Contemporary Engineering Sciences, Vol. 8, 2015, no. 3, 137 - 151
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ces.2015.515

Examination Expertise Sharing in Academic Social Networks Using Graphs: The Case of ResearchGate

Omar Alheyasat

Computer Engineering Department
Faculty of Engineering
Al-Balqa' Applied University
Al-Salt, Jordan

Abstract

ResearchGