user trust network to identify influence users took long time to build trust list, which is incomplete. Qian [11] proposed weighted LeaderRank technique by allowing users with more fans get more scores from the ground node that is, replacing the standard random walk by a biased random walk. Tang [12] proposed a new approach to incorporate users' reply relationship,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

conversation content and response immediacy which capture both explicit and implicit interaction between users to identify influential users of online healthcare community. Yuxian [13] used multi-attribute in TOPSIS method to identify influential nodes in complex networks. Although the previous research has examined the problem of discovering a group of influential users, it did not quickly identify influence users using minimum computing power and it was not dynamic to the situation.

In contrast to these earlier studies, a new method is proposed to identify influence users using homophily characteristics along with the technique for order preference by similarity to ideal solution (TOPSIS).

2.1 A Directed Graph

Formally, assume a social network is modelled as directed graph G (V:E) where nodes V represent users, edges E represent social relationships among users and N represent size of network. Suppose user x adopts an innovation at time t1. We say that user x influences user y if and only if at time t2 when user y adopts the innovation, user x has already adopted it at an earlier time t1, at which time x and y were already friends. We therefore assume that social influence occurs when the information of a friend adopting the innovation has the influence to flow to neighbouring nodes in the social network.

2.2 SOCIAL INFLUENCE

Social influence refers to the behavioural change of individuals affected by others in a network. Social influence is an intuitive and well-accepted phenomenon in social networks [14]. The strength of social influence depends on many factors such as the strength of relationships between people in the networks, the network distance between users, temporal effects, characteristics of networks and individuals in the network. ‘

Standard Network Graph Metrics such as centrality closeness, eigenvector closeness and Betweenness closeness are related to social influence in terms of the structural effects of different edges and nodes.

2.3 Edge Measures

Edge measures relate the social influence based concepts and measures on a pair of nodes.

Tie strength: According to Granovetter's seminal work [15], the tie strength between two nodes depends on the overlap of their neighborhoods. In particular, the more common neighbors that a pair of nodes A and B may have, the stronger the tie between them. If the overlap of neighborhoods between A and B is large, we consider A and B to have a strong tie. Otherwise, they are considered to have a weak tie. It is formally defining the strength S(A,B) in terms of their Jaccard coefficient.

$$S(A,B) = \frac{|nA \cap nB\ |}{|nA \cup nB|} \qquad (1)$$

Equation (1) describes the strength of Tie, where nA and nB indicate the neighborhoods of A and B, respectively. Alternatively, it is called embeddedness. The embeddedness of an edge is high if two nodes incident on the edge have a high overlap of neighborhoods. Mark stated that when two individuals are connected by an embedded edge, it makes it easier for them to trust one another, because it is easier to find out dishonest behaviour through mutual friends [16].

Edge Betweenness: It measures the total amount of flow across the edge. The information flow between A and B are evenly distributed on the shortest paths between A and B. edge Betweenness leads to graph partitioning. By removing the edges of high Betweenness scores to turn the network into a hierarchy of disconnected components or clusters of nodes. More detailed studies on clustering methods [17] are presented in the work by Girvan and Newman.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2.4 Node Measures

Node-based centrality measure the importance of a node in the network. Centrality has attracted a lot of attention as a tool for studying social networks. A node with high centrality score is usually considered more highly influential than other nodes in the network [18]. Many centrality measures have been proposed based on the precise definition of influence. Centrality measures can be classified into: radial and medial measures. Radial measures assess random walks that start or end from a given node whereas the medial measures assess random walks that pass through a given node. The radial measures are further categorized into volume measures and length measures based on the type of random walks. Volume measures fix the length of walks and find the volume (or number) of the walks limited by the length. Length measures fix the volume of the target nodes and find the length of walks to reach the target volume.

Degree: It is radial and volume-based measure. The simplest and most popular measure is degree centrality.

