ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, Vol 9(32), DOI: 10.17485/ijst/2016/v9i32/95171, August 2016 ISSN (Online) : 0974-5645

Discovering and Ranking Influential Users in Social Media Networks Using Multi-Criteria Decision Making (MCDM) Methods

A. Muruganantham[1*] and Meera Gandhi[2]

[1]Sathyabama University, Chennai - 600119, Tamil Nadu, India; amuruganantham@gmail.com
[2]Department of Computer Science and Engineering, Sathyabama University, Chennai - 600119, Tamil Nadu, India; drmeeragandhii@gmail.com

Abstract

Background: Social media networks created highly interactive platforms through which individuals and communities share, discuss, collaborate. It is important to discover and rank the influential users. **Methods:** In an online social media customers or users trust the opinion of other known customers or users, especially those with prior experience of a product or service, rather than company suggestions or recommendations. In a dynamic business situation, a customer or user in an e-commerce site like Amazon tends to trust the buying experiences of his/her known friends rather than the buying recommendations from Amazon. **Findings:** This paper provides a comprehensive study of various Multi-Criteria Decision Making (MCDM) methods to understand or discover and rank influential users in an online social media network such as Facebook. Experiment results were demonstrated using tradition metrics such as Page Rank, Betweenness and Closeness centrality measures and compared with MCDM based methods. It is proved that MCMD based methods are precise, dynamic and capable of identifying or ranking the influence users preciously than the standard benchmarked traditional metrics. **Applications/Improvements:** A well-managed campaign with influential users, enterprises can get sustainable profit or growth rather than doing generalized campaign on their product or services. Our experimental performance results can be compared with benchmark results.

Keywords: Influence Users, Multi-Criteria Decision Making (MCDM) Methods, SDI, Social Media Network, TOPSIS

1. Introduction

Social media networks represent a revolution in Web 2.0 users behaviour that is spreading at an unprecedented rate during last few years. Online users at platforms such as Facebook and Twitter create large social media networks of millions of users that interact and group each other. Online users create social ties constituting groups based on existing relationships in real life, such as on relatives, friends, colleagues, or based on common interests, shared tastes, etc. Information is spreading out much faster than before throughout different online social communities. One of the challenges for researcher is to provide methods to collect[1] and process[2] data from online social networks in an automatic fashion, and strategies to reveal the features that characterize these types of complex networks[3]. In addition, these methods should be capable of working in such large-scale scenarios[4].

Analysing and identifying the community structure is a topic of great interest for its economical and marketing implications[5]. For example, it could be possible to improve the advertising campaign by identifying and targeting the most influential users of each community, exploiting effects such as the word-of-mouth and the spread of information within the community itself[6]. Similarly, exploiting the affiliations of users to communities might be effective to provide them useful recommendations on the base of common interests shared with their friends[7]. It is shown in[8] that customers or users tend to get recommendations from the users they know and trust, such as friends or

* Author for correspondence

relatives rather than from automated recommendation systems used in e-commerce websites like Amazon. Websites like Facebook, Twitters allow promoting of products by targeting influential users in their networks and allowing consumers to make comments on products. Using a social media network to propagate information can significantly affect customer decision making[9,10], whether or not to visit an e-commerce sites for buying or not buying a product from a particular retailer. Numbers of previous studies are found for identifying influence users using tradition structural analysis network metrics such as Degree, Page Rank, Betweenness and Closeness centrality[11–14]. Alternate class of methods for social media network analysis are also found and it investigates the behaviour of individuals in the network and their interactions.

In this paper, various MCDM methods applied to a social media network like Facebook to discover and rank the influence users and we compared the MCDM methods results against the traditional metrics and proved that MCDM methods are precise, dynamic and capable of identifying the influencers preciously than the standard benchmarked traditional metrics.

Although many research works proposed multiple techniques of generating social media network and ranking user influence using traditional structural network metrics, no direct approaches are not found to understand and rank the users based on their behavioural aspects. A common analysis is to understand the importance of users in social media network using a traditional measure such as Closeness centrality, Eigenvector centrality. Closeness centrality is the average length of all shortest paths from a user to all the other users in the social media network[15]. A node with higher Closeness centrality helps to reach rest of users very rapidly to the entire network. Another metric Eigenvector centrality helps to understand a user's popularity in asocial media network. Eigenvector centrality for a user is proportional to the sum of the Eigenvector centralities of all users directly connected to it[15]. Eigenvector centrality metric is also known as Google PageRank. The tradition metrics such as Closeness centrality, Eigenvector centrality understand the structure of the network rather than the behaviour of nodes and their dynamic interactions. It is important to consider analysis of users' behaviour in a network. Few researchers consider a weight for each tie and show the importance and strength of a link between two users based on their past interactions. In [16] used Spearman's rank correlation coefficient for comparing users' influence and compared In-degree, Retweets, and Mentions to analyse topics of the most influential users in Twitter network. In[17] revealed that user influence is not gained accidentally but requires that users need to be consistently active in the network.

In[18] analysed user influence as a combination of link strength (volume of interactions) and incoming and outgoing clustering value (based on Closeness) defined for each user in the network while they filter the spam and inactive nodes according to their activities and their interaction with other users.

