Investigating Social Network as Complex Network and Dynamics of User Activities

Hradesh Kumar
Department of Computer Science & Engineering
Krishna Institute of Engineering & Technology
Ghaziabad, India

Sanjeev Kumar Yadav
Department of Computer Science & Engineering
Krishna Institute of Engineering & Technology
Ghaziabad, India

ABSTRACT

In modern era activities of users on social network applications are growing exponentially. To understand the user behavior on online social networks, determining the parameters of activities, identifying structure of inter-relationship and exploration of dynamics are major issues of the paper. Randomness of information flow in social network is inheriting property and user response varies with time and space so these allow us to select appropriate stochastic modeling technique. Most of the events happened at the social network are random and simultaneously i.e. which post/message will be shared/visible on wall of whom and when. Random graphs and dynamic stochastic model i.e. epidemic model is observed to be very close for analyzing user activities and mechanism of information spreading.

Keywords
Complex Network, Random graph, Erdos-Renyi (ER) model, Social network

1. INTRODUCTION

Social network is a type of complex network. A social network is a set of people or groups of people having interaction among them. There are so many social networks application i.e. Facebook, MySpace, LinkedIn, Flicker, Twitter and YouTube [1]. To understand dynamics of the user activities they should know mechanism of interaction among the online social network users [2]. Online social networks are basically designed for two primary purposes; these are sharing & interaction of data and support the social activities of users. Main aim of social network analysis research is to understand the dynamics of network and its structural properties [3].Facebook is one of the most popular social networks. There are 1.44 billion monthly active users till March 2015 at the Facebook [4]. Facebook profile has so many attributes such as birth date, home town, contact information, college, employers, high school etc. The average Facebook user makes more than four attributes set as publicly [5].

Online Social networks are investigated to find the relationship between modeling phenomena and characteristics of real networks i.e. average path length, clustering coefficient and others. One common question is how to connect the local structure on phenomena with the global dynamics. Visualization methods and tools are used to analysis evidence parameter of nodes. The position of nodes are determined by the graph layout method to remove this issues but it become more complicated when number of nodes increases in the networks. To address this issue of large number of nodes community structure of the network is used. Community Structure connects the local network to the global. There is a problem related interpretation of the result. Social networks come under assortative and biological & technological

network comes under disassortative. Watts and Strogatz model [6] generates the network which has high clustering and small average distance. Barabasi and Albert model generates the network which has power low degree distribution. Wang presented a network model which works for assortative and disassortative network. A new node attached the preferentially node which is already exists in the network in the second step node select a neighbor node s of the node i with the probability

$$\frac{k_s^\beta}{\sum_{j\in\Gamma_i} k_j^\beta}$$

Where β is preferential exponent parameter, Γ_i is neighbor node set of node i.

The general problem of network is efficiency in terms of growth of networks, relevant and user specific information propagation, optimize use of hardware. So they desire to maximize the efficiency of the network. One of the idea based on entropy model proposed by the Shanon about the maximizing the information capacity of channel can be applied. There are good amount of literature on social network computation for mentioned parameters. Like biological network at macroscopic level by Lotka and scaling in organism by McDonald and phenomenon of synchronization by Strogatz on social network interaction among people by Scott exploring the interface between man and machine by Wiener.

The rest of paper is organized as follow. Section 2 describes structure of complex network; it includes some important terms related to the complex network. Section 3 describes the relation between structure and dynamics of social networks, explains some definitions and user activities related to social networks. Section 4 describes the SII model and finally we conclude the paper in Section 5.

2. STRUCTURE OF COMPLEX NETWORK

Recent development in the characterization and modeling is the structural properties of a network [7]. Complex networks treated structurally and mathematically with the help of graph theory. Reachability and connectivity between nodes of a graph are quantity of interest which is well served as graph theory. The natural framework is the exact mathematical treatment of complex networks.

2.1 Motif
Motif is a meaningful number to represent pattern of interconnection in graph G either directed or undirected. It represents inverse relation to random graph. The concept of motifs was proposed by Alon and coworkers [7].