Let A be the adjacency matrix of a network, and deg(i) be the degree of node i. The degree centrality C_i^{DEG} of node i is defined to be the degree of node:

$$C_i^{DEG} = \deg(i) \qquad (2)$$

Closeness: It is radial and length based measures. Unlike the volume based measures, the length based measures count the length of the walks. The most popular centrality measure in this group is closeness centrality [19]. It measures the centrality by computing the average of the shortest distances to all other nodes. The closeness centrality C_i^{CLO} of node i is defined as follows:

$$C_i^{CLO} = e_i^T S1 \qquad (3)$$

Here S be the matrix whose (i, j)th element contains the length of the shortest path from node i to node j and 1 is the all one vector.

Node Betweenness or Betweenness Centrality: nodes of high Betweenness occupy critical positions in the network structure, and are therefore able to play critical roles. It is often enabled by a large amount of flow, which is carried by nodes which occupy a position at the interface of tightly-knit groups. Such nodes are considered to have high Betweenness. The Betweenness centrality C_i^{BET} of node i is defined as follows:

$$C_i^{BET} = \sum_{j,k} \frac{b_{ijk}}{b_{jk}} \qquad (4)$$

Here b_{ijk} is the number of shortest paths from node j to k, and b_{jk} be the number of shortest paths from node j to k that pass through node i.

Eigenvector Centrality: It is defined as a function of number and strength of connections to its neighbors and as well as those neighbors' centralities. Let x(i) be the Eigenvector centrality of a node vi. Then,

$$x(i) = \frac{1}{\lambda} \sum_{j=1}^{N} A_{i,j}\, x(j) \qquad (5)$$

Here λ is a constant and A denotes the adjacency matrix. In nutshell, The Eigenvector centrality network metric takes into consideration not only how many connections a vertex has (i.e., its degree), but also the degree of the vertices that it is connected to. The Structural Network Metrics of a social networking site such degree, closeness centrality, Betweenness centrality and eigenvector centrality are calculated based on the equations (2), (3), (4) and (5) respectively.

2.5 TOPSIS Method

TOPSIS (technique for order preference by similarity to an ideal solution) method is presented in Chen and Hwang [20], with reference to Hwang and Yoon [21]. TOPSIS is a multiple criteria method to identify solutions from a finite set of alternatives. The basic principle is that the chosen alternative should

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

have the shortest distance from the positive ideal solution and the farthest distance from the negative ideal solution. The procedure of TOPSIS can be expressed in a series of steps:

(1) Calculate the normalized decision matrix. The normalized value n_{ij} is calculated as

$$n_{ij}= x_{ij} / \sqrt{\sum_{j=1}^{m} x^2_{ij}}$$

(2) Calculate the weighted normalized decision matrix. The weighted normalized value v_{ij} is calculated as

$$v_{ij}=w_i n_{ij},\ j=1, \ldots,m,\ i=1, \ldots,n,$$

where w_j is the weight of the i^{th} attribute or criterion, and $\sum_{j=1}^{m} w_j =1$

(3) Determine the positive ideal and negative ideal solution.

$$A^+ = \{ v^+_1, \ldots v^+_n\}=\{(\max_j v_{ij} \mid i\ \varepsilon\ I),$$

$$(\min_j v_{ij} \mid i\ \varepsilon\ J) \},$$

$$A^- = \{ v^-_1, \ldots v^-_n\} =\{ (\min_j v_{ij} \mid i\ \varepsilon\ I),$$

$$(\max_j v_{ij} \mid i\ \varepsilon\ J) \},$$

Where I is associated with benefit criteria, and J is associated with cost criteria.

(4) Calculate the separation measures, using the n-dimensional Euclidean distance. The separation of each alternative from the ideal solution is given as

$$d^+_j = \{\sum_{i=1}^{n} (v_{ij} - v^+_i)^2\}^{1/2},\ j=1,\ldots,m.$$

Similarly, the separation from the negative ideal solution is given as

$$d^-_j = \{\sum_{i=1}^{n} (v_{ij} - v^-_i)^2\}^{1/2},\ j=1,\ldots,m.$$

(5) Calculate the relative closeness to the ideal solution. The relative closeness of the alternative A_j with respect to A^+ is defined as

$$R_j= d^-_j / (d^+_j + d^-_j),\ j=1,\ldots,m.$$

Since $d^-_j \geq 0$ and $d^+_j \geq 0$, then, clearly, $R_j\ \varepsilon\ [0, 1]$.

(6) Rank the preference order. For ranking Decision Making Units (DMUs) using this index, we can rank DMUs in decreasing order.

The basic principle of the TOPSIS method is that the chosen alternative should have the ''shortest distance'' from the positive ideal solution and the ''farthest distance'' from the negative ideal solution. The TOPSIS method introduces two ''reference'' points, but it does not consider the relative importance of the distances from these points.

3. PROPOSED METHOD

Although many works proposed multiple techniques of generating social network and estimating influence probabilities, they did not directly deal with the problem of identifying influential users from a social network. To address the problem of identifying influential users, it is important to understand the properties of the social networking site and develop the mechanism to extract the social network structure and develop formulations to compute the influence possibilities in the dynamic business situation. A new method of identifying influential users in social network is based on homophily characteristic of users. The TOPSIS method is applied to users selected based on homophily characteristics along with centrality measures rather than it is applied to entire users, who are identified only through centrality measures. The tendency of individuals to form association with individuals of similar socio-cultural background becomes the basic governing structural component of any social network and it has been the focus of many social

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

network studies as described by HalilBisgin [22]. A link between individuals (such as colleague) is correlated with those individuals being similar in nature. For example, friends often tend to be similar in characteristics like age, working department, social background and education level. Our proposed method rank the user influence by combing the content-based (Homphily characteristics) and structural network-based metrics such as Betweenness Centralities, Closeness Centrality and Eigen vector centralities etc. Besides identifying the top influencers using general centrality metrics like Betweenness centrality (BC), Closeness Centrality (CC) and eigenvector centrality (EC) are further classified using value of homophily attribute, which is appropriate to a specific business situations. For example, users from same region/town are interested to a specific topic. Among those users, the influencers are identified from their structural properties using TOPSIS method. The users with same or similar homophily are highly influence users than others in the social network.

The flow chart of the proposed method is shown in Figure 1. The specific steps of the method are illustrated as the following:

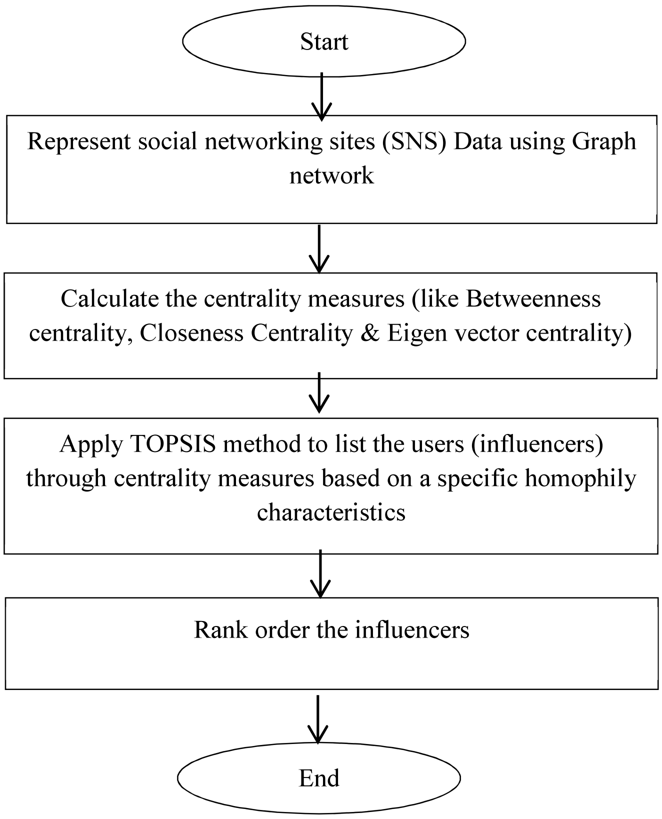

Figure 1: The Flow Chart Of Propose Method

Step 1 Construct network.

Social network sites (SNS) data is constructed as network using any conventional tool like NodeXL. Users with relationships are represented as undirected graph as shown in Figure 2.

Step 2 Calculate the different centrality values.

Degree centrality (DC), closeness centrality (CC) and Betweenness Centrality (BC) and eigenvector centrality (EC) are calculated as mentioned in the equations (2), (3), (4) and (5).