In[17] used Dynamic Graph Analysis. They analysed each users in a social media network using the number of daily mentions as an indicator for computing different ranks such as PageRank, drank, and starrank. Khrabrov and Cybenko used several primitive indices in combination, such as Contiguous Longest Increasing Subsequences (CLIS) and GrowFall for analysis of influence ranks during a period of time and explained how the influence rank of a user changes with time. They also analysed the rate of increase in the number of mentions for influencing users in a social media network.

2. Modelling the User Influence in Social Media Networks

In this paper, we consider a user influence to convince others to change their mind about making a specific decision and it depends on various factors, such as reputation, social personality and psychological influencing abilities. In a community social media network site like healthcare community site, the postings of one person on a particular disease is depends on how that person is articulated his/her disease infection and how much it is relevance to his age group users. On the one side, the quality of a posting is important for the followers of a specific user as well as the subject of a posting or a tweet. A topic might be interesting for a group of users while it is bothersome or boring for another group of users in this healthcare community network. On the other hand, the quality of the personality of a person like popular film celebrity is another aspect of being influential in a network.

2.1 Social Media Network Modelling

In order to analyze a social media network, the network is formally defined as SN = (G, P, L,C), is a tuple in which G is a graph representing the network structure, P is the set of postings posted by the individuals in the network, L is the set of Likes about the postings and C is the set of comments posted by individuals about the postings. In the graph G =< V, E >, V is the set of vertices representing individuals and E is the set of edges representing ties between individuals. In some models, e = fa; bg just represents a tie and the direction of the relation is not important. Every edge is represented by an ordered set, e =< a,b > representing a directed edge from the node a to b. In the network SN, a is called a follower of b if in the graph of their network, there exists an edge < a, b > in E. Stated more formally:

$$follow(a,b) \leftrightarrow \exists\ e = < a,b > \in E$$

The following definitions are important in order to model the user influence in the network

SN = (G,P, L,C):

- $|X|$ returns the cardinality of the set X and Po(X) is the power set of the set X.
- F(v): is a mapping function from V to set of vertices $F{:}V \rightarrow Po(V)$ ($Po(V)$ is the power set of the set V). This function returns the set of nodes which are following the node $v \in V$ in a social media network.
- P(v): is a mapping function from V to the power set of postings $P{:}V \rightarrow Po(P)$. This function returns the set of postings posted by the node $v \in V$ in the SN.
- L(p) is a function returning the set of individuals who liked the posting $p \in P. L{:}P \rightarrow Po(V)$.
- C(p) is a function returning the set of individuals who commented on the posting $p \in P. C{:}P \rightarrow Po(V)$.
- RT(p) is a function returning the set of individuals who retweeted the posting $p \in P$ in the network. $RT : P \rightarrow Po(V)$.
- LCRT is a function that determines whether a particular user has liked, commented on or retweeted a particular posting in the network. $LCRT : V \times P \rightarrow \{1,0\}$.

$$LCRT(v,p) = \begin{cases} 1 & if\ (v \in L(p) \cup C(p) \cup RT(p)) \\ 0 & otherwise \end{cases} \quad (1)$$

2.2 Social Influence

In a social media network, a user behavioral change is affected by other users is called Social influence. Social influence is an intuitive and well-accepted phenomenon in social networks[19]. 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 tradition network metrics such as Degree, Closeness centrality, Eigenvector centrality, Betweenness centrality and Page Rank mentioned in the equations (1), (2), (3), (4) and (5) are related to social influence in terms of the structural effects of social media networks.

2.3 Basic Formulations

In a social media network, the tradition standard network metrics such as Degree, Closeness centrality, Eigenvector centrality are defined[20] as,

2.3.1 Degree

The simplest and most popular measure is degree centrality. The degree k of a node is the number of edges connected to it.

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:

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

2.3.2 Closeness

It is a natural distance metric between all pairs of nodes, defined by the length of their shortest path. The farness of a node x is defined as the sum of its distance from all other nodes, and its closeness was defined by Bavelas as the reciprocal of the farness. 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:

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

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.

2.3.3 Node Betweenness or Betweenness Centrality

Betweenness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes. 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:

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

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.

2.3.4 Eigenvector Centrality

It is a measure of the influence of a node in a network. It assigns relative scores to all nodes in the network based on the concept that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. Google's PageRank is a variant of the eigenvector centrality measure. Let x(i) be the Eigenvector centrality of a node vi. Then,

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

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.

3. MCDM Methods to Discover and Rank the Influential Users

Although many research works proposed multiple techniques of generating social media network and ranking user influence using traditional structural network metrics, no direct approaches are not found to understand and rank the users based on their behavioral aspects. Multi Criteria Decision Making (MCDM) methods are pertaining to structure and solve decision and planning problems involving multiple criteria and it support decision makers to rank influence users with multiple choices (characteristics) exist for a problem to be solved. MCDM methods have been applied to different applications and find the best solution to choose the best alternative.

3.1 Classification of MCDM Methods

Various MCDM methods were emerged over the period of time and each method has its own characteristics. One way is to classify them based on the type of the data they use. Another way of classifying MCDM methods is according to the number of decision makers involved in the decision process.

The Figure 1 depicts the hierarchical view MCDM methods and its important types. The widely used MCDM methods have been described in following headings.

3.2 The ELECTRE Method

ELECTRE[21] is a family of multi-criteria decision analysis methods that originated in Europe in the mid-1960s. The acronym ELECTRE stands for: ELimination Et Choix Traduisant la REalité (ELimination and Choice Expressing REality). It was first applied in 1965, the ELECTRE method was to choose the best action(s) from a given set of actions, but it was soon applied to three main problems: choosing, ranking and sorting. The method became more widely known when a paper by B. Roy appeared in a French operations research journal and it evolved into ELECTRE I (electre one) and the evolutions have continued with ELECTRE II, ELECTRE

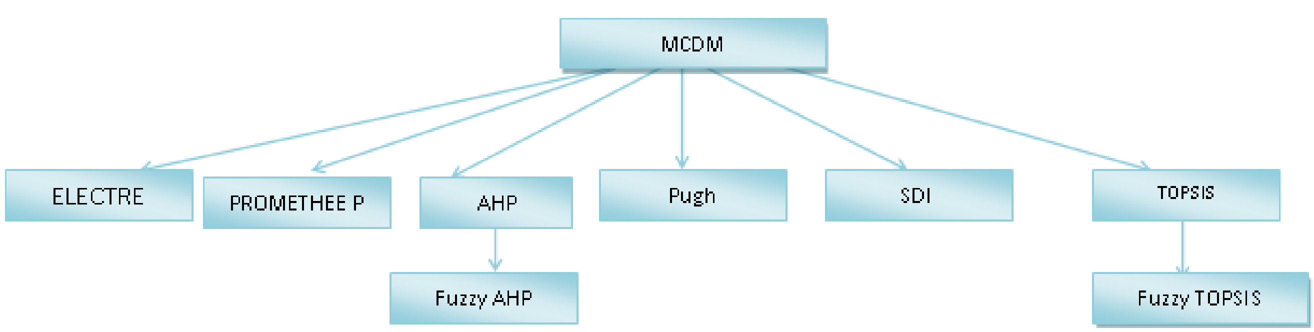

Figure 1. Hierarchical structure of MCDM methods.

III, ELECTRE IV, ELECTRE IS and ELECTRE TRI (electre tree), to mention a few.

ELECTRE is a two-step technique. First, the construction of one or several outranking relations, which aims at comparing in a comprehensive way each pair of actions; second, an exploitation procedure that elaborates on the recommendations obtained in the first phase. The nature of the recommendation depends on the problem being addressed: choosing, ranking or sorting.

The ELECTRE method yields a system of binary outranking relations between the alternatives. Because this system is not necessarily complete, the ELECTRE method is sometimes unable to identify the most preferred alternative. It only produces a core of leading alternatives. This method has a clearer view of alternatives by eliminating less favourable ones. This method is especially convenient when there are decision problems that involve a few criteria with a large number of alternatives.

3.2 PROMETHEE

PROMETHEE[22] is similar to ELECTRE in that it also has several iterations and is also an outranking method.

The PROMETHEE family of outranking methods such as PROMETHEE I for partial ranking of the alternatives and the PROMETHEE II for complete ranking of the alternatives, were developed in 1982. Several versions of the PROMETHEE methods such as the PROMETHEE III for ranking based on interval, the PROMETHEE IV for complete or partial ranking of the alternatives when the set of viable solutions is continuous, the PROMETHEE V for problems with segmentation constraints, the PROMETHEE VI for the human brain representation.

Though the PROMETHEE is easy to use and it does not provide a clear method by which to assign weights and it requires the assignment of values but does not provide a clear method by which to assign those values.

3.3 Analytic Hierarchy Process (AHP)

AHP[23] is a structured technique for organizing and analyzing complex decisions, based on mathematics and psychology. It was developed by Thomas L. Saaty[24] in the 1970s and has been extensively studied and refined since then. Many researchers have preferred to use AHP to find the weights of criteria[25,26]. Due to the fact that criteria weights in the decision-making problems are various, it is not correct to assign all of them as equaled[27]. To solve the problem of indicating the weights some methods like AHP, eigenvector, entropy analysis, and weighted least square methods were used. For the calculation of criteria weight in AHP the following steps are used:

- Arrange the criteria in n × n square matrix form as rows and columns.
- Using pairwise comparisons, the relative importance of one criterion over another can be expressed as follow:

If two criteria have equal importance in pairwise comparison enter 1; if one of them is moderately more important than the other enter 3 and for the other enter 1/3; if one of them is strongly more important enter 5 and for the other enter 1/5; if one of them is very strongly more important enter 7 and for the other enter 1/7, and if one of them is extremely important enter 9 and for the other enter 1/9. 2, 4, 6 and 8 can be entered as intermediate values. Thus, pairwise comparison matrix is obtained as a result of the pairwise comparisons. Then

- Take the square of n × n matrix.
- To find the first eigenvector, sum the rows of the matrix.
- Sum the row totals.
- Normalization is done by dividing the row sum by the total rows.
- The obtained result is first eigenvector.

3.4 Pugh or Decision Matrix Method

Pugh or Decision Matrix Method[28] (DMM) is qualitative selection method used to make fast qualitative comparisons by comparing an option against a datum or reference option. DMM technique is interesting for looking at large numbers of factors and assessing each relative importance. Furthermore, DMM[29] is a method for alternative selection using a scoring matrix. A basic decision matrix consists of establishing a set of criteria options which are scored and summed to gain a total score which can then be ranked.

Let C be the criteria vector of a DMM $C=(c_1, c_2, \ldots c_n)$ where C_j belongs to the criteria domain of the problem and n is the total number of criteria.

Let W be the weights criteria vector of DMM $W=(w_1,w2,\dots wn)$ where $wj\varepsilon[0,N] \mid N \neq \infty$

Let Ai be the rating vector of i alternative $Ai=(a1, a_2,\dots, a_n)$ where $a_m € \{-1,0,1\}$

Consider the matrix D be defined by $D=(a_{ij})$ where a_{ij} is the rating of alternative to i to the criterion j,a_i, $€ \{-1,0,1\}$. D is called the rating matrix of the DMM

Consider the vector S be define by S= W X D being $D=(s_1, s_2,..s_m)$ where s_k is the product of weight i by alternative j and m is the number of alternatives.

$$(s_1, s_2,..s_m) = (w_1, w_2,\dots, w_n) \ x \begin{pmatrix} a11 & \dots & am1 \\ .. & .. & \dots \\ a1n & \dots & amn \end{pmatrix}$$

The highest s_k will be the team's proposal for the problem analysed.

3.5 SDI Method

SDI Matrix method[30] is Quantitative multi-criteria decision analysis developed by Statistical Design Institute using Value Assessments for product specifications. It is based on quantitative selection for detailed design. SDI method rate and rank multiple user-defined design options against multiple user-defined criteria using a quantitative approach. SDI's approach to design trade-off is to evaluate how well the Design Option meets each Design Criterion by performing detailed, quantitative comparisons. The Total Score for each Design Option is calculated as the weighted sum of each Design Option Score per Design Criterion.

The procedure of SDI method is expressed in a series of steps:

- Create matrix with Design Options and Criteria
- Assign Importance to Criteria
- Enter LSL, Target, and/or USL values from known specifications for each Criteria
- Enter Design Option data for each Criteria
- Select the Criteria Scoring Method (Linear Inspec, Quadratic Inspec, In/Out Spec, Minimize Maximize, or 1-PNC))
- Compute and Evaluate Scores

3.6 TOPSIS

TOPSIS[31] (technique for order preference by similarity to an ideal solution) method was developed by Hwang and Yoon as an alternative to the ELECTRE. TOPSIS is a multiple criteria method to identify solutions that is closest to the ideal solution from a finite set of alternatives. The selected alternative in TOPSIS should have the shortest distance from the ideal solution and the farthest distance from the negative-ideal solution in some geometrical sense.

The TOPSIS method carried out in a series of steps:

- Create an evaluation matrix consisting of m alternatives and n criteria i.e, $(n_{ij})_{mxn}$
- Calculate the normalized decision matrix. The normalized value n_{ij} is calculated as $n_{ij}= x_{ij} / \sqrt{\sum_{j=1}^{m} \quad x^2_{ij}}$
- Calculate the weighted normalized decision matrix. The weighted normalized value v_{ij} is calculated as
- $v_{ij}=w_i n_{ij}$, $j=1, \dots,m$, $i=1, \dots,n$,
- where w_i is the weight of the i^{th} attribute or criterion, and $\sum_{j=1}^{m} \quad w_j =1$
- Determine the positive (best) ideal and negative (worst) ideal solution.
- $A^+ = \{ v^+_1, \dots v^+_n\} =\{ (\underset{j}{max} \ v_{ij} \mid i \ \varepsilon \ I), (\underset{j}{min} \ v_{ij} \mid i \ \varepsilon \ J) \}$,
- $A^- = \{ v^-_1, \dots v^-_n\} =\{ (\underset{j}{min} \ v_{ij} \mid i \ \varepsilon \ I), (\underset{j}{max} \ v_{ij} \mid i \ \varepsilon \ J) \}$,
- where, I is associated with benefit criteria, and J is associated with cost criteria.
- Calculate the separation measures, using the n-dimensional Euclidean distance. The separation of each alternative from the ideal solution is given as
- $d^+_j = \{\sum_{i=1}^{n} \ (v_{ij} - v^+_i)^2\}^{½}$, $j=1,\dots,m$.
- Similarly, the separation from the negative ideal solution is given as
- $d^-_j = \{\sum_{i=1}^{n} \ (v_{ij} - v^-_i)^2\}^{½}$, $j=1,\dots,m$.
- Calculate the relative closeness to the ideal solution i.e., similarity to the worst. The relative closeness of the alternative A_j with respect to A^+ is defined as
- $R_j= d^-_j / (d^+_j + d^-_j)$, $j=1,\dots,m$.
- Since $d^-_j \geq 0$ and $d^+_j \geq 0$, then, clearly, $R_j \ \varepsilon \ [0, 1]$.
- Rank the alternatives according to Rj (j=1,2,..m).

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. Fuzzy TOPSIS (Fuzzy Technique for Order of Preference by Similarity to Ideal Solution) and Fuzzy AHP[32] (Fuzzy Analytic Hierarchy Process) to aid the selection decision process better than the existing TOPSIS and AHP methods.

The advantages and disadvantages of popular MCDM methods are summarized in Table 1.

4. Experimental Study

In this section, we present the details of experiments on real datasets to evaluate the performance of popular MCDM methods. We applied various MCDM methods to a Facebook data set and ranked the user influences.

4.1 A Facebook Dataset

An extracted Facebook network dataset, download from our Facebook account is used for our experimental study. The extracted Facebook data in xlsx file is represented through nodeXL as network graph structures and found 260+ users connections.

4.2 Traditional Network Metrics Calculation

The NodeXL tool represents the imported dataset in the form of nodes and edges. Figure 2 displays the uploaded Facebook dataset in NodeXL.

NodeXL built-in software routines computes automatically the Centrality measures such as degree, in-degree, out-degree, clustering coefficient, and closeness, Betweenness, and eigenvector centrality etc. as highlighted in the Figure 2.

4.3 Discovering and Ranking Influencers

The Influencer users can be extracted based on traditional structural network metrics from the computed values. In

Table 1. Summary of MCDM Methods

Method	Advantages	Disadvantages
PROMETHEE	Easy to use	Does not provide a clear method on weights assignments
ELECTRE	Outranking is used	Process and outcome is difficult to explain; Time consuming
AHP	Easy to use; scalable; hierarchy structure helps to fit many sized problems;	Irregularities in ranking; Additive aggregation is used and more pair-wise comparisons are needed
SDI	Quantitative based technique for detailed evaluations of various design options	The best design option may exceeds the bounds of one or more the design criteria
Pugh	Selection/decision making is subjective opinions about one alternative versus another can be made more objective	Selection is based on the design team's knowledge
TOPSIS	Easy to use and program; the number of steps remains same independent of problem size.	Difficult to weight and keep consistency of judgment.

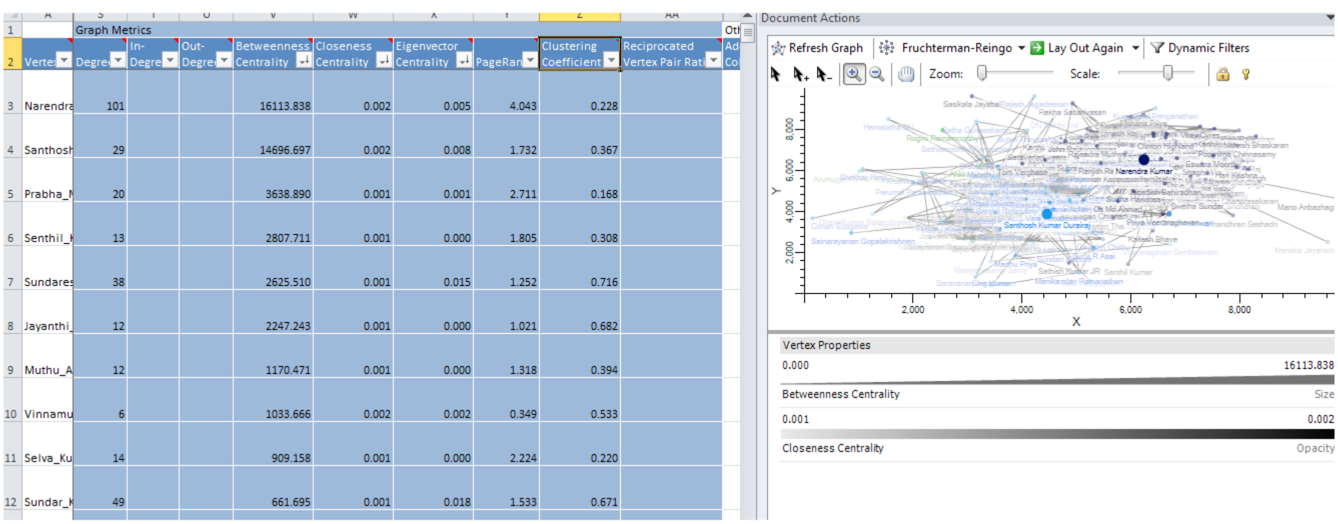

Figure 2. Uploaded Facebook dataset in NodeXL.

Figure 3, the influence users are sorted descending order of Betweenness centrality values. The highest Betweenness centrality value indicates that the users (vertices) with high value can spread any information to the rest of users in the Facebook network very rapidly than the vertices having zero Betweenness centrality as shown in Figure 3. In a social media network, a connection to a popular individual (having highest Eigenvector Centrality) is more important than a connection to a loner because the Eigenvector centrality network metric combines how many connections a vertex has (i.e., its degree) and also the degree of the vertices that it is connected to. Figure 4 shows the ranking of users based on the Eigenvector. Similar approach is followed to rank the users based on the PageRank values as mentioned in Figure 5. Depending upon the business situation or context, an appropriate centrality measure says Eigenvector centrality or PageRank or Degree is chosen as measure to rank the influence users. This approach will work for static business situations or context. MCDM methods use various centrality measures as inputs and rank the influence users by choosing appropriate business context or situations. Considering all centrality measures are equality importance, the value of importance or priority is assigned as 1 for all three centrality measures: Betweenness centrality, Closeness centrality and Eigenvector centrality. Subsequently TOPSIS or SDI method is applied to rank influencer by considering multiple centrality measures synthetically as shown in the Figures 8 and 9 rather than ranking the influence users using one centrality measure.

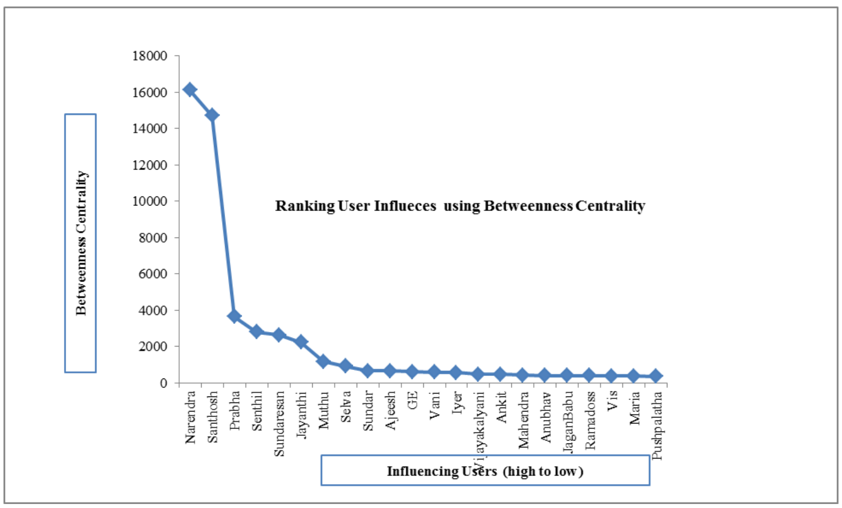

Figure 3. Ranking influencers using betweenness centrality.

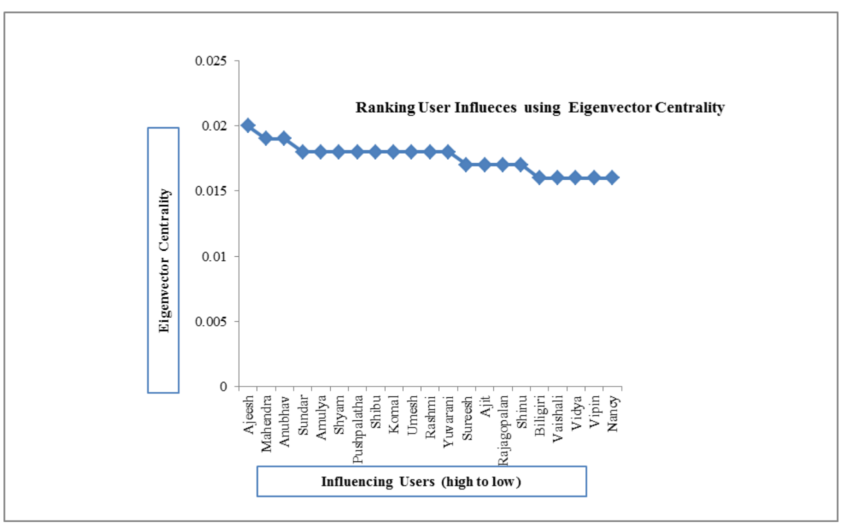

Figure 4. Ranking influencers using eigenvector centrality.

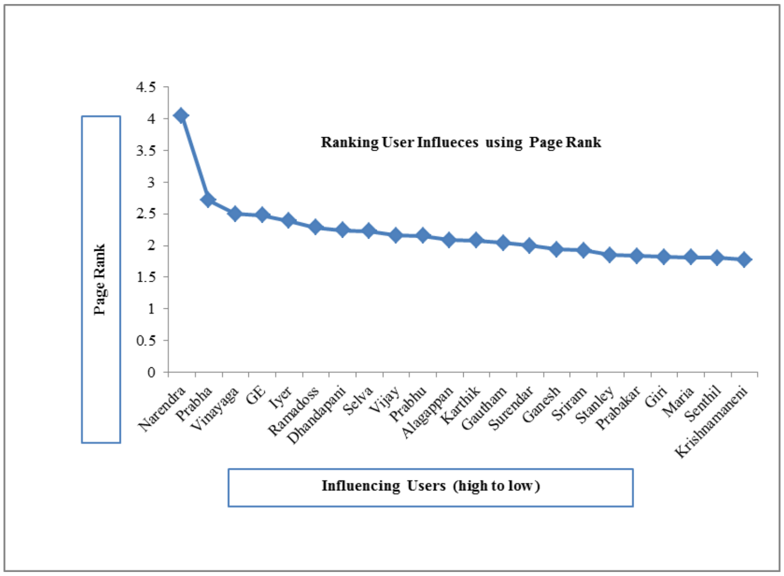

Figure 5. Ranking influencers using page rank.

			Ranking User Influences																
				1	2	3	4	5	6	7		171	172	173	174				
			Importance	Santhosh	Narendra	Sundaresan	Prabha	Ajeesh	Mahendra	Anubhav	...	Vidhya	Vkey	Karthikeyan	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 6. Ranking influencers computation using MCDM: TOPSIS method.

			LSL	Target	USL		1	2	3	4	5	6	7		171	172	173	174				
			[redacted]			Importance	Narendra	Santhosh	GE	Vinayaga	Iyer	Ramadoss	Dhandapani	...	Shyam	Anantha	Pootha	Pugazhenthi	...	Criteria Scor	Max	Min
Criteria:	1	Betweenness Centrality				1	16113.84	14696.7	619.495	362.254	567.883	402.856	273.903	...	333.073	33.674	1.127	20.739	...	maximize	16113.84	0
	2	Closeness Centrality				1	0.002	0.002	0.002	0.002	0.002	0.002	0.002	...	0.001	0.001	0.001	0.001	...	maximize	0.002	0.001
	3	Eigenvector Centrality				1	0.005	0.008	0.004	0.004	0.004	0.004	0.004	...	0.018	0	0.001	0	...	maximize	0.02	0
	4	Page Rank				1	4.043	1.732	2.473	2.499	2.389	2.283	2.239		1.484	0.821	0.656	0.82		maximize	4.043	0.561
						Score	**13**	**10.59**	**7.15**	**7.12**	**7.04**	**6.88**	**6.8**		**4.74**	**0.31**	**0.31**	**0.3**				

Figure 7. Ranking influencers computation using MCDM: SDI method.

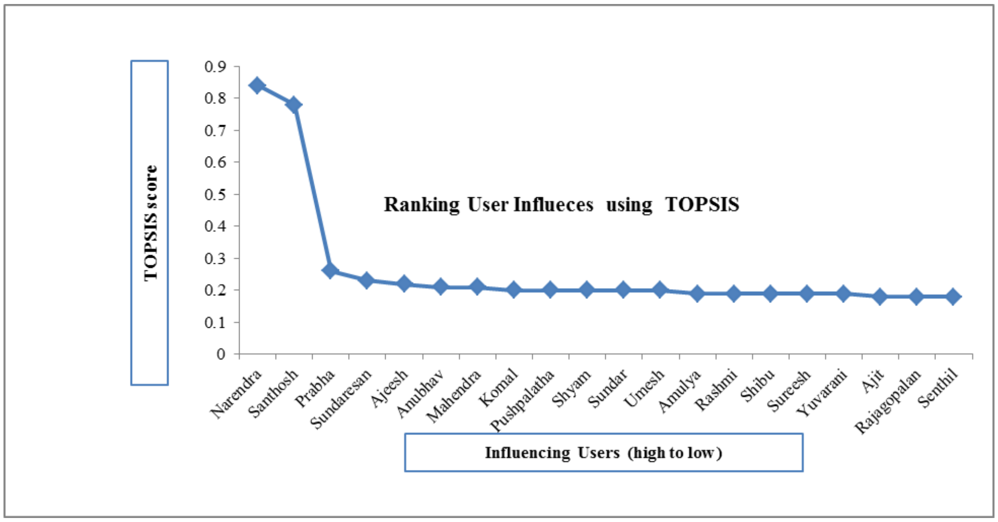

Figure 8. Ranking influencers using TOPSIS.

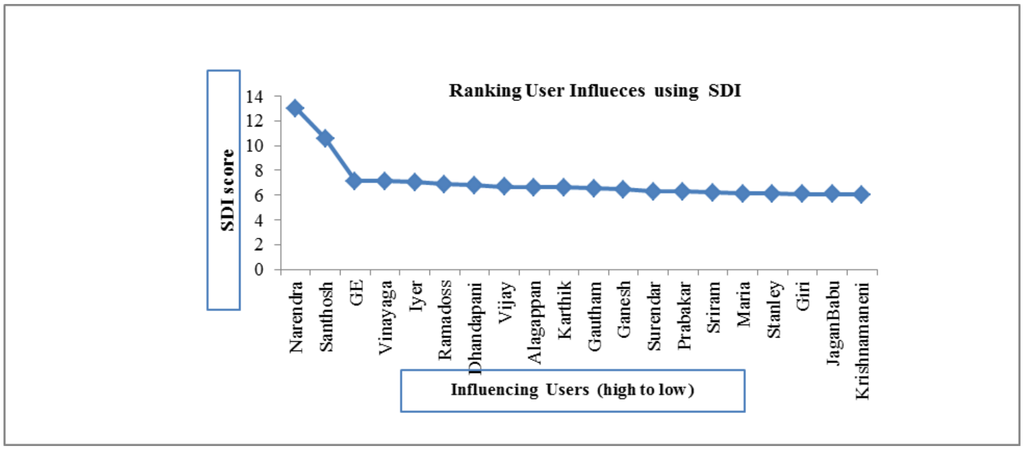

Figure 9. Ranking influencers using TOPSIS.

In TOPSIS, the influencers are ranked closes to ideal values as mentioned in +ideal or -ideal column in Figure 6. Similarly, SDI method defined the acceptable maximum and minimum values for various centralities measures as shown in Figure 7. MCDM based methods apply more than one centrality measures equally and discover and rank the influencers rather than using one centrality metrics, which is followed in a tradition approach.

5. Conclusions and Future Works

This paper proposes a model for discovering the influencers in a social media network. Rather using one centrality measure, MCDM based method uses more than one centrality measures to identify the influence users in a social media network. Depending upon the business situation or content, an appropriate number of centrality measures are chosen and applied using MCDM methods to discover and rank the influence users. For example, a business organization wants to rank the user influences based on popularity and maximum influence in a social media network. The Page Rank and Eigenvector centrality measure combination is chosen as centrality measures to rank the user influences using MCDM methods. The popular MCDM based methods such as TOPSIS & SDI ranks influence users using more than one centrality measures as shown in Figure 7,8,9. Ranking influence users using more than one centrality measures and applying MCDM methods are meaningful because they describe more than one business situation/context and preciously understand the influence users unlike the traditional structured network metric rank influence users based on one attribute at a time and it is not sufficient to represent all business situations or contexts.

The above work can be improved in the future by including the below mentioned challenges,

- Dynamic properties of social media network can be included while the centrality measures are calculated
- MCDM methods experimental performance results can be compared with benchmark results.

6. References

1. Ferrara E, Fiumara G, Baumgartner R. Web data extraction, applications and techniques: a survey. Technical Report; 2010.
2. Heer J, Shneiderman B. Interactive dynamics for visual analysis. ACM Queue. 2012 Apr; 10(2):1–30.
3. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D. Complex networks: structure and dynamics. Physics Reports. 2006 Feb; 424(4–5):175–308
4. Backstrom L, Huttenlocher D, Kleinberg J, Lan X. Group formation in large social networks: membership, growth, and evolution. Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, New York; 2006 Aug. p. 44–54.
5. Goldenberg J, Libai B, Muller E. Using complex systems analysis to advance marketing theory development: modelling heterogeneity effects on new product growth through stochastic cellular automata. AMS Review. 2001 Jan; 9(3):1–18.
6. Goldenberg J, Libai B, Muller E. Talk of the network: a complex systems look at the underlying process of word-of-mouth. Marketing Letters. 2001 Aug; 12(3):211–23.
7. Zhou X, Xu Y, Li Y, Josang A, Cox C. The state-of-the-art in personalized recommender systems for social networking. Artificial Intelligence Review. 2012 Feb; 37(2):119–32.

8. Salehi-Abari A, White T. The impact of naive agents in heterogeneous trust-aware societies. Multi-Agent-Based Simulation X; 2010. p. 110–22.
9. Kim Y, Srivastava J. Impact of social influence in e-commerce decision making. Proceedings of the Ninth International Conference on Electronic Commerce; 2007 Aug. p. 293–302.
10. Ma H, Yang H, Lyu M, King I. Mining social networks using heat diffusion processes for marketing candidates selection. Proceeding of the 17th ACM conference on Information and Knowledge Management; 2008 Oct. p. 233–42.
11. Fink M, Spoerhase J. Maximum betweenness centrality: approximability and tractable cases. Algorithms and Computation; 2011 Feb. p. 9–20.
12. Goh K, Oh E, Kahng B, Kim D. Betweenness centrality correlation in social networks. Physical Review E. 2003 Jan; 67(1).
13. Puzis R, Elovici Y, Dolev S. Finding the most prominent group in complex networks. AI Communications. 2007 Dec; 20(4):287–96.
14. Suri NR, Narahari Y. Determining the top-k nodes in social networks using the Shapley value. Proceedings of the 7th international joint conference on Autonomous Agents and Multiagent Systems. 2008; 3:1509–12.
15. Borgatti S. Centrality and network flow. Social Networks. 2005; 27(1):55–71.
16. Cha M, Haddadi H, Benevenuto F, Gummadi K. Measuring user influence in twitter: The million follower fallacy. Proceedings of the 4th International Conference on Weblogs and Social Media; 2010. p.1–8.
17. Khrabrov A, Cybenko G. Discovering influence in communication networks using dynamic graph analysis. IEEE International Conference on Social Computing; 2010. p. 288–94.
18. Rad A, Benyoucef M. Towards detecting influential users in social networks. E-Technologies: Transformation in a Connected World. 2011 Jan; 78: 227–40.
19. Easley D, Kleinberg J. Networks, crowds, and markets: Reasoning about a highly connected world, Cambridge University Press; 2010.
20. Centrality [Internet]. [Cited 2015 Oct 01]. Available from: https://en.wikipedia.org/wiki/Centrality.
21. ELECTRE [Internet]. [Cited 2015 Oct 01]. Available from: https://en.wikipedia.org/wiki/ELECTRE.
22. Velasquez M, Hester PT. An analysis of multi-criteria decision making methods. International Journal of Operations Research. 2013; 10(2):56–66.
23. Analytic hierarchy process [Internet]. [Cited 2015 Oct 01]. Available from: https://en.wikipedia.org/wiki/Analytic_hierarchy_process.
24. Saaty TL. The analytic hierarchy process: Planning, priority setting. McGrawHill, New York; 1980 Jan.
25. Shimizu Y, Jindo T. A fuzzy logic analysis method for evaluating human sensitivities. International Journal of Industrial Ergonomics. 1995; 15(1):39–47.
26. Tsaur SH, Chang T-Y, Yen C-H. The evaluation of airline service quality by fuzzy MCDM. Tourism Management. 2002 Apr; 23(2):107–15.
27. Chen MF, Tzeng GH. Combing grey relation and TOPSIS concepts for selecting an expatriate host country. Mathematical and Computer Modelling. 2004 Dec; 40:1473–90.
28. Decision-matrix method [Internet]. [Cited 2015 Oct 01]. Available from: https://en.wikipedia.org/wiki/Decision-matrix_method.
29. Salmeron JL, Smarandache F. Redesigning decision matrix method with an indeterminacy-based inference process. Advances in Fuzzy Sets and Systems. 2006; 1(2):263–71.
30. SDI [Internet]. [Cited 2015 Oct 01]. Available from: http://www.stat-design.com/.
31. TOPSIS [Internet]. [Cited 2015 Oct 01]. Available from: https://en.wikipedia.org/wiki/TOPSIS#References.