2.2 Community Structure

It is very special and empowers full form of social network where a group of individual and this group is treated as a tight couple node for analyzer purpose. Mathematically it is more useful to measure cohesiveness of communities and further individual of sub graph based on degree of cohesiveness and finally it helps in modeling of adaptive networks. Topology helps to understand global properties of social system and model as graph whose node represent the dynamical individual units and links represent the interaction between two individuals. Approximation helps to model interaction between two nodes while this interaction or links is depend on space, time and other details.

A clique is a complete graph of structure. All nodes are adjacent to each other. 2-clique graph is not complete graph but one node is intermediary to reach any node. So many schemes in nature and in technology are made by a large number of highly connected units. All features of clique are different from the lattices and random graphs (S.Boccaletti, V.Latora, Y.Moreno, 2006).

2.3 Small World Network

Fast and speedily communication in social networks application is most desirable property so mathematically finding structure in real networks like links so different areas of networks can be connected fastly. Assuming network as hypercube lattice of N dimension so commonly average number of nodes need to be passed to be reach any chosen node is found growth with $S^{1/N}$ where S is network size. There is a small path between any two nodes. Small world network is an example of random graph. Generally small world network property founds in the social network. In small world network clustering coefficient is high and path length is short. In small world network all nodes are not connected to each other but user can reach any nodes with the help of some nodes. The examples of small world networks are social network, railway network, internet network, gene network etc. Six degree of separation is a phenomenon of small world network. ENL (equal number of links) algorithm is helpful to generate the small world network. The small world property is introduced by Watts and Strogatz. In [8] author shows the difference between ENL network and WS model. The distance between two nodes in small world network is given by

$$D \; \alpha \; \log N \qquad (1)$$

Where D is the distance between two nodes and N is number of nodes in the network.

2.4 Random Graph

Random graph is the combination of graph theory and probability theory [9]. It is introduced by Erdos and Renyi in 1959. ER model shows the phase transition in the network. Many structural changes in the network when edges probability crosses to a threshold, is called as phase transition. This phenomena like as in physics changes the pressure and temperature. Facebook friends' activity follows the monotonic phase transition property. Because once Facebook profile created then friends continuously increases till a certain limit.

In scale free network degree distribution follows the power law. It applied in many real complex networks. Scale free network property proposed by Barabasi and Albert.

2.5 Characteristics Properties of Nodes

Network has n multivariate variable. Variable n includes the average degree of neighboring nodes, degree, node betweenness, the average path length and clustering coefficient.

Let N be the number of the nodes in the network. i_{th} node of network denoted by v_i. The degree of node denoted by a_i, the averaging degree of neighboring nodes denoted by D_{nn}, the node betweenness denoted by q_i, average path length denoted by P_i, clustering coefficient is denoted by h_i [10].

$$q_i = \frac{\sum_{i_s=1; i_s \neq i}^{N} \sum_{i_t=1; i_t \neq i}^{i_s - 1} \frac{g_i(i_s i_t)}{N_{i_s i_t}}}{(N-1)(N-2)/2} \qquad (2)$$

$$P_i = \sum_{i \neq j} \frac{d(v_i, v_j)}{N-1} \qquad (3)$$

$$h_i = \frac{E_i}{K_i(K_i - 1)/2} \qquad (4)$$

3. RELATION BETWEEN STRUCTURE AND DYNAMICS OF SOCIAL NETWORKS

Modeling of a system is complex even nontrivial task it includes many factors such as interconnection delay or hierarchical structure. An important example is system consisting the gene regulatory network and neurons in the brain and idea of interest is to correlate the structure properties of social network to functional properties and analyzing mechanism behind dynamics. Recently researcher's works out to relating the dynamics to specific connectivity's such as random network and small world topology.

3.1 Networks Traffic Models

There are so many communication networks like internet, body's central nervous system and cellular phone network and every network require maximizing the information exchange. Theory of traffic is based on Poisson statistics. This model analysis is based on the non Markov fractal nature [11]. There are two methods -

Fractional Brownian motion

Stable Levy motion

3.2 Modulation and Renewal Aging

The first waiting time is unlike from the consecutive comment time. It is not same that the laminar area like total period phase. Extended and short length reduced equally when waiting time is exponential. Renewal process never be aged. The weight of short time supply is reduced and longtime laminar area is increased with the help of regularization supply. Allegrini gave the technique of elderly research analysis [11].

3.3 Aging Network

Waiting time supply is a decided pointer of difficulty (Bruce j. west, 2008). This work achieved by Allegrini and variation and super figures and approach to difficulty. Elderly is due to rebirth of process and work is given by Bianco and Grigolini. The active process is liable for the irregular fluorescence; variation under super figures is slow physical process means realization not sole. Actions establish the anomalous by process and anomalous diffusion is multicasting [12].

3.4 Subordination and Complexity Matching

Difficulty matching is determined by the crucial event of perturbed and crucial event of worrying system. Allowing to demotion to the regular fluctuation dissipation crucial events are not constantly evident. Key actions performed lead role in the information transport when actions are not evident. The waged is based on the continuous time random walk. CTRW is freely interpreted in terms of Poisson and fluctuation indulgence method. The actions are non-Poisson called as non ergodic [13].

3.5 Modeling of User Activity

Fabricio Benevenuto et al. explained that the there are three type of interaction first one is visible interaction second one is invisible interaction and third one is silent interaction. In visible interaction user can write the message and get the reply. But in silent interaction user can browse the profiles and browse the pictures [14].

Weight w(m, n) of relationship between person m and person n in the i_{th} social network is calculated for each social network as

$$w_i(m,n) = \frac{Q_i(m,n)}{Q_i(m)}$$

(5)

Where $Q_i(m,n)$ represent amount of message sent by m person to n person and $Q_i(m)$ represent total number of message by m.

3.6 Measurement of User Activity

Power law applies in many natural phenomena like earthquake, precipitation etc. Power law applicable in calculating the number of website visitor, number of hyperlinks of a web page, number of friends on social network, number of activities performed by the users [15].

Power function k(x) $= \alpha x^{-\tau}$, where α and τ are positive constant. τ is called scaling exponent. We can rewrite the function by taking logarithms on both sides it look like

$$\log(k(x)) = -\tau \log(x) + \log(\alpha)$$

(6)

3.7 Correlation of User Activity

There are some users those have more than one account on different social network site. Oana Gaga et al. study about the correlating of user activity. They take data from three social network sites Yelp, Flickr and Twitter. To identify the user is same on different sites they use three technique first one is location second one is timestamp of post and third one is writing style [16]

3.8 Application of User Activity

Social networking site also use for marketing purpose. Facebook is one of the most popular marketing website. Facebook profile, Group, Fans page these all are features for marketing purpose. Indonesia is the second largest Facebook user country [17]. Facebook also has paid application for marketing.

Game is also an application of social network websites. Texas Holdem Poker is so much popular that 1.7 million daily user of this game [18]. They compare three games in their study. First one Game is hugged, second one is Fighter club and third one is Familiars.

Political parties also used social network sites to campaign. Fahri Amirullah et al. study about the one election which was held in Indonesia. Their research conducted from October 2012 to January 2013. They show the differences among the big parties those used social network for campaign and small parties did not used social network sites [19].

4. EPIDEMIC MODELS FOR USER ACTIVITY-INFORMATION PROPOGATION

There are several of epidemic model for information spreading on social network are vastly used [7, 10 and 20]

4.1 Social Network Topology Characters

1) Sheng Wen et al. shows that social network is semi directed network that means some links are directed and some links are undirected in social network.

2) In social network some users have large number of friends; those person appears in many users profile. This activity due to principle of large degree of preferential.

3) In social network weight of edge is determined by the message passing through a link. Generally it follows the Gaussian distribution.

Spreading speed and scale are affected by the topological factors. Sheng Wen derived a topology for a social network. User behaves like a node in the social network. Random variable R_i(t) denote the state of person i time t.
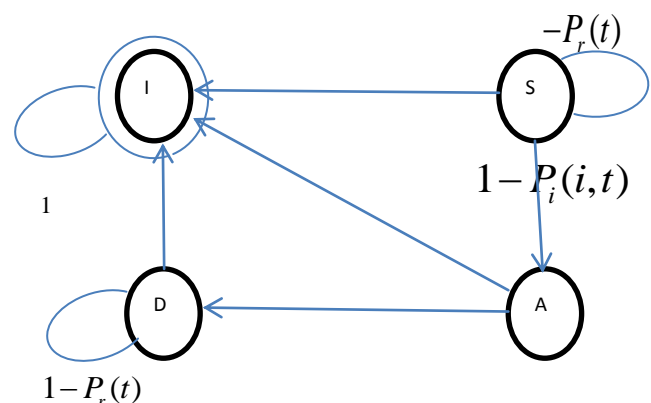

Fig. I State transition diagram

Where immunized state denoted by I, susceptible state denoted by S, active state denoted by A and dormant state denoted by D.

$$R_i(t) = \begin{cases} H & Healthy \begin{cases} S & Susceptible \\ I & Immunized \end{cases} \\ \\ IF & Infected \begin{cases} A & Active \\ D & Dormant \end{cases} \end{cases} \qquad (7)$$

A person reach at the infected state only if person at the susceptible state. Infection probability is denoted by $P_i(i,t)$. A person is infectious if and only if person is at the active state. A user is stay at the dormant state until not immunized. At the end all users reaches at the immunized state. Recovery probability is denoted by $P_r(t)$.

In SII model Sheng Wen propose a N by N square matrix

$$\begin{bmatrix} P_{11} \cdots\cdots\cdots\cdots P_{1N} \\ P_{lm} \\ P_{N1} \cdots\cdots\cdots\cdots P_{NN} \end{bmatrix} \qquad P_{lm} \in [T, F] \qquad (8)$$

Where P_{lm} represent the probability of user m reading message from user l. $P_{lm} = 0$ means no message sharing between user m and user l. Wen et al. introduce two model for properly model this process $read_i(t) = T$ means user checking a message at time t.

$$P(read_i(t)) = \begin{cases} F & otherwise \\ T & t \bmod T_i = 0 \end{cases} \qquad (9)$$

Probability of message checking is determined by T_i. Arbitrary time is the gap between last message seen time and current message seen time. Arbitrary time is denoted by the t'.

$$\begin{cases} t - T_i \le t' \le t & if\, read_i(t) = T \\ t - (t \bmod T_i) \le t' < t & otherwise \end{cases} \qquad (10)$$

Joint infected probability is calculated with the help of message from neighbors at each t'. A user can spread only worm copies to the neighbor.

Table I Study of data collection and result analysis of social networks Ref. [21, 22, 23, 24, 25, 26]

S.NO.	Author Name	Title of Paper	Data Collection technique	Result Analysis
1	Nikos Salamanos, Elli Voudigari, Theodore Papageorgiou and Michalis Vazirgiannis	Discovering Correlation between Communities and Likes in Facebook	They implemented a crawler which is developed in python to collect the data. Crawler works in Breadth first search technique	They show that the Like's categories between the different communities and whole population have high correlation.
2	Pinghui Wang and Wenbo He, Junzhou Zhao	A Tale of Three Social Networks	They collected the data from three social network foursquare, Facebook and twitter.	Activities are highly correlated among three online social networking sites by which information leakage chances increases.
3	Darren Quinn, Liming Chen, Maurice Mulvenna	Does Age Make A Difference In The Behavior Of Online Social Network Users?	A Facebook user profile was created and using the find classmate facility a list of potentially suitable friends were identified by the Facebook.	Younger user have more friends and utilize the more features frequently. Older user used 17.20% Hyper media while younger user used only 7.93%.
4	Reza Farahbakhsh, Xiao Han, Angel Cuevast and Noel Crespi	Analysis of publicly disclosed information in Facebook profiles	They developed a HTML crawler that is able to collect publicly available information from a Facebook user's profile	On average 4 attributes publicly available, men show larger public exposure as compared women, most of the user age range 18-25 out of those 50% users Birthday publicly available.
5	Atikan Muangngeon, Supaporn Erjongmanee	Analysis of Facebook Activity Usage through Network and Human Perspectives	They select population of 60,992 students of Kasetart university in the first sem. Of the academic year 2012.	Like, Chat, Profile & Wall, Photos and News Feed & Home are five significant activities in both network and human perspective. Notification is important in aspect of Network but not

			Human.
6	Annalisa Socievole, Floriano De Rango and Antonio Caputo	Wireless Contacts, Facebook Friendships and Interests: Analysis of a Multi-layer Social Network in an Academic Environment	They selected 35 students of Calabria university for experimental purpose. Experiment duration was one week. From 28 January 2014 to 5 February 2014.	Bluetooth contact network graph is close to the Facebook friendship graph. Communities vary from layer to layer still there are some similarity between Facebook friendship layer and Bluetooth contact layer.
7	Ehsan Khadangi', Ali Zarean, Alireza Bagheri, Ali Bagheri lafarabadi	Measuring Relationship Strength in Online Social Networks based on users'activities and profile information	They fill out the some questionnaires about their friends then users profile information collected by the software.	Multi-layer perceptron measure link strength with 72% accuracy and Decision Tree measure link strength with 67% accuracy. But when they use 2 class Support vector machine and Multilayer perceptron then accuracy became 87% for MLP and 84% for SVM.
8	Lili Nemec Zlatolas , Tatjana Welzer, Marjan Hericko, Marko Holbl	Privacy antecedents for SNS self-disclosure: The case of Facebook	They collected data of Facebook users from Slovenia in May 2014. This survey includes 47 questions. Total participants were 828.	Privacy awareness, privacy social norms, privacy policy, privacy control, privacy value and self-disclosure are shown by a single model.
9	Juliana, Edson, Itana, Silvia	Experimentation and Analysis of undergraduate students performance and satisfaction in a blended model of an introductory computer science and programming course	This survey was based on Likert scale. Students answered each question online.	Failure rate of classroom course is 16.14% which is the higher than the distance education. Dropout rate of distance education is 8.7%.
10	Khanh Nguyen Duc A. Tran	An Analysis of Activities in Facebook	They used data available from Max-Planck Software Institute for Software System.	Wall-posting activity follows the Power law. A node tendency to connect with the node which has similar node degree but this is not clear in the case of Facebook. Correlation between activity and user's popularity is based on message from others and wall post to others.

5. CONCLUSION

Social Networks are found to be scale free and stochastic dynamical systems. These networks are represented by random graphs where each node is entity on social networks. Structural and functional properties are found to be random in nature and in consequence dynamics of social network is uncertain and big challenge is to share and propagates information of interest in time to relevant node. Mechanism adopted by Facebook to share post and messages on storyboard is under critical observation because many times users do not get appropriate and relevance information in time. Epidemic model of information diffusion is in closely resembled and need to be customized with time and space.

In future, researcher and other organization can observe behavior and trend of user based on their online activities in real time. Strong recommendation systems would be applied in different areas. Online user(s) behavior prediction would also help us to identify unethical and harmful elements in social networks finally it would assist to make secure and prosperous environment around us.

6. REFERENCES

[1] Khanh Nguyen Duc A. Tran, " An Analysis Of Activities In Facebook", The 8th Annual IEEE Consumer Communications and Networking Conference,2011

[2] Pinghui Wang, Wenbo He, Junzhou Zhao, " A Tale Of Three Social Networks", IEEE Computer Society,2014

[3] Darren Quinn, Liming Chen, Maurice Mulvenna, " Does Age Make A Difference In The Behaviour Of Online Social Network Users?", IEEE International Conferences on Internet of Things, and Cyber, Physical and Social Computing,2011

[4] https://en.wikipedia.org/?title=Facebook

[5] Reza Farahbakhsh, Xiao Han, Angel Cuevas And Noel Crespi, " Analysis Of Publicly Disclosed Information In Facebook Profiles", IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining,2013

[6] Quiang guo, Tao zhou, "Growing scalefree small world network with tunable assortative coefficient", Elsevier,2006.

[7] S.Boccaletti, V.Latora, Y.Moreno, M.Chavez, U.Hwang, "Complex network structure and dynamics" ,Elsevier, 2006.

[8] Roberto Serra, Marco Villani, Luca Agostini, "A Small-world Network Where All Nodes Have the Same Connectivity, with Application to the Dynamics of Boolean Interacting Automata", Complex Systems Publications,2004.

[9] https://en.wikipedia.org/wiki/Random_graph

[10] Tomoyuki Yuasa,Susumu Shirayama, "A new analysis method for simulations using node categorizations", Springer,2012.

[11] Bruce j. west, Elvis l., paolo grigolini, "Maximizing information exchange between complex network", Elsevier, July 2008.

[12] Allegrini, Grigolini, Hamilton, Palatella and Raffaelli , 2002. Phys. Rev. E 65.

[13] Sokolov, I.M., 2000. Phys. Rev. A 63, 011104.

[14] Fabricio Benevenuto, Tiago Rodrigues, Meeyoung Cha, Virgilio Almeida, " Characterizing User Behavior In Online Social Networks", Acm Sigcomm Imc.,2009

[15] Aniket Mahanti, Niklas Carlsson, Anirban Mahanti, Martin Arlitt and Carey Williamson, " A Tale of the Tails: Power-Laws in Internet Measurements", IEEE Network,2013

[16] Oana Goga, Howard Lei, Sree Hari Krishnan Parthasarathi, Gerald Friedland, Robin Sommer, and Renata Teixeira, " On Exploiting Innocuous User Activity for Correlating Accounts Across Social Network Sites", International Computer Science Institute,2012.

[17] Putu Wuri Handayani And Wahyu Lisdianingrum, "Impact Analysis On Free Online Marketing Using Social Network Facebook: Case Study Smes In Indonesia", ICACSIS 2011.

[18] Ben Kirman, Shaun Lawson and Conor Linehan, "Gaming on and off the Social Graph: The Social Structure of Facebook Games", International Conference on Computational Science and Engineering, 2009.

[19] Fahri Amirullah, S.Komp, Dr. Yani Nurhadryani, S.Si., Mt, " Campaign 2.0 : An Analyze Of The Utilization Social Network Sites Of Political Parties In Indonesia", ICACSIS 2013

[20] Sheng Wen,Wei Zhou, Jun Zhang, Yang Xiang, Wanlei Zhou And Weijia Jia, " Modeling Propagation Dynamics Of Social Network Worms", IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS,2013

[21] Nikos Salamanos, Elli Voudigari, Theodore Papageorgiou and Michalis Vazirgiannis, "Discovering Correlation between Communities and Likes in Facebook", IEEE International Conference on Green Computing and Communications, Conference on Internet of Things, and Conference on Cyber, Physical and Social Computing, 2012

[22] Atikan Muangngeon, Supaporn Erjongmanee, "Analysis of Facebook Activity Usage through Network and Human Perspectives" ,IEEE, 2015

[23] Annalisa Socievole, Floriano De Rango and Antonio Caputo, "Wireless Contacts, Facebook Friendships and Interests: Analysis of a Multi-layer Social Network in an Academic Environment" ,IEEE, 2014

[24] Ehsan Khadangi, Ali Zarean, Alireza Bagheri, Ali Bagheri Lafarabadi, " Measuring Relationship Strength In Online Social Networks Based On Users' Activities And Profile Information", 3rd International Conference on Computer and Knowledge Engineering,2013

[25] Lili Nemec Zlatolas , Tatjana Welzer, Marjan Hericˇko, Marko Holbl, " Privacy Antecedents For Sns Self-Disclosure: The Case Of Facebook",Elsevier,2014.

[26] Juliana Cristina Barga, Edson Pimentel, Itana Stiubiener, Silvia Dotta, "Experimentation and analysis of undergraduate students performance and satisfaction in a blended model of an introductory computer science and programming course" ,IEEE, 2014.