Step 3 Identify influencers using TOPSIS method on users selected on a specific homophily characteristics

Based on the centrality values of social network data, the top influencers are listed from the users, who are having similar homophily values.

Step 4 Rank the users with same homophily characteristic

Rank order the users based on the TOPSIS score. For example, the users with homophily attribute like Current Organization, relationship, home town, working department etc. are ordered based on the TOPSIS score.

Journal of Theoretical and Applied Information Technology
10th March 2015. Vol.73 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

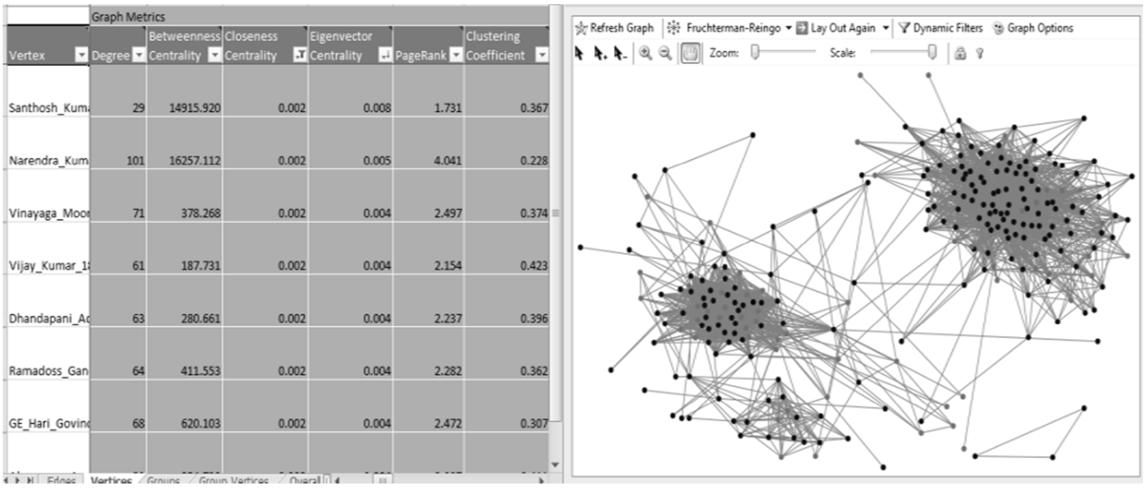

Figure 2: Network Structure Of Facebook Data In Nodexl

4. EXPERIMENTAL STUDY

In this section, we present the details of experiments on real datasets to evaluate the performance of our approach. We applied our approach to a Facebook data set and identified influencers.

4.1 A Facebook Dataset

A Facebook network dataset was extracted from our Facebook account. The extracted Facebook data was represented through NodeXL as network graph structures and found 260+ users connections. .

There are other network analysis tools like Pajek, UCINet, and NetDraw that provide graphical interfaces, rich libraries of metrics, and do not require coding or command line execution of features. However, we find that these tools are designed by Borgatti [23] and Shneiderman [24], have complex data handling, and inflexible graphing and visualization features that inhibit wider adoption.

4.2 Centrality Measures Calculation

The NodeXL represents the imported dataset in the form of edge lists, i.e., pairs of vertices which are also referred to as nodes. Each vertex is a representation of an entity in the network. Each edge, or link, connecting two vertices is a representation of a relationship that exists between them. This relationship may be directed or not. Some relationships are bidirectional (like marriage); others can be uni-directional (like lending pen). The right-hand side of Figure 3 displays the imported Facebook dataset network structure.

Journal of Theoretical and Applied Information Technology
10th March 2015. Vol.73 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

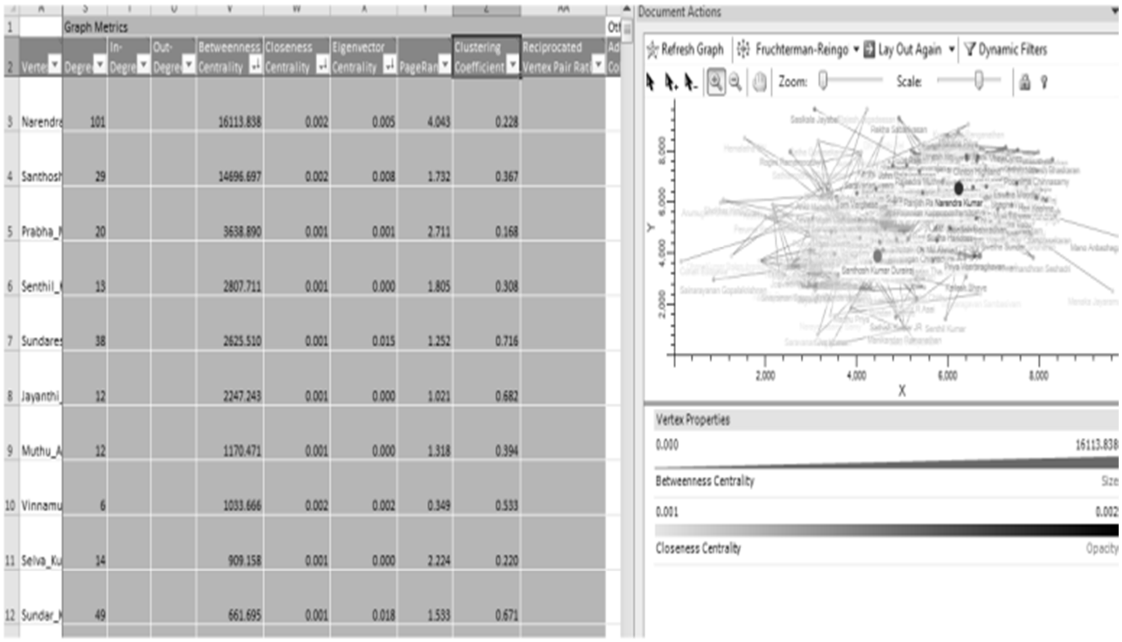

Figure 3: Imported Facebook Dataset In Nodexl

NodeXL has a number of software routines for calculating Centrality measures such as degree, in-degree, out-degree, clustering coefficient, and closeness, Betweenness, and eigenvector centrality etc. The computed centrality measures of Facebook dataset are also shown in the left-hand side of Figure 3.

4.3 Identification of Influence Users

From the Facebook dataset in NodeXL, the users are extracted based on the specific homophily characteristics. For example, "Unmarried" users are listed with various centralities measures. From the "Unmarried" users list, the influencers are identified using TOPSIS method. The below shown figures 5, 6 and 7 rank the influencers identified through homophily attributes such as Married, home town, and sex. In this approach, business organizations understand preciously the influence users for a specific business situation or context rather than ranking the influencers based only centrality measures as shown in figure 4. It helps to target specific marketing campaign on the influence users, who are identified on specific business situation/topic. A marketing campaign specific to a city say "Chennai" can be targeted to only users who are identified from the homophily characteristics-hometown="Chennai". Figure 8 & 9 shows that how the top ranked influence users are listed with TOPSIS with homophily against TOPSIS without homophily characteristics. Homophily with TOPSIS based classification ranks and list the users based on the specific interest or topic. Depending upon the specific business situation/subject, the influencers users identified based on the business specific homophily characteristics.

4. CONCLUSIONS

In this paper, a new method is proposed to identify the influence users in the social network sites (SNS) like Facebook using an improved TOPSIS method. Instead of applying the TOPSIS method directly on the users, who are identified through various centrality measures, the users are listed based on TOPSIS method along with homophily characteristics. Due to this improvement, the potential influencers are identified and ranked preciously based on the specific topic or interest. For example, a business organization understands the influencers (as

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

shown in Figure 6 or 9) in a specific city/town and engages those users with specific marketing campaign. Our approach has many potential applications in the context of understanding influence users. The influence user identified by our approach is meaningful because they are topic/business situation specific.

The current work has still few limitations and can be improved in the future.

(1) The centrality (structural) network metrics are taken into account while it is not considered the dynamic properties, which are evolved over the time period.

(2) The current work is limited to social network site-Facebook and it can be extended other online social network site like Blogs, E-mail, Twitter, Myspace etc.

				Options														
				1	2	3	4	5	6	7		171	172	173	174			
			Importance	Santhosh	Narendra	Sundaresan	Prabha	Ajeesh	Mahendra	Anubhav	...	Vidhya	Vkey	Karthikey	Prakash	Goal	- Ideal	+ Ideal
Criteria:	1	Betweenness Centrality	1	14915.9	16257.1	2620.258	3783.14	664.51	427.572	411.892	...	9.066	7.028	4.396	4.235	maximize	3.214	16257.1
	2	Closeness Centrality	1	0.002	0.002	0.001	0.001	0.001	0.001	0.001	...	0.001	0.001	0.001	0.001	maximize	0.001	0.002
	3	Eigenvector Centrality	1	0.008	0.005	0.015	0.001	0.02	0.019	0.019	...	0.001	0.001	0.001	0.001	maximize	0	0.02
			Score	0.84	0.83	0.23	0.22	0.22	0.21	0.21	...	0.01	0.01	0.01	0.01			

Figure 4: TOPSIS Computation With No Homophily Characteristics

				Options:											
				1	2	3	4	5	...	80	81	82			
			Importance	Narendra	Ajeesh	Anubhav	Pushpalatha	Senthil	...	Vaishnavi	Tom	Karthikey	Goal	-Ideal	+Ideal
Criteria:	1	Betweenness Centrality	1	16257.1	664.509	411.892	367.344	2984.56	...	3.134	2.695	4.396	maximize	0.194	16257.1
	2	Closeness Centrality	1	0.002	0.001	0.001	0.001	0.001	...	0.001	0.001	0.001	maximize	0.001	0.002
	3	Eigenvector Centrality	1	0.005	0.02	0.019	0.018	0	...	0.001	0.001	0.001	maximize	0	0.02
			Score	0.84	0.21	0.2	0.19	0.18	...	0.01	0.01	0.01			
							Homophily: Relationship Status=Married								

Figure 5: TOPSIS Computation Based On Relationship Status

				Options:													
				1	2	3	4	5	6	7	...	33	34	35			
			Importance	Narendra	GE	Dhandapa	Vijay	Gautham	JaganBabu	Surendar	...	Kala	Swetha	Ramesh	Goal	- Ideal	+ Ideal
Criteria:	1	Betweenness Centrality	1	16257.1	620.103	280.661	187.73	207.411	413.61	223.568	...	1.816	213.049	140.726	maximize	0.287	16257.1
	2	Closeness Centrality	1	0.002	0.002	0.002	0.002	0.001	0.001	0.001	...	0.001	0.001	0.001	maximize	0.001	0.002
	3	Eigenvector Centrality	1	0.005	0.004	0.004	0.004	0.004	0.003	0.003	...	0.001	0	0	maximize	0	0.005
			Score	1	0.24	0.24	0.24	0.22	0.17	0.17	...	0.06	0.01	0.01			
								Homophily: HomeTown=Chennai									

Figure 6: TOPSIS Computation Based On Hometown

				Options:											
				1	2	3	4	5		150	151	152			
			Importance	Santhosh	Narendra	Sundaresan	Ajeesh	Mahendra	...	Jaya	Pootha	Sharfaas	Goal	- Ideal	+ Ideal
Criteria:	1	Betweenness Centrality	1	14915.9	16257.1	2620.258	664.509	427.572	...	1.557	1.127	0.332	maximize	0.194	16257.1
	2	Closeness Centrality	1	0.002	0.002	0.001	0.001	0.001	...	0.001	0.001	0.001	maximize	0.001	0.002
	3	Eigenvector Centrality	1	0.008	0.005	0.015	0.02	0.019	...	0.001	0.001	0.001	maximize	0	0.02
			Score	0.82	0.81	0.25	0.25	0.24	...	0.01	0.01	0.01			
							Homophily: Sex=Male								

Figure 7: TOPSIS Score Based On Sex Status

Journal of Theoretical and Applied Information Technology
10th March 2015. Vol.73 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

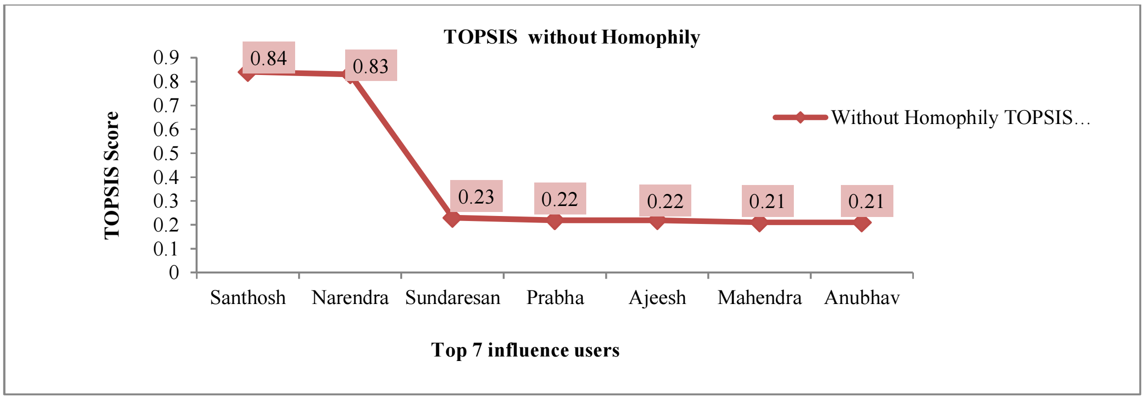

Figure 8: TOPSIS Score Without Homophily

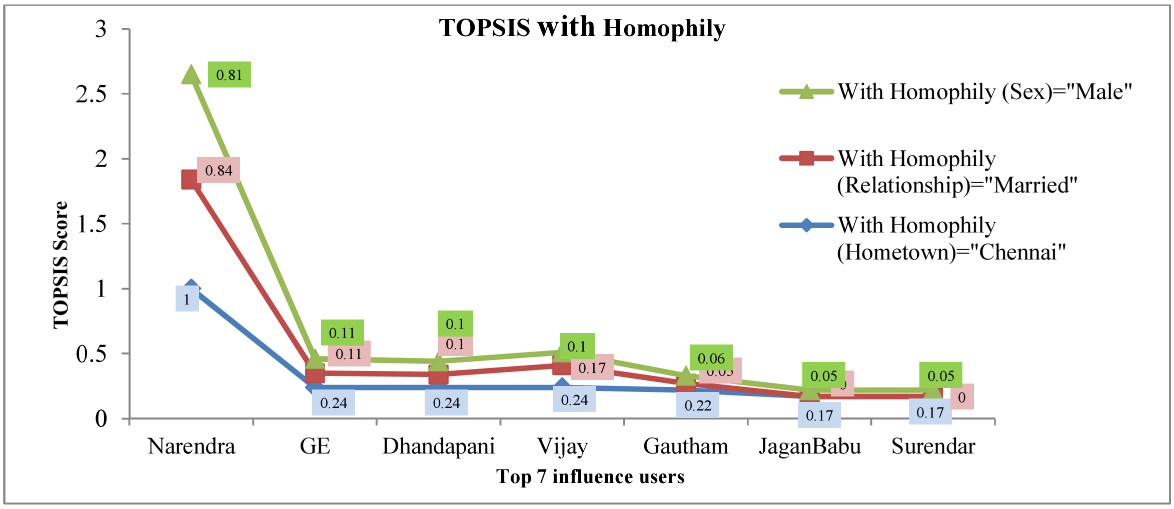

Figure 9: TOPSIS Score With Homophily

REFERENCES

[1] Goyal A., Bonchi F, and Laks V. S. Lakshmanan, "Learning influence probabilities in social networks", *WSDM '10 Proceedings of the third ACM international conference on Web search and data mining*, 2010, pp. 241-250.

[2] King I., Li J., and Chan T., "A Brief Survey of Computational Approaches in Social Computing", *IJCNN'09 Proceedings of the International joint conference on Neural Networks*, IEEE Press, 2009, pp. 2699-2706.

[3] Bharathi S., Kempe D., and Salek M., "Competitive influence maximization in social networks", *Lecture Notes in Computer Science*, 2007, pp. 306-311.

[4] Tang J., Sun J., Wang C., and Yang Zi, "Social influence analysis in large-scale networks," *Proceedings of the 15th ACM SIGKDD International conference on Knowledge discovery and data mining*, 2009, pp. 807-816.

[5] Pang B., Lee L., and Vaithyanathan S, "Thumbs up? Sentiment classification using machine learning techniques", *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2002, pp. 79-86

[6] Turney P. D., "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews", *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL '02)*, 2002, pp. 417-424.

[7] Weng J., Lim E. P., Jiang J., and Qi He, "TwitterRank: finding topic sensitive influential twitterers", *WSDM '10*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Proceedings of the third ACM international conference on Web search and data mining, 2010, pp. 261-270.

[8] Duanbing Chen, Linyuan Lü, Ming-Sheng Shang, Yi-Cheng Zhang, Tao Zhou, "Identifying influential nodes in complex networks", *Physica A: Statistical Mechanics and its Applications*, Volume 391, Issue 4, 2012, pp. 1777–1787.

[9] Kaiquan Xu, Xitong Guo, Jiexun Li, Raymond Y.K. Lau, Stephen S.Y. Liao (2012), "Discovering target groups in social networking sites: An effective method for maximizing joint influential power" , *Electronic Commerce Research and Applications*, Volume 11, Issue 4, 2012, pp. 318–334

[10] Zhigu Zhu, "Discovering the influential users oriented to viral marketing based on online social networks",*PhysicaA: Statistical Mechanics and its Applications*, Volume 392, Issue 16, 2013, pp. 3459–3469.

[11] Qian Li, Tao Zhou, Linyuan Lü, Duanbing Chen,"Identifying influential spreaders by weighted LeaderRank", *PhysicaA: Statistical Mechanics and its Applications*, Volume 404, 2014, pp. 47–55.

[12] Tang Xuning, Yang Christopher, "Ranking User Influence in Healthcare Social Media", *ACM Transactions on Intelligent Systems and Technology*, Vol. 3, No. 4, 2012, Article No. 73.

[13] Yuxian Dua, Cai Gaoa, Yong Hub, Sankaran Mahadevanc, Yong Denga,"A new method of identifying influential nodes in complex networks based on TOPSIS", *PhysicaA: Statistical Mechanics and its Applications*, Volume 399, 2014, pp. 57–69.

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

[15] Girvan M. and Newman M.E.J, "Community structure in social and biological networks", *Proceedings of the National Academy of Sciences of the United States of America*, Volume 99, Issue 12, 2002, pp. 7821–7826.

[16] Mark Granovetter, "Economic Action and Social Structure: The Problem of Embeddedness", *American Journal of Sociology*, Volume 91, Issue 3, 1985, pp. 481-510.

[17] Girvan M. and Newman M.E.J, "Community structure in social and biological networks", *Proceedings of the National Academy of Sciences of the United States of America*, Volume 99, Issue 12, 2002, pp. 7821–7826.

[18] Borgatti S. P. and Everett M.G, "A graph-theoretic perspective on centrality", *Social Networks*, 2006, pp. 466-484.

[19] Freeman L.C, "Centrality in social networks: Conceptual clarification", *Social Networks*, 1979, pp. 215-239.

[20] Chen S.J, Hwang C.L., "Fuzzy Multiple Attribute Decision Making: Methods and Applications", 1992, *Springer-Verlag, Berlin*.

[21] Hwang C.L., Yoon K., "Multiple Attribute Decision Making Methods and Applications", 1981, *Springer, Berlin Heidelberg*.

[22] HalilBisgin, Nitin Agarwal and Xia oweiXu, "Investigating Homophily in Online Social Network", *IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Web Technology*, 2010, pp. 533-536.

[23] Borgatti, S., Everett, M. G., and Freeman, L. C.,*UCINET 6, Analytic Technologies,* 2006.

[24] Shneiderman, B. and Aris, A., "Network visualization with semantic substrates", IEEE Symposium on Information Visualization and IEEE Trans. Visualization and Computer Graphics, Volume 12, Issue 5, 2006, pp. 733-740.