Multi-Perspective Elicitation of Influential Parameters and Measures in Social Network

Sakshi Agarwal, Shikha Mehta

Abstract*: Recently, there has been enormous curiosity in the phenomenon of influence spread in social and information networks. Researchers from various domains are trying to develop effective and efficient model to estimate social influence. In these models, selection of parameters and measures of the social network for influence estimation is arbitrary. Moreover, these models use different parameters to estimate the influence with respect to same dataset. However, domain specific user's parameters in influence analysis for social network have not been addressed aptly so far. As, there is no study that precisely focuses on what parameters and measures are suitable for which application, this work presents multi-perspective elicitation of influential parameters and measures in social network. In this paper, we categorized social influence parameters in three classes based on domain specific user's characteristics i.e. uniform properties of user & their relationship with others time dependent user behaviour and user interaction. Through experiments, we analysed the impact of these three parameter type in the field of diffusion. Statistics revealed that there are various parameters, which have not yet been explored. Like time dependent user behaviour parameters and user interaction based parameters have enough scope for research. Therefore, this work intended to open the new road map for researchers and practitioners who aim to develop new technique in this area i.e. social influence.

***Index Terms*: Influence parameters, Social influence analysis, Social influence measures, Social network analysis**

I. INTRODUCTION

With the rapid growth of social media portals, there is a large quantity of user-generated content available on the Web. Additionally, the attractiveness and the status of social media websites like Twitter, MySpace and Facebook are growing persistently. The rising quantity of information spread through Social network pushes the members of these networks to fight for the consideration of influence. This deliberation of influence is relying on other people to spread their significance i.e. social influence.

Social influence is defined as change in the activities of an individual due to another person's intentional or unintentional actions and activity. It creates an indirect hidden relationship between the impelled person and the influencer, which may continue by the person in future as well. To understand the person's behaviour on online social network, presence has been considered as a basic element [80]. This presence of a user influence is classified in three parts based on the type of change in the user's behavior [33], [78]: conformity, compliance and obedience. Conformity is the unwitting change in a person's behavior to become or behave like others by following their actions or activity. This leads to change in one's behavior as soon as influencer changes their belief and values. Second area is compliance, which is a change in person's behavior due to explicit request from others. It is a deliberate change in actions performed by a person i.e. recommended to act upon by another person. Obedience is a decisive change in behavior of a person in which that person believed that he/she does not have any choice to refuse to comply the recommended actions of influencer. Overall, user is the main key component for the influence analysis such as user behaviour, user actions or interactions between users etc. Predicting all these behaviors in social influence is useful in wide variety of applications i.e. viral marketing, health-care [26], social psychology [33], drive-risk analysis [65], home education [67] etc. In recent years, many researchers formulated various models and techniques for acquiring and estimating influence in online social networks.

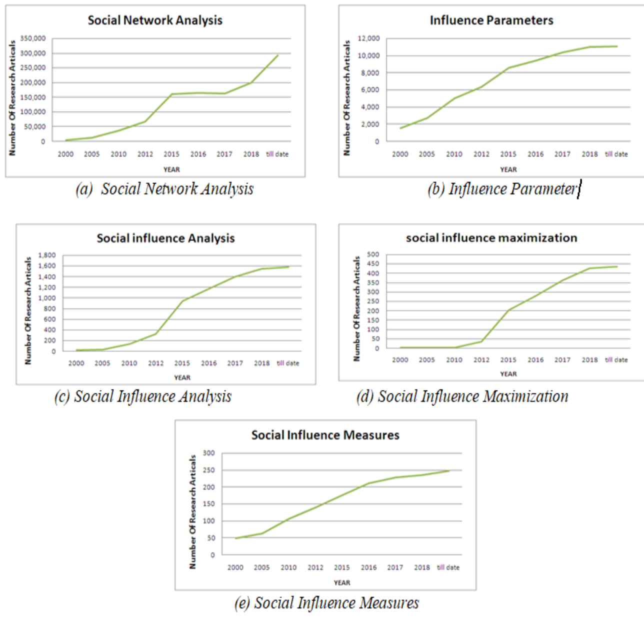


Fig. 1. Statistical Usage Analysis of various keywords used in Social Networks

In this regard, they have used different measures and parameters to analyze social networks [70], [77]. Using these measures, they extracted knowledge about the social networks to estimate the information diffusion.

Revised Manuscript Received on June 04, 2019

Sakshi Agarwal, Department of Computer Science & Engineering and Information Technology, Jaypee Institute of Information Technology, Noida, India.

Shikha Mehta, Department of Computer Science & Engineering and Information Technology, Jaypee Institute of Information Technology, Noida, India.

In literature, we observed that social influence interpretation depends upon variety of influence measure and user's behavioral parameters, it is not necessary that behavior of a user is always same with respect to each domain/platform. For example, users interact differently on a research platform such as Researchgate, Quora etc. by write, like or comment on a post. Whereas, the same user can upload, like, comment or share pictures, text, create groups, pages or events on social media platforms such as facebook or twitter etc. Therefore, influence analysis not only depends upon the domain/user's characteristics, instead of it depends upon the domain specific user characteristics i.e. these measures and parameters are domain specific and dataset dependent [1].

We performed the statistical analysis of the various keywords i.e. social influence measures, social network analysis and influence parameters etc. to obtain the current state of art in the domain of social networks. We performed the usage analysis of keywords through the Google scholar i.e. how many articles include that particular keyword in their work as shown in fig. 1. Fig. 1 (a) shows the usage count of keyword "Social Influence Analysis" i.e. nearly, 350,000 articles have been included this keyword so far. It is clear from the fig. 1 that usage of all influence related keywords are rapidly increasing. Although, all keywords usage count are less than keyword "social network analysis". Therefore, these statistics indicate the scope of research in the field of social influence. Thus, it creates the need to understand the depth of social influence measures, which will further lead to enrichment in the efficiency and accuracy estimation of social influence. This further depends on selection of appropriate parameters according to nature of the social network. As per our knowledge, there is no study in literature that focuses on identification of domain specific parameters and measures in the area of social influence. Therefore, selection of suitable set of parameters and measure is an intensifying research problem in the field of social influence.

The main goal of this paper is to elicit various influence measures and parameters. In this paper, we also provide a simplified three level view of these measures. With this in mind, we pointed out the inclination and concept of existing influence measures. We also categorized the influence parameters in three categories i.e. Type 1, Type 2 and Type 3. Type 1 is based on user's uniform properties, Type 2 is based on time dependent user behavior and Type 3 is based on user interaction. This can help researchers in the selection of appropriate set of parameters to attain the more accurate influence analysis by embedding them in different social measures. Therefore, this work is planned to provide the road map for researchers and practitioners who are going to develop new models in this domain i.e. social influence. Our work also will be adjuvant for developers who wish to utilize current obtainable models on particular network applications, as we show a domain specific user's characteristic-wise parameter categorization of parameters w.r.t social influence. The rest of this paper is organized as follows: In Section 2, we present various measures used to detect influence in social networks. Then the categorization of the influence parameters is given in Section 3. Next, we present the literature analysis with respect to influence measures and type of parameters in Section 4 and Section 5. In the last section, we summarize the key findings of the survey work and point out some open questions.

II. SOCIAL INFLUENCE ASSOCIATED TERMINOLOGY

A social network can be presented as a graph $G = \{V,E\}$, where V is the set of all users (nodes), and E is the set of edges. The relationship between the two users is represented by links/edges. These links correspond to social relationships. A user can be influenced by another user/edge in two ways: Locally or globally [1]. In local influence, user A gets influenced by some other user B through the properties of their relationship i.e. a direct edge between user A and user B. Whereas global influence is the effect on user A by some other user C through user C's properties. In this respect, we present the various social network measures and their level-wise categorization.

A. Social Influence Measures

In the literature, many studies have been carried out using social theories [50] to estimate the social influence. Each theory [55] estimates the social influence using different measures and parameters of social network with respect to diverse application domains. Therefore, there are various measures present in the literature to estimate the social influence [82] that is useful for various other applications as well such as shortest path [72]. We have extended the categorization [1] of these measures in three levels: local level, global level and network level as shown in fig 2.

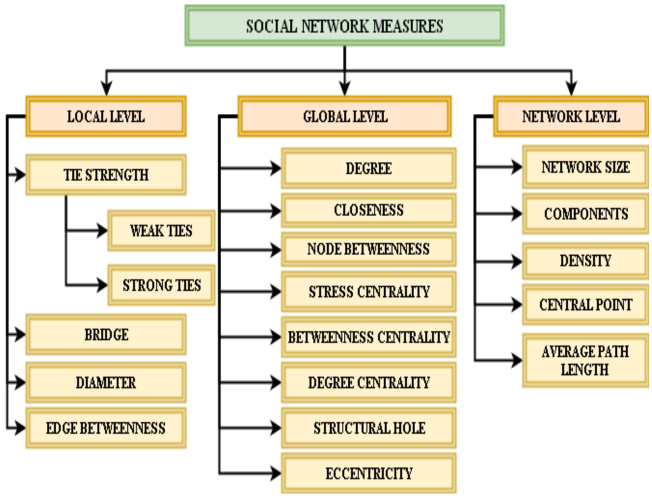


Fig. 2. Level-wise categorization of Social influence measures

2.1. Social influence measures

In the literature, many studies have been carried out using social theories [50] to estimate the social influence. Each theory [55] estimates the social influence using different measures and parameters of social network with respect to diverse application domains. Therefore, there are various measures present in the literature to estimate the social influence [82]. We have extended the categorization [1] of these measures in three levels: local level, global level and network level as shown in fig 2.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Local view concerns with the position and value of the edge in the network. It tells how powerfully and strongly any user-pair provides their contribution to the network i.e. how probable this connection is to capture information, how trouble-free it is for a user-pair to manage and control the information. Various measures used at local level are as follow:

- Tie strength: Estimation of the overlapping neighbourhood between user A and user B is considered as the tie strength of the user-pair AB. Based on the tie strength value, it is further classified into two categories: weak ties and strong ties. If the common neighbourhood of user A and user B is small then the connection present between the user-pair AB is known as weak ties. If the common neighbourhood of user A and user B is big then the interpersonal connection present between the user-pair AB is considered as strong ties. However, removal of this user-pair does not divide the network into disconnected components.
- Bridge: Bridge is an edge in the network, which plays a significant role as an only mediator or connector for two users or components of the network to exchange the information between them. As a result, non-existence of this edge will divide the network into more than one connected components.
- Diameter: The diameter is the longest shortest path distance i.e. minimum number of intermediate users present between any user-pair of the network.
- Edge betweenness: Edge betweenness measures the centrality or power of an edge in the network. It computes and identifies which edges are the most central to the network. The edges with high betweenness hold the maximum information and become more influential.

Global view concerns with the position of the user in the network. It indicates how powerfully and strongly any user provides their contribution to the network [73] i.e. how probable any user is to capture information, how trouble-free it is for a user to manage and control the information, etc.

- Degree: Degree is a method of measuring user's activity i.e. how active is a user in the network. It is further classified into three categories i.e. In-degree, Out-degree and Total-degree. In-degree is the count of all edges that are pointing towards the user. Out-degree is the count of all edges that point outward from the user and Total-degree is the count of all edges incident (connected to) on a user.
- Closeness: Closeness is a method to identify user efficiency. It is a measurement of how a user is close to every other user in the network.
- Node betweenness: Node betweenness is a method to calculate the controlling power of a user. Betweenness calculate by the count of how many shortest paths between every pair of users go through the subjected user.
- Betweenness centrality: Betweenness centrality is the measure of how much control a user can employ centrally [28]. Users with high betweenness may have appreciable influence within a network by virtue of their control over data sharing between users through the subjected user.
- Stress centrality: The stress centrality of a user is the total number of shortest paths traversing through that particular user in the network without multiple edges. If a user traverses by the large number of shortest paths then it is considered as the highly stressed user.
- Structural hole: Structural hole is a user in the network, which plays an important role as an only mediator or connector for two users or components of the network to exchange the information between them. As a result, absence of such user split the network into two or more than two components. Therefore, this measure has an important place in the field of social influence.
- Degree centrality: Degree centrality of a user represents the ratio of the number of edges passing through the user and the total number of users present in the network.
- Eccentricity: The maximum distance user A has with respect to any other user B of the network measures as the eccentricity of user A.

Network level measures represent the overall composition of the network. They indicate the united effect of all users & edges and provide the information about how the network/structural properties are likely to capture information.

- Network size: Network size is the count of total number of users or the count of total number of edges in the network.
- Components: Every network is not necessarily fully connected. If a network is not completely connected, then it divides into a bunch of connected "pieces" or cluster of users connected to each other but not with the rest of the network. If we have multiple components in a network, clustering can be used to extract all the components of the network.
- Density: Density is the ratio of the count of edges with respect to the total number of possible edges in the network.
- Central point: If eccentricity of a user is equivalent to the radius of the network i.e. minimum eccentricity from all the vertices of the network, then it is considered as the central point of the graph.
- Average path length: Computation of the effectiveness of the information transfer within network is average path length i.e. ratio of sum of all user-pair's shortest distance with respect to count of total number of user-pair in the network.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

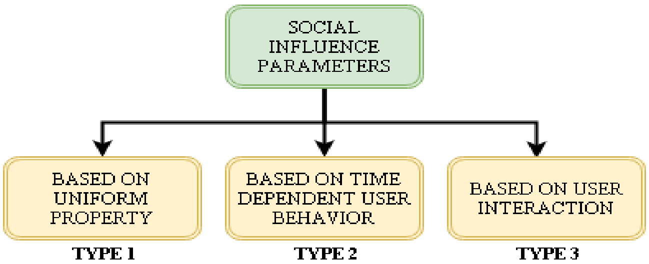
III. INFLUENCE PARAMETERS IN SOCIAL NETWORK

Numerous applications have been fabricated using the hidden details of social influence among people, such as marketing, prediction, recommendations etc. For this, many studies and theories have been projecting to measure the impact of the social influence over a large variety of datasets using various parameters of network. However, there is no extensive attempt available in the literature that focuses on dataset/domain specific social influence analysis with respect to diversity of parameters. Different social influence algorithms utilize diverse set of parameter. However, not all dataset contains all type of parameters. This creates the scope of research to perform the dataset/domain or user's specific parameter categorization [55]. Therefore, we need a framework to segregate variety of parameters available based on datasets specific user characteristics. In this respect, we categorized the parameters in the three parts: Type 1, Type 2 and Type 3 as shown in fig 3.

Fig. 3. Social influence parameter categories

A. Type 1 Parameters

In the network, every edge between user-pair AB is assigned with some uniform properties to the user-pair AB as well as to the edge itself. This is the first parameter category i.e. Type 1 which is based on uniform properties of users and edges. These uniform properties based Type 1 parameters of



an individual user-pair are accountable to create or update relationship between other user-pairs of the network. Like, creation of an edge between two persons whose opinions are of the same mind (is known as process of selection) or alteration in the opinion of a person to have the same opinion with one of their neighbors (is known as process of influence). For example, two LinkedIn users have some similar profile information i.e. school, college or company details. Based on this type of parametric information we can infer the influential behavior of these users with respect to each other. Every user v at time t has an n-dimensional vector v(t), where the i^{th} coordinate of v(t) describes the i^{th} parameter value of the user v. Digg, Flixster [13], Twitter dataset [14], [15], [17], [40], [54], Homogeneous network -> coauthor network [2], [9], [14], Linkedin [3], Flickr, Gowalla, Weibo [9], Team fortress2 game, Infectious exhibition [14], Facebook [3], [11], [19], Microsoft Web search engine [8], Messenger Instant messaging network [8], Paper citation n/w, film-director-actor-writer n/w [2], Social news aggregation Web-site [10], Email, PGP, Blog Dataset [19], DBLP bibliography data [12], Amazon product co-purchasing network [12], Opinion network e.g. election data [5], social network VK4 [43], Argus smart-phone app by Azumio tracking dataset [46], Dutch mobile telecommunications operator dataset [56], SNARE Dataset [68], Pew Research Center's Internet and American Life Project [24], School year-group networks in our study [38] are some datasets that have type 1 parameters. Table I shows the Type 1 parameters with respect to some datasets.

B. Type 2 Parameters

Type 2 parameters are those parameters which are based on changes in the social activity i.e. user behavior. These parameters are time dependent. Therefore, Type 2 parameters are observed from user's historical action log investigation i.e. how social events progress in the perspective of the network with respect to time. For example, two author who are co-author of the same paper have performed some actions i.e. total count of new links in the latest three time stamps, Time stamp information of each interaction of co-authored papers, count of groups author join in latest three time stamp etc. Using such action information, we can infer that if one author changes their research area than there will be high probability of change in research area by their peer co-author. These time-varying parameters supply the great estimation of social influence. Flicker [4], [6], [9], Twitter [6], [14], [15], [16], [17], [59], Wikipedia [7], Messenger Instant messaging n/w [8], Co-Author [9], [14], Dutch mobile telecommunications operator dataset [56], Gowalla, Weibo [9], Team fortress2 game, Infectious exhibition [14], Arnetminer [6], Digg, Flixster [13] are some datasets that have type 2 parameters. Table II gives the details of Type 2 parameters with respect to some datasets.

C. Type 3 Parameters

In addition to the Type 1 and Type 2 parameters, influence can be also enlightening by the interactions between users. Therefore, Type 3 parameters are based on the user interactions. Typically, on-line communities hold auxiliary interaction information about users. For example, a twitter user has a Wall page, where his/her contacts can discuss certain topics using tweets and re-tweets. Based on these tweets and re-tweets on the user's Wall, it can deduce which contacts are close friends and which are contacts only. On whole, we can also use follower and following information of a user on Twitter to figure out the intensity of a connection and influence power of the user in the network.

Wikipedia, Live Journal [7], LinkedIn [3], [18], Facebook [3], [18], Twitter [6], [16], [18], [59], DBLP bibliography data [12], [30], Amazon product co-purchasing n/w [12], Flickr [6] [18], Arnetminer [6] are some example datasets with respect to type 3 parameters. Table III shows examples of the Type 3 parameter.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Table I. List of Type 1 parameters based on uniform property

Name of Datasets	Parameter List				
Twitter dataset [14][15][17][40] [54]	User propagation weight	Content interestingness	Decay factor	Number of friends	User propagation rate
	Semantics of the relationship	Length of shortest path b/w users	No. of direct/indirect paths connecting two nodes	Number of mentions	Number of re-tweets
	Number of tweets per unique user for different topics	Avg Tweet rate	Percentages of urls in tweets for different topics	Quantity of a user's followers	Number of tweets
Linkedin[3]	Same school	Same company	Same geographical region	Same industry	Same job title
	Logarithm of normalized counts of common groups b/w user i&j	Logarithm of normalized count of common connections b/w users	Same function area	-	-
Flickr, Gowalla, Weibo [9]	Whether node v is an opinion leader or not	Count of new links in the latest three time stamps	Node v is a structural hole spanner or not	No. of groups user joined	No of friend
	No. of groups joined in recent 3 time stamp	No. of common neighbor between two users	Tie b/w two users strong or weak	-	-
Facebook [3][11] [19]	Relationship status	Logarithm of normalized counts of common groups user i &j joined	Logarithm of normalized counts of common friends b/w user i&j	Genders-> male, female	Age range
School year-group networks [38]	Same class	Same neighborhood(SES)	Same elementary school	Same sex	academic achievement (GPA)
	AlterSES, GPA, Female, Minority	Similarity SES, GPA, Female, Minority	Ego of SES, GPA, Female, Minority	-	-

Table II. List of Type 2 parameters based on time dependent user behavior

Name of Dataset	Parameter List			
Flicker [4][6][9]	Time when social tie created between user x and user z	Time when action A performed by user i.	Whether a user includes a photo to his favourite list	Count of new links in the latest three time stamps
Twitter[6][14][15][16][17][59]	Time t when a user talks about the matter "X" on his micro blogs (tweets).	Time when declared friendship was recorded	Time stamp information of each tweet	Time-stamp of a message

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

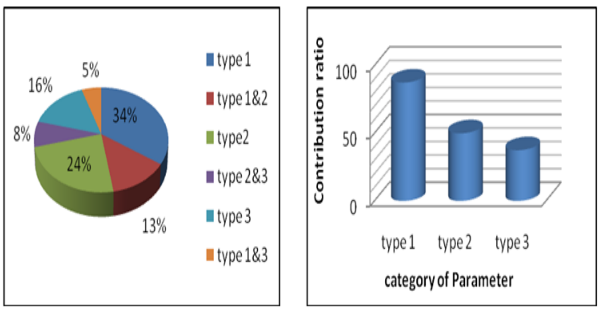
	No of groups user join in latest three time stamp	Retweet percentages for critical time stamp	Url based tweet percentages for critical time stamp	User's opinion at time *t*
Wikipedia[7]	A user's action vector v(t) to specify the total no. of times that he or she has updated every article up to that point in time	Time stamp of 1st assembly for two users x & y as the time instance at which one of them first makes a post on the user discussion page of the other	Time of updating the discussion page by editor allied with a particular user	Time when user talk to a random person
Co-Author[9][14]	Count of new links in the latest three time stamps	No. of groups user join in recent three time stamp	Time when the declared friendship was recorded b/w authors	Timestamp information of each interaction of co-authored papers

Table III. List of Type 3 parameters based on user interaction

Name of Dataset	**Parameter List**				
Wikipedia, Live Journal [7]	Editing articles by an editor	Updating the discussion page related with a specific user	Pairs of editors who have communicated with one another	Sample from user's own history	Sample from a neighbour's history
	Talk to a random person	Talk to someone based on a common activity	Sample from the world's history	Start a new activity	-
LinkedIn [3][18]	User j included in user i's online LinkedIn address book	Written a recommendation for user j by user i	User i viewed user j's profile	Establishe d a connection between users	Profession al connections in LinkedIn
	Contents are re-shared by users	Users comment on these contents	-	-	-
Facebook [3][18]	User i has posted something on user j's wall	User i tagged user j in a picture	"friending" connections in Facebook	-	-
Twitter[6][16][18][59]	User talks the matter "X" on his microblogs (tweets)	"following" connections in a Twitter	Distribution of user activity across the day	Opinion sequence on a specific topic	-

IV. ANALYSIS OF RESULTS

Meticulous study was act upon in order to obtain the depth understanding of the field i.e. social influence. We have studied 80 papers in the field of social influence. Based on the intensive study we categorized the parameters in three types and conducted the type dominance analysis on these three types of parameters. Fig. 4(a) shows research work contribution analysis of parameter types in the field of social influence. It can be inferred from the figure 4(a) that Type 1 are majorly used parameters in the past studies. It also shows lesser amount of research has been done using other parameter types i.e. Type 2, Type 3 and using the combination of different parameter types.



(a) (b)

Fig. 4. (a) Research Work dominance ratio of Parameter types (b) Dataset distribution for parameter types

Therefore, this analysis provides scope of research in the domain of social influence to discover a number of new outcomes using different type of parameters.

Similarly, fig 4(b) shows the dataset distribution with respect to different type of parameters. It shows which type of parameter is commonly obtainable in different datasets. Type 3 parameters are rare parameters. Approximately 85% of dataset contains Type 1 parameter and nearly 50% dataset have user action specific parameters.

Therefore, analysis shows that small amount of research has been using Type 2 and Type 3 parameters i.e. time dependent user behavior and user interaction respectively. These parameters have strong potential to analyze influence such as, user interaction activities i.e. "following" connection in a Twitter for user j by user i and time dependent user behaviour i.e. time t when a user i discussed the topic "X" on his micro blogs (tweets) are the significant information for influence analysis. This type of information plays an important role in influence analysis, which has not received enough attention so far. Results indicates that contribution ratio of parameters other than Type 1 is very low whereas some datasets may only contribute for social influence through Type 2 and Type 3 only. Therefore, this provides scope of study to researchers and practitioners who intend to develop new models in the area of social influence by using various types of parameters and measures.

In this work, we analyzed the impact of different parameters with respect to three different social networks i.e. Facebook, Twitter and LinkedIn the field of diffusion. The description of dataset and parameters is given in the table I, II,III and IV.

Table IV. Description of dataset

Name of Dataset	No. of Nodes	No. of Edges
Facebook	4039	88234
Twitter	8130	176814
LinkedIn	6208	63286

Diffusion is the process to quantify that how much information can be propagated through k number of seed nodes. In this work, we analyze the impact of parameter in diffusion models using seed node values k ranging from 10 to 200.

For this, we used the Independent cascade model [81] as diffusion model and identified the impact of all three type of parameters over the three different datasets i.e. Facebook, Twitter and LinkedIn as shown in figure 5, figure 6 and figure 7 respectively. Percentage accuracy is the evaluative parameter used in this work for performance analysis of all three types of parameters over three datasets. Percentage accuracy is calculated as follow:

Accuracy (in percentage) = $((R_E - R_o) / R_E) * 100$

Where, R_E is the expected propagation in the network and R_o is the obtained propagation in the network. Therefore, we analyzed that up to how much extend parameters contribute in achieving desired outputs.

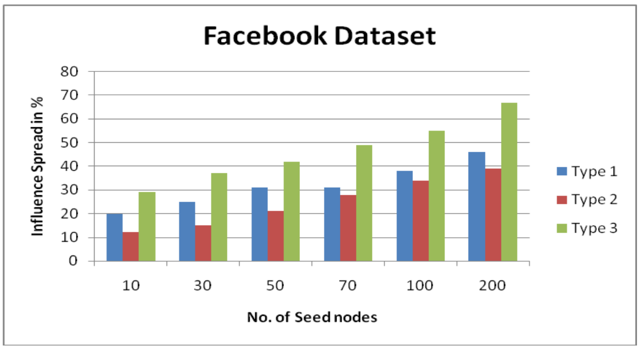


Fig. 5. Performance evaluation of parameters on Facebook dataset

Figure 5 illustrates the impact of parameters over the Facebook dataset. It can be depicted through the figure that type 3 performed better that type 2 and type 1. Nature of the facebook network is more interaction based like commenting on post, writing on friend's wall etc. Whereas, twitter is the social network where people can express their thoughts and views in the form of tweets and followers can read or re-tweet on that. Therefore, in twitter dataset type 2 performed well i.e. time dependent user behavior.

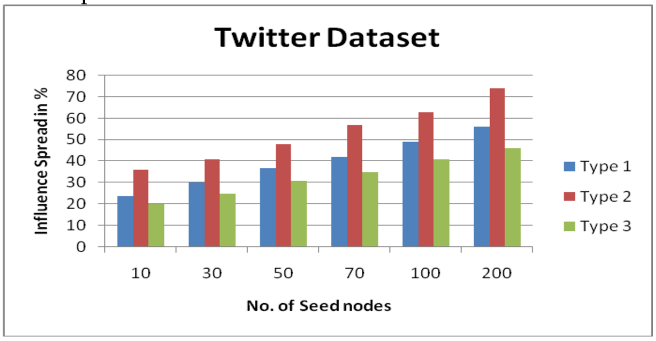


Fig. 6. Performance evaluation of parameters on Twitter dataset

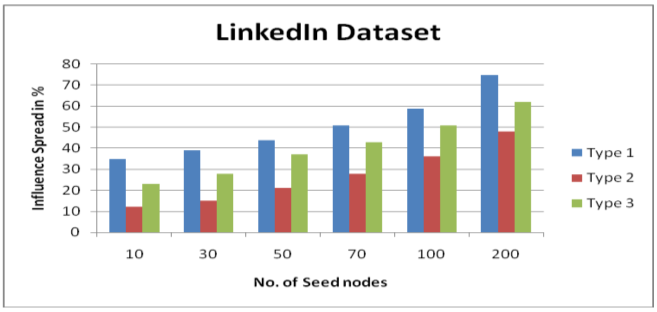


Fig. 7. Performance evaluation of parameters on LinkedIn dataset

Similarly, linkedIn illustrated different result as it performed better with respect to type 1 i.e. user based uniform properties. LinkedIn is the professional platform where people can share information with each other. Therefore, in this dataset performance of uniform properties performed well such as job location, company name, current company name etc. Therefore, the estimation of information propagation is more appropriate using profile similarities information. Overall, through experimental results we explored that performance of diffusion model is dependent on the type of the dataset. Therefore, selection of appropriate parameter set according to nature of the dataset is an important aspect in area of social influence.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Exhaustive study performed in order to discover the relation between the types of parameters and the influence measures. Table V, shows component summary of the various models i.e. which parameters have been used and which measures have been applied for the influence analysis.

Table V. List of Various Models used in social influence analysis

Approach/Model	Influence Measures Used in Influence Analysis			Type of parameter Used in Influence Analysis
	Local	Global	Network	
Link prediction based on the matrix alignment [21]	Diameter, Tie Strength	Degree, Closeness, Node Betweenness	-	Type 1
An incremental and practical classification method using Bayesian framework[20]	-	Degree, Degree Centrality	Network Size, Density, Central Point	Type 1 & Type 2
Influence Rank[22]	-	Degree, Closeness	Central Point	Type 1
Bipartite Markov Random Field[23]	-	Degree, Node Betweenness, Stress Centrality	Network Size, Components, Average Path Length	Type 2
Social dimension based approach with sparse social dimensions[25]	Tie Strength, Diameter, Edge Betweenness	Degree	Components	Type 1
Quasi-social networks extraction[27]	-	Degree, Node Betweenness, Stress Centrality, Closeness	-	Type 1
Regression model[29]	-	Eccentricity, Betweenness Centrality	-	Type 1
Approximation algorithm based on greedy strategy[30]	Tie Strength	Degree	-	Type 1 & Type 3
Frequent pattern mining approach[31]	-	Degree, Betweenness Centrality	-	Type 2
Longitudinal social network analysis[32]	Tie Strength	Degree	Components	Type 1
Longitudinal statistical models[35]	Tie Strength	Degree, Node Betweenness	Components	Type 1
Linear Regression[36]	Tie Strength	Degree, Node Betweenness	-	Type 1
Network Traffic Routing Models[42]	-	Betweenness Centrality, Eccentricity	Components, Central Point, Average Path Length	-
Hierarchical Linear Regression [49]	-	Closeness, Node Betweenness	-	Type 1
Constant Time Optimization Approach [52]	-	Degree, Structural Hole	Components, Central Point	-
Randomizes Experiments[53]	Tie Strength	Degree, Node Betweenness, Stress Centrality	-	Type1 & Type 3
National Statistics Socio- economic Classification (NS-SEC)[58]	-	Closeness, Node Betweenness, Stress Centrality, Betweenness Centrality, Eccentricity	-	Type1 & Type 3
HEALER & DOSIM[60]	Tie Strength	Degree	-	Type 2
Influential Checkpoints (IC) framework[62]	-	Degree	Network Size, Central Point, Density	Type 2 &Type 3
Dynamic Independent Cascade model [63]	-	Degree, Closeness, Node Betweenness	-	Type 1
Stochastic Dynamic Programming using Lawrence Degree Heuristic and Adaptive Hill- Climbing [64]	Edge Betweenness	Degree, Eccentricity	Network Size, Components	-
Competitive independent cascade (CIC) model[66]	-	Degree, Stress Centrality, Betweenness Centrality, Degree Centrality	-	Type 1
Novel Multidimensional Models[71]	Tie Strength, Bridge, Diameter	Degree, Node Betweenness	Average Path Length	-
Model of strategic formation of influence networks with two strategic agents [74]	Tie strength	Degree, Node Betweenness, Stress Centrality	Components	Type 1
A stochastic multi-agent model for opinion dynamics[75]	Tie strength, Diameter, Edge Betweenness	Degree, Closeness, Node Betweenness, Eceentricity	-	Type 2

These results further establish that most of the existing models have also been developed either using Type 1 parameters or in reference to Type 1 parameters. The reason may also be the ease of availability of type 1 parameter in datasets. There is no model, which utilizes only Type 3 parameters for measuring the social influence. Models use

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Type 3 parameters either along with Type 1 or Type 2 parameters. It's a very unexpected observation as the social networks exist because of user interactions and current social influence models barely use such parameters for measuring the social influence. Since social web has huge collection of interaction data, there is lot of scope to develop the models for measuring the social influence using user interaction parameters.

V. RELATED WORK

Social networks are being use by millions of persons who share millions of posts and messages on various topics. The information shared by these social networks provides valuable information in order to foretell or discover actions and events in the real world. In order to predict and understand the diffusion behavior in detail, it is essential to discover and measure the spread of the information. Therefore, numerous varieties of methods [79] and computational models [37] have been proposed and several researchers have attempted to determine parameters occupied in the influence analysis. Like, Sancheng et al. [61] presented a study of social influence analysis and the characteristics of the social influence. They also described the structural design of social influence analysis from the perspective of big data. Lee et al. [82] performed a comparitive study on 512 undergraduate students to analyse the similarities and dissimilarities between the user's traits and technological habits. Mariam Adedoyin-Olowe et al. [69] discussed the various data mining, which can be applied for social influence analysis rather than data pre-processing, data analysis and data interpretation processes only. From the past studies, we scrutinize that researchers used a variety of parameters based on various datasets in social networks to predict or discover actions in the real world. For example, Stark et al. [38] explored relation and their impact among self-identified ethnic minority and majority scholars through longitudinal social network analysis (stochastic actor-oriented models) with the data of 1175 scholars (aged 13). Through their experiment, they depict that those scholars who are good in studies strongly willing for friends with high grades. However, those scholars with lower grades show extra social influence on their friends to regulate their grades and do well in studies. Similarly, Cascio et al. [51] described through experiments that Socioeconomic status (SES) is one key cultural aspect that influences susceptibility to social cues, with those from lower SES backgrounds tending toward greater interdependence and those from higher SES backgrounds tending toward greater independence. Albi et al. [39] surveyed various latest improvements in the statistical modeling of opinion-based applications. Author used partial differential equations of kinetic type in interacting multi-agent systems to model the solutions for opinion-based applications. They also discussed the optimal management of opinion formation in the network and influence impacts of added social aspects i.e. confidence and count of links in social networks. Paul et al. [34] discovered the usefulness of social influence concept in the field of virtual world. In their study, they investigated the foot-in-the-door effect and door-in-the-face effect in the virtual world. In addition, a number of researchers are applying their knowledge towards the maximization of influence [52], [62], [63], [70]. As Influence maximization is a critical research topic intended to signify the status of a social network. Tsai et al. [41] intended their research towards the Influence maximization, which is one of the significant research subjects in present situation of the social networks. They proposed an algorithm to discover a seed node set, which will have a highest influence towards a propagation model. Through their experimental outcomes, they showed that the proposed algorithm gives improved outcomes than simple Genetic Algorithm. Likewise, Loewenstein et al. [44] surveyed the role of surprise in social influence. Surprising people can provide an opportunity to influence them i.e. surprise not only impact individually on beliefs and attitudes but also put impact on the content of culture. One such recipe discussed in their work i.e. the repetition-break plot structure, to explore the psychological and social possibilities of examining surprise. Andrighetto et al. [45] the author explored whether self-objectification triggered by doing peculiar work actions would raise people's conforming behavior or not. Author depicted the future work in the field of social influence in their paper i.e. which features of a node or the network are more relevant in attaining high influence impact on other nodes in the network. Soryani et al. [76] attempted to analyse several accomplished research on data and records of Social Networks. In their analysis, they categorized the social network information domain in seventeen research subtopics or subareas. Althoff et al. [46] author presented numerous usual experiments and observational studies regarding how an online social network parameter influence offline and online user actions by a big activity tracking dataset. Fan et al. [47] presents a system to determine influence behavior i.e. OCTOPUS, which discovers important insights using topic-aware social influence analysis services. OCTOPUS has some unique features i.e. keyword-based influential user discovery, personalized influential keywords suggestion and interactive influential path exploration.

Similarly, Vaswani et al. [48] author proposed a pair wise-influence semi-bandit feedback model and develop a LinUCB-based bandit algorithm. This algorithm not only makes their structure agnostic to the fundamental diffusion model, but also statistically capable to learn from data. Yao et al. [57] contributed towards influence analysis by reducing the spread of a present unwanted item by restricting a number of nodes from participating in a network from a topic modeling perspective. Therefore, various studies and experiments have been done in the past to discover the various causes, impact and spread behavior of social influence.

VI. CONCLUSION

Social network is an essential characteristic for numerous of today's computing applications. Countless prospering internet sites as well as applications utilize social networking attributes to attract their clients. They encourage their clients to act together through social network like post news/information, discuss or extend their thoughts through influence. Researchers from various application domains are trying to estimate the social influence power efficiently. As a result, effective and efficient social influence analysis is the necessity of social network. In this paper, we focused on the conceptual categorization of the measures for influence analysis and categorized the types of parameter with respect to various social networks. These parameters used by variety of techniques and algorithms for social influence estimation. Through our study, it is inferred that social influence analysis is not possible using same set of parameters for all type of datasets. Therefore, selection of suitable set of parameters for influence analysis plays an important role in the field of social networks, as parameters are dataset/application dependant. Along with this parametric selection problem, appropriate measure selection also plays a significant role in social network influence analysis.

Therefore, this generates various open questions in the field of social influence: are we using the right measures to estimate the influence? Is appropriate parameter selection a dataset/application based process or not? Do we have common or unique set of parameter and measures for influence analysis w.r.t. every dataset and application domain? As a result, this work exposed the various open research gaps in social influence analysis. Therefore, in the future this crucial and provocative research can be useful in discovery of efficient and quantifiable social influence methods and models to enable various applications more effective.

REFERENCES

1. Sun, Jimeng, and Jie Tang. "A survey of models and algorithms for social influence analysis." Social network data analytics (2011); 177-214.
2. Tang, Jie, et al. "Social influence analysis in large-scale networks." Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2009.
3. Xiang, Rongjing, Jennifer Neville, and Monica Rogati. "Modeling relationship strength in online social networks." Proceedings of the 19th international conference on World Wide Web. ACM, 2010.
4. Goyal, Amit, Francesco Bonchi, and Laks VS Lakshmanan. "Learning influence probabilities in social networks." Proceedings of the third ACM international conference on Web search and data mining. ACM, 2010.
5. Holme, Petter, and Mark EJ Newman. "Nonequilibrium phase transition in the coevolution of networks and opinions." Physical Review E 74.5 (2006): 056108.
6. Tan, Chenhao, et al. "Social action tracking via noise tolerant time-varying factor graphs." Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2010.
7. Crandall, David, et al. "Feedback effects between similarity and social influence in online communities." Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2008.
8. Singla, Parag, and Matthew Richardson. "Yes, there is a correlation:-from social networks to personal behavior on the web." Proceedings of the 17th international conference on World Wide Web. ACM, 2008.
9. Tang, Jie, Sen Wu, and Jimeng Sun. "Confluence: Conformity influence in large social networks." Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2013.
10. Muchnik, Lev, Sinan Aral, and Sean J. Taylor. "Social influence bias: A randomized experiment." Science 341.6146 (2013): 647-651.
11. Aral, Sinan, and Dylan Walker. "Identifying influential and susceptible members of social networks." Science 337.6092 (2012): 337-341.
12. Zhou, Yang, and Ling Liu. "Social influence based clustering of heterogeneous information networks." Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2013.
13. Barbieri, Nicola, Francesco Bonchi, and Giuseppe Manco. "Topic-aware social influence propagation models." Knowledge and information systems37.3 (2013): 555-584.
14. Zuo, Xiang, et al. "The power of indirect ties." Computer Communications 73 (2016): 188-199.
15. Kim, Hyoungshick, Konstantin Beznosov, and Eiko Yoneki. "Finding influential neighbors to maximize information diffusion in twitter." Proceedings of the 23rd International Conference on World Wide Web. ACM, 2014.
16. Guille, Adrien, et al. "Information diffusion in online social networks: A survey." ACM SIGMOD Record 42.2 (2013): 17-28.
17. Asur, Sitaram, and Bernardo A. Huberman. "Predicting the future with social media." Proceedings of the 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology-Volume 01. IEEE Computer Society, 2010.
18. Wang, Zhi, et al. "Social media recommendation." Social Media Retrieval. Springer London, 2013. 23-42.
19. Kim, Hyoungshick, and Eiko Yoneki. "Influential neighbours selection for information diffusion in online social networks." Computer Communications and Networks (ICCCN), 2012 21st International Conference on. IEEE, 2012.
20. Boutet, Antoine, Hyoungshick Kim, and Eiko Yoneki. "What's in Twitter, I know what parties are popular and who you are supporting now!." Social Network Analysis and Mining 3.4 (2013): 1379-1391.
21. Scripps, Jerry, et al. "A matrix alignment approach for link prediction." Pattern Recognition, 2008. ICPR 2008. 19th International Conference on. IEEE, (2008): 1-4.
22. Song, Xiaodan, et al. "Identifying opinion leaders in the blogosphere." Proceedings of the sixteenth ACM conference on Conference on information and knowledge management. ACM, 2007.
23. Shi, Xiaolin, et al. "User grouping behavior in online forums." Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2009.
24. Peng, Kun, and Bolin Cao. "Being online daters or not: Effects of individual factors, peers influence, and social reality." (2017).
25. Tang, Lei, and Huan Liu. "Scalable learning of collective behavior based on sparse social dimensions." Proceedings of the 18th ACM conference on Information and knowledge management. ACM, 2009.
26. Rosenquist, J. Niels, et al. "The spread of alcohol consumption behavior in a large social network." Annals of internal medicine 152.7 (2010): 426-433.
27. Provost, Foster, et al. "Audience selection for on-line brand advertising: privacy-friendly social network targeting." Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2009.
28. Newman, Mark EJ. "A measure of betweenness centrality based on random walks." Social networks 27.1 (2005): 39-54.
29. Mednick, Sara C., Nicholas A. Christakis, and James H. Fowler. "The spread of sleep loss influences drug use in adolescent social networks." PloS one5.3 (2010): e9775.
30. Kempe, David, Jon Kleinberg, and Éva Tardos. "Maximizing the spread of influence through a social network." Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2003.
31. Goyal, Amit, Francesco Bonchi, and Laks VS Lakshmanan. "Discovering leaders from community actions." Proceedings of the 17th ACM conference on Information and knowledge management. ACM, 2008.
32. Fowler, James H., and Nicholas A. Christakis. "Dynamic spread of happiness in a large social network: longitudinal analysis over 20 years in the Framingham Heart Study." Bmj 337 (2008): a2338.
33. Elias, Steven M., and Anthony R. Pratkanis. "Teaching social influence: Demonstrations and exercises from the discipline of social psychology." Social Influence 1.2 (2006): 147-162.
34. Eastwick, Paul W., and Wendi L. Gardner. "Is it a game? Evidence for social influence in the virtual world." Social Influence 4.1 (2009): 18-32.
35. Christakis, Nicholas A., and James H. Fowler. "The spread of obesity in a large social network over 32 years." New England journal of medicine 357.4 (2007): 370-379.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

36. Cacioppo, John T., James H. Fowler, and Nicholas A. Christakis. “Alone in the crowd: the structure and spread of loneliness in a large social network.” Journal of personality and social psychology 97.6 (2009): 977.
37. Tang, Jie. "Computational Models for Social Network Analysis: A Brief Survey." Proceedings of the 26th International Conference on World Wide Web Companion. International World Wide Web Conferences Steering Committee, 2017.
38. Stark, Tobias H., Lars Leszczensky, and Sebastian Pink. "Are there differences in ethnic majority and minority adolescents’ friendships preferences and social influence with regard to their academic achievement?.” Zeitschrift für Erziehungswissenschaft 20.3 (2017): 475-498.
39. Albi, Giacomo, et al. "Recent advances in opinion modeling: control and social influence." Active Particles, Volume 1. Birkhäuser, Cham, 2017. 49-98.
40. Yin, Jessie. "A Survey on Influence Analysis in Social Networks." (2011).
41. Tsai, Chun-Wei, Yo-Chung Yang, and Ming-Chao Chiang. "A genetic newgreedy algorithm for influence maximization in social network." Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on. IEEE, 2015.
42. Brown, Philip N., and Jason R. Marden. "Studies on robust social influence mechanisms: Incentives for efficient network routing in uncertain settings." IEEE Control Systems 37.1 (2017): 98-115.
43. Ivanov, Sergei, et al. "Content Recommendation for Viral Social Influence." Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2017.
44. Loewenstein, Jeffrey. "Surprise, Recipes for Surprise, and Social Influence." Topics in cognitive science (2018).
45. Andrighetto, Luca, et al. "Objectified conformity: working self-objectification increases conforming behavior." Social Influence (2018): 1-13.
46. Althoff, Tim, Pranav Jindal, and Jure Leskovec. "Online actions with offline impact: How online social networks influence online and offline user behavior." Proceedings of the Tenth ACM International Conference on Web Search and Data Mining. ACM, 2017.
47. Fan, Ju, et al. "OCTOPUS: An Online Topic-Aware Influence Analysis System for Social Networks." ICDE, 2018.
48. Vaswani, Sharan, et al. "Model-independent online learning for influence maximization." International Conference on Machine Learning. 2017.
49. Pinciotti, Caitlin M., and Holly K. Orcutt. "Understanding gender differences in rape victim blaming: the power of social influence and just world beliefs." Journal of interpersonal violence (2017): 0886260517725736.
50. Liu, Wenlin, et al. "Social network theory." The International Encyclopedia of Media Effects (2017).
51. Cascio, Christopher N., et al. "Cultural context moderates neural pathways to social influence." Culture and Brain 5.1 (2017): 50-70.
52. Kolumbus, Yoav, and Sorin Solomon. "Phase Transition in the Maximal Influence Problem: When Do We Need Optimization?." arXiv preprint arXiv:1708.02142 (2017).
53. Taylor, Sean J., and Dean Eckles. "Randomized experiments to detect and estimate social influence in networks." arXiv preprint arXiv:1709.09636 (2017).
54. Zhang, Yuchi, Wendy W. Moe, and David A. Schweidel. "Modeling the role of message content and influencers in social media rebroadcasting." International Journal of Research in Marketing 34.1 (2017): 100-119.
55. Henry, Didier, Erick Stattner, and Martine Collard. "Social media, diffusion under influence of parameters: survey and perspectives." Procedia Computer Science 109 (2017): 376-383.
56. Risselada, Hans, Peter C. Verhoef, and Tammo HA Bijmolt. "Dynamic effects of social influence and direct marketing on the adoption of high-technology products." Journal of Marketing 78.2 (2014): 52-68.
57. Yao, Qipeng, et al. "Topic-aware social influence minimization." Proceedings of the 24th International Conference on World Wide Web. ACM, 2015.
58. Moulton, Vanessa, et al. "The influence of grandparents’ social class on children’s aspirations." British Journal of Sociology of Education 38.4 (2017): 500-517.
59. Chen, Chengyao, et al. "Exploring Interpersonal Influence by Tracking User Dynamic Interactions." IEEE Intelligent Systems 32.3 (2017): 28-35.
60. Yadav, Amulya, et al. "Influence maximization in the field: The arduous journey from emerging to deployed application."Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems. International Foundation for Autonomous Agents and Multiagent Systems, 2017.
61. Peng, Sancheng, Guojun Wang, and Dongqing Xie. "Social influence analysis in social networking big data: opportunities and challenges." IEEE Network 31.1 (2017): 11-17.
62. Wang, Yanhao, et al. "Real-time influence maximization on dynamic social streams." Proceedings of the VLDB Endowment 10.7 (2017): 805-816.
63. Tong, Guangmo, et al. "Adaptive influence maximization in dynamic social networks." IEEE/ACM Transactions on Networking (TON) 25.1 (2017): 112-125.
64. Lawrence, Trisha. "Stochastic Dynamic Programming Heuristics for Influence Maximization-Revenue Optimization." arXiv preprint arXiv:1802.10515 (2018).
65. Beck, Kenneth, and Samantha Watters. "How do significant others influence our driving? A descriptive study of ego–alter dyads in a college population." Traffic injury prevention 18.4 (2017): 381-386.
66. Yang, Dingda, et al. "Relative influence maximization in competitive social networks." Science China Information Sciences 60.10 (2017): 108101.
67. Spiegler, Thomas. "Parents’ motives for home education: The influence of methodological design and social context."International Electronic Journal of Elementary Education 3.1 (2017): 57-70.
68. Gremmen, Mariola Claudia, et al. "First selection, then influence: Developmental differences in friendship dynamics regarding academic achievement." Developmental psychology53.7 (2017): 1356.
69. Ramshree, K. S. "A Survey of Data Mining Techniques for Social Network Analysis." Software Engineering and Technology 9.1 (2017): 4-6.
70. Spiliopoulou, Myra. "Evolution in social networks: A survey."Social network data analytics. Springer, Boston, MA, 2011. 149-175.
71. Parsegov, Sergey E., et al. "Novel multidimensional models of opinion dynamics in social networks." IEEE Transactions on Automatic Control 62.5 (2017): 2270-2285.
72. Agarwal, Sakshi, and Shikha Mehta. "Approximate shortest distance computing using k-medoids clustering." Annals of Data Science 4.4, pp. 547-564, 2017.
73. AlFalahi, Kanna, Yacine Atif, and Ajith Abraham. "Models of influence in online social networks." International Journal of Intelligent Systems 29.2 (2014): 161-183.
74. Grabisch, Michel, et al. "Strategic influence in social networks." Mathematics of Operations Research (2017).
75. Bolzern, Paolo, Patrizio Colaneri, and Giuseppe De Nicolao. "Opinion influence and evolution in social networks: a Markovian agents model." arXiv preprint arXiv: 1711.09802(2017).
76. Soryani, Mohammad, and Behrooz Minaei. "Social networks research aspects: A vast and fast survey focused on the issue of privacy in social network sites." arXiv preprint arXiv: 1201.3745 (2012).
77. Bonchi, Francesco. "Influence Propagation in Social Networks: A Data Mining Perspective." IEEE Intelligent Informatics Bulletin 12.1 (2011): 8-16.
78. Cialdini, Robert B., and Noah J. Goldstein. "Social influence: Compliance and conformity." Annu. Rev. Psychol. 55 (2004): 591-621.
79. Qi, J., Monod, E., Fang, B., & Deng, S., Theories of social media: Philosophical foundations. *Engineering*, *4*(1), (2018), 94-102.
80. Gao, W., Liu, Y., Liu, Z. and Li, J., How does presence influence purchase intention in online shopping markets? An explanation based on self-determination theory. Behaviour & Information Technology, pp.1-14, (2018).
81. S. Agarwal and S. Mehta, "Social Influence Maximization Using Genetic Algorithm with Dynamic Probabilities", 2018 Eleventh International Conference on Contemporary Computing (IC3), Noida, India, pp. 1-6, 2018.
82. Wang, Yichuan, and Chiahui Yu. "Social interaction-based consumer decision-making model in social commerce: The role of word of mouth and observational learning." International Journal of Information Management 37.3 (2017): 179-189.

AUTHORS PROFILE

Sakshi Agarwal is working as Assistant Professor in Jaypee Institute of Information Technology, NOIDA, India. She completed her Master in Technology from Indraprastha Institute of Information Technology, Delhi in 2012. Her research interests are Information Retrieval, Big data analysis, Social Network Analysis, Optimization Algorithms.

Shikha Mehta is working as Associate Professor in Jaypee Institute of Information Technology, NOIDA, India. She completed her Doctor of Philosophy from Delhi University, Delhi in 2013. She is an active researcher and guiding 4 Ph.D. students. Her research interests are large scale optimization, Nature Inspired Algorithms, Soft Computing, Big data analysis, Social Network Analysis.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication