

CENTRALITY MEASURE OF SOCIAL NETWORKS USING GEPHI AS A VISUALIZATION TOOL

Akhtar Ali Jalbani¹, Farah Naveen¹, Mukhtiar Memon¹, Gordhan Das Menghwar¹, Mansoor Hyder Debar¹

¹Information Technology Centre, Sindh Agriculture University, Tandojam, Pakistan
{ akjalbani, farahnaveen, mukhtiar.memon, gdas, mansoor.hyder }@sau.edu.pk

ABSTRACT: *The use of Social Network is increasing day by day because of its simplicity and way of communication to other users, hence it is important to know the users involvement in the community. The actual involvement will provide the exact interest of the person. The information extracted from social networks requires some criteria to understand the important factors of the community. In this study, educational social network communities on Facebook are examined using Gephi as visualization tool. The involvements of the persons in the social networks are measured with the help of centrality measure.*

1. INTRODUCTION

In 20th century, social scientists gave main concept of social network that connects complex relationship of social members. First Social networking site was launched in 1997. Social network is defined as: “a process of obtaining filter and mine data in various representative forms that can be achieved by collection of information by visualizing the network to obtain valuable results [1]”.

Many social networking sites such as MySpace, Facebook, Twitter provides hundreds of technologies and attractive features for users that give main interest and practices. Some of these sites help to define views, interests or activities. Social networking analysis provides new information and different tools to connect, sharing of online videos and photos. These sites are considered to be a web based service in which the individual make their profiles and making a list of friends and other users who want to connect and share information in such a way that these connection vary from site to site. The backbone of social networking site provides a wide variety of feature that includes user profiles that display list of friends.

Sunden [2] considered profiles on social network profiles as unique pages that provide information about the person connected to the social community. In this research message boards are focused where one can leave messages on others profiles message board. The other great features of social network communities are photo and video sharing that attract many people in respect to their age, nationality and educational level. Social networks are connected through friendship, common interests, financial exchange, knowledge or relationship of belief. The most popular terms of relationship are “contact”, “Friends” and “Fans”. Data in the social networking site can be explored through displaying nodes and ties, its attribute, color, sizes and other node properties. It provides a map that is made up of ties such as friendship. It also determines how much organization shows interaction and provides informal connection between users and also shows the connection and association between different users of an organization. Many social networking sites include visual representation of network that will help to understand the network data and provides the accurate results of the analysis.

User profiles are considered to be an integral part of social networking site in such a way that profiles gives the positive outcomes for their users and create a great sense of presence in which the users keep updated. The common and important goal of online social communities is forming connection

between users. These connections can be visualized using network tools such as Gephi [3].

Gephi is software tool used to analyze the social network graph. There are many different software tools that are available today for the utilization of network with their advantages but Gephi facilitates visualization of network. It is open source software and free to use and analyze the network graph. Open source software is considered to be an internal algorithm and shows direct knowledge of many other investigations. To understand and explore the graph considered to be a basic tool for a people in which user interacts and communicate with structure representation of network and colors to reveal hidden properties

In this research, we examine the Facebook as an online social network using Gephi as a visualization tool to monitor, evaluate and examine centrality of measure for the Facebook friend’s network.

2. LITERATURE REVIEW

Gross [4,5] focused on information revelation and privacy in online social network such as Facebook. Facebook allows user to make their profiles, share personal data and information to other friends. In this paper they proposed understandable pattern of information revelation and privacy implications in Facebook profiles. For the application of this research they analyzed more than 40,000 university students Facebook profiles. Another interesting criteria for evaluation of social networking site is focused in [6]. The researcher comprehensively discussed Facebook in terms of privacy and person’s relation.

Usage of online social network such as Friendster, MySpace and Facebook is reasonable but Facebook showed the rapid growth that provides interaction and communication between users efficiently [7]. The Facebook stands out the most unique and personally identifiable window platform that provides millions of information of young adults and also focuses on privacy and security patterns. New members of the Facebook seems to be less aware of privacy risk and do not have any ability to manage and maintain their profiles. Authors focused on privacy concerned “behavior” and compare member with their personal attitude that provides great information among the networks. They took age and status that represent most significant approach for determining Facebook membership because most of the Facebook members are not concerned with the visibility of their profiles. They reviewed many Facebook members’ profiles and examined the data using data mining techniques. Samuel [8] focused on the personality impression on Facebook profiles that gives important features to individual

users to know each other. Another research [9] defined the profiles relationships and number of friends in Facebook that describe the profile importance that help and encourage relationship and connection between users. They defined different theories to explore online social networks that give theoretical framework of Facebook profiles.

Joinson [10] explored the use of Facebook with gratifications of users. Almost they use 137 users' generated words to describe how Facebook can be used. In this research they define different factors by identifying unique users and gratification. These factors involve in social connection shared different identities and update their status to deal with different social investigations and privacy concern.

Raacke *et al.* [11] evaluated the use of Internet that would allow user to friendly interact with each other by using friend networking sites. They identified most promising factors that why most people are like to interact with each other by using these friend networking sites, what types of uniqueness the college students have and what type of gratification are required to users when using these sites. By using gratification method they identified valuable information from friend's behavior and their interest of sharing ideas with social community.

Hart *et al.* [12] focused on Facebook user experiences and their recent success. They used heuristic evaluation technique to addressed traditional usability methods and unique user experiences. Pempek *et al.* [13] assessed online social networking sites involves millions of young adults and also experienced how and why college students are more excited to use the social networking site. By using these popular online social networking sites 92 undergraduate's students reported the daily usage of Facebook and in order to access the use of social networking sites they proposed daily use activities. Results show that students in their daily routine use approximately 30 minutes Facebook throughout the day.

Cheung *et al.* [14] explored the importance and usage of online social networking and also focused different factors that drive the students to use Facebook. They developed a new type of interaction phenomena from little theory-driven empirical research. In addition, they examine the presence of social effects, social influence and usage of five key values of gratification paradigm when using online social networking sites.

Catanese *et al.* [15] demonstrated online social networking sites by collecting and analyzing different data of participants in the network. Data is organized on undirected graph and can be compared to factors such as degree distribution, centrality measures, scaling laws and friendship distribution. They used the web data extraction technique to extract user's relationship, in addition all information is extracted including the data that is not publically accessible. The data is graphically represented using the concept of graph theory in which users represented mainly as nodes and edges are relationship between nodes.

Shah *et al.* [16] focused on the usage of Facebook among undergraduate students. They focused on their academic performance by sharing views and comments. It helps student to get their behavioral job environment. They examined two things that social capital and academic performance helps to increase rich use-cognitive absorption aspect of using social

media instead of humiliating academic performance, second is usage of social media aids in building social capital. They also provide the effect of one's social ties between students by using these social media technologies.

Latha *et al.* [17] demonstrated primary focus of online social networking sites for example making new friends, information sharing, photos and video sharing. They introduced friend's recommendation list and proposed framework that measure the local and global network characteristics. In addition, researchers introduced graph generation model which define link prediction in Facebook and Twitter. They used Gephi tool to link these nodes in different manner.

Rieder [18] proposed method to export data from different standard file format of online social networking sites such as Facebook and proposed NetVizz API for data collection and extraction from Facebook. The same approach is adopted in [19].

3. DATA EXTRACTION USING NETVIZZ

Facebook and other online social networking sites have public Application Programming Interfaces (API's) which allow computer programs to extract data from websites. All information of the Facebook profiles can be exported in GEPHI tool using NetVizz. The NetVizz application was initiated in 2009 to conduct practical study on Facebook profiles. The data of the profiles were extracted using the NETVIZZ API. In this research, we also used the same API to extract data from Facebook profiles. The API is simply written in PHP that is maintained by Digital Media Initiative. One should use his/her own loginID and password to get the NetVizz API for his/her profile. Figure 1 shows the actual NetVizz API that can be downloadable from the Facebook profile.

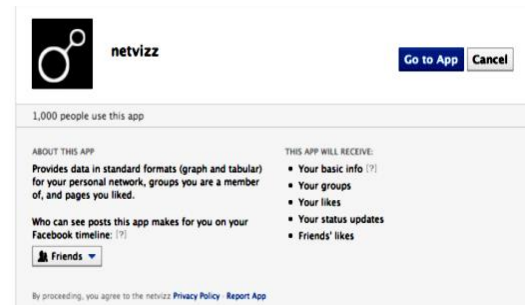


Figure 1. NetVizz API in Facebook

Facebook uses some strict policies for the application that extract personal data from Facebook profile. The construction of any Facebook API reflects these concerns in the following ways:

1. Every data profile is "signed" with the credentials and all polices are defined for each platform defines the scope of which data can be accessed.
2. Another focus of these API's to handle the user privacy settings that play important role and it defines what data can be extracted from the profile's pool.
3. Without permission no one is allowed to access different data elements so these can be implemented using various requests by the user when user first uses the application as shown in Figure 1,
4. User Interfaces are not available for these types of API's.

From above facts, it has been clearly defined that API's only extract the data in secure way without disturbing the Facebook profiles; this is due to the privacy settings of the profiles. As there is no way to visualize the data inside Facebook profiles. The application like NetVizz exports all selected data of the profiles and that data can be used to visualize using third party tool such as GEPHI. The NetVizz extracts data for three different sections of Facebook profiles, these are: personal networks, Groups and Pages. The personal networks are further divided into different friendship network and bipartite. Friendship network are simple undirected graphs: the users are considered to be nodes and friendship connects to the edges, gender, interface language and ranking based on the posts and likes of the user. Other friendship network may be considered based on locations, institutes, high schools, workplaces, and clubs. Figure 2 shows the undirected graph. The bipartite are like networks and can be generated from the users and liked entities. The networks are examined by using graph analysis toolkit such as GEPHI.

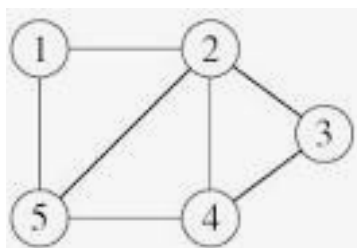


Figure 2. Undirected Graph

Groups may be considered in the similar way as friendship network, although the NetVizz has some limitations to extract the limited number of data that can be exported from any personal Facebook profile. For larger groups that data can be exported into sub groups and then can be combined into one group in the graphical tool to visualize the data. The social graphs can be identified based on the basis of interaction between group members through their posts. For example one user likes or comments on the post to the user basically they are communicating like a directed edge between two users as illustrated in Figure 3.

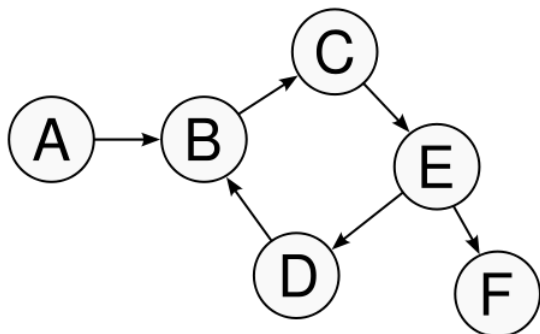


Figure 3. Directed Graph

Pages are characterized as bipartite network. Users activity can be monitored by identifying likes and posts in the directed graphs.

4. RESULTS AND DISCUSSION

In this research various Facebook profiles were analyzed; due to the privacy of the profiles; limited information was

accessed from the personal profiles such as like pages, fan pages, comments and so on. The other information related to the friends list or contact list has not been focused.

Research focused on the identification of the person's interests in the Facebook groups. The centrality measure is used identify the main interest of the user, this can only by investigating their profiles. Hence user profiles data is used for this research.

4.1 CENTRALITY MEASURE

One important feature of social network is to exchange comments on the web. The various measures such as degree centrality represents by counting links that are accidentally tied upon a node. The Facebook friends can provide the degree of centrality by measuring the closest friends or direct friends and others are the indirect friends such as friends of friends. Hence one can simply measure in terms of network terminology by counting in-degree and out-degree measurements. The equation-1 shows the mathematical way to represent for graph G: = (V,E) with n vertices, the degree of centrality for vertex is as follows:

$$C_D(v) = \frac{\text{deg}(v)}{n - 1}$$

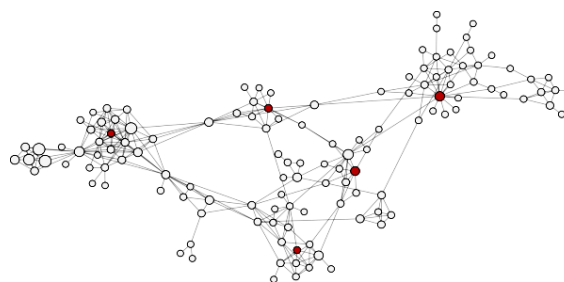


Figure 4. Nodes for Centrality Measure

From above figure-4, the degree centrality shows that how much depth or closeness of the Facebook friends has been calculated. The above scenario is implemented for the university profile where many Facebook friends and friends-of-friends have been connected. It is very difficult for someone who can figure out how many directly or indirectly connected friends have been linked with the user. The GEPHI tool is used with NetVizz to visualize links by conducting assessment in terms of in-degree and out-degree connection. The above figure shows centrality has high closeness degree. This measurement has many different meanings by exploring each pair of nodes with connected edges. The degree of centrality can be evaluated and profile users can easily locate most demanding pages. The following equation show how can we calculate the response rate for various users:

$$\text{Response Rate} = \frac{\text{No. of Likes} + \text{No. of Comment on page}}{\text{No. of impressions}} \text{----- (1)}$$

The number of impressions can be measured by counting number of times post displayed on the page wall that can be shown on fans or within the fan box. This measure gives you the approximate value and may not be 100% accurate, because that post may not been seen or read by the fans.



Figure 5. Response Rate for Comments

$$\begin{aligned} \text{Likes} &= 23 \\ \text{Share} &= 1 \\ \text{Comment on SAU official page} &= 1 \\ \text{Response rate} &= 23 + 1/1 \\ \text{Response Rate} &= 24\% \end{aligned}$$

Response rate metric was calculated on various comments on official web of SAU and found that it community is mostly involved in liking and sharing the comments but writing comments are adopting slowly and gradually as shown in Figure 5.

4.2 BETWEENESS CENTRALITY

It is another centrality measure that provides the measurement of vertex of the graph. This measure has various advantages such as identifying shortest path and fractions.

The figure-6 can be used to measure the betweenness centrality as follows:

1. Short path may be computed for each pair of vertices.
2. Each pair again determines the fraction of shortest paths that passes through the vertex.
3. All vertices are then added or sum has been computed.

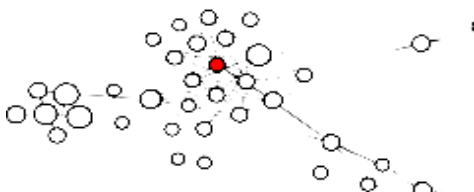


Figure 6. Betweeness Centrality Measure

Betweeness centrality was also measured based on various official web comments as shown in Figure 6. In addition, in-degree and out-degree nodes are also identified as show in Figure 6. It has been observed that community is very strong in sharing the information placed on the web instead of the individual Facebook profiles. Research study was carried out for the visualization of complex social networking site such as Facebook. This study was based on three different Facebook profiles and focuses on the visualization and network analysis techniques. Various attributes for measurement was considered and the result was discussed based on the centrality measures. All the measures were set for specific profiles those are member of education and universities fan pages, likes and favorites. It has been observed that most of the community is involved in liking and sharing the comments especially on the University official web page. It is also observed that the University

official web was not very old when this research study was conducted. Only 1675 members were liked that page. From other personal profiles of the members were not studied due to the privacy settings of their profiles.

5. SUMMARY & CONCLUSION

This research focused on the visualization of social networking site such as Facebook. Facebook is measured based on network centrality theory in which study identified homophily and heterophily networks by considering the centrality measures. The Facebook profiles are accessed by using powerful Facebook API. The data extracted from profile based on the privacy settings of the individual profiles. The study was conducted on the Sindh Agriculture University Tandojam official Facebook profile and replicated that research on two other individual Facebook personal profile groups that is related to education sector.

From this study, it has been observed that users are interested to know various activities happened on the campus. The strong and weak areas have also been identified by using centrality measure such as considering page likes, comments and sharing activities. This gives an idea that how users are using information posted on the Facebook profiles. It has been observed that page likes and sharing is strongly supported to make profiles more social instead of writing comments on the profile page.

REFERENCES

- [1] Jackson, Joyce. "Data Mining; A Conceptual Overview." Communications of the Association for Information Systems 8.1, 2002.
- [2] Sunden, J., Material Virtualities. New York:peter Lang, 2003
- [3] Bastian, M, Sebastien H, and Mathieu J. "Gephi: an open source software for exploring and manipulating networks." ICWSM 8: pages: 361-362, 2009.
- [4] Gross,R., Re-identifying facial images. Technical report, Carnegie Mellon University, Institute forSoftware Research International, 2005
- [5] Sonia, A., Is Friendster the new TIA? Online article published on <http://www.pacificresearch.org/> , January, 2004.
- [6] Ralph, G; Alessandro, A., Information revelation and privacy in online social networks, Proceedings of the 2005 ACM workshop on Privacy in the electronic society. Pages 71-80. 2005
- [7] Acquisti A. and Ralph G.. Imagined communities: Awareness, Information sharing, and privacy on the Facebook. Lecture Notes in Computer Science, Vol: 4258, pages:36-58, 2006
- [8] Samuel D.G, Sam G, Simine V; Personality Impressions Based on Facebook Profiles. Proceedings of International conference on weblogs and social media Colorado USA, 2007.
- [9] Lampe,C.,B. N.Ellision and C. Stienfield. Changes in use and perception of Facebook. Proceedings of the ACM conference on Computer supported cooperative work, 2008.
- [10] Joinson, A. N., Looking at' Looking up' or Keeping up with people? Motives and use of Facebook. Proceedings

- of the SIGCHI Conference on Human Factors in Computing Systems, 2008.
- [11] Raacke, J and J. Bonds-Raacke., MySpace and Facebook: Applying the uses and Graftification Theory To Exploring Friend- Networking Sites. *Cyberpsychol Behav* vol.11(12).pp:169-174, 2008.
- [12] Hart.Jennifer.,Charlene .R., Faisal Taher., Corina Sas and Alan Dix. Exploring the Facebook experience. A new approach to usability.Computing Department. ACM ISBN 978-1-59593-704-9: 2008.
- [13] Pempek.Tiffany.A, Yevdokiya. A. Yermolayeva and Sandra L, Calvert., College student's social networking experiences on Facebook .*Journal of applied Developmental psychology* vol. 30, pp: 227-238, 2009.
- [14] Cheung.Christy M.K., Pui-Yee-Chee., Matthew K.O. Lee. 2010. Online social networks:Why do students use Facebook? *Computers in Human behavior*. Vol. 27, PP: 1337-1343, 2010.
- [15] Catanese,S.A., D. M. Pasquale and F. Emilio. 2011. Crawling Facebook for social network analysis purpose. *Proceeding of WIMS'11 Norway*, 2011.
- [16] Shah, V., S. Subramanian.,S.Rouis and M. Limayem. A study on the impact of Facebook usage on students social capital and Academic performance. *Proceedings of AMCIS2012*, 2012.
- [17] Latha.R.Hema and K.Sathiyakumari.. Predicting link strength in online social network. *International journal of Engineering Research and application(IJERA)* Vol:2(6), pages.703-707, 2012.
- [18] Rieder, B.; Studying Facebook via Data Extraction: The Netvizz application in proceedings of the 5th Annual ACM Web Science Conference, Pages. 346-355, 2013.
- [19] Naveen F.; Interactively Visulize and explore social network using Gephi. Masters Thesis, ITC, SAU Tandojam, 2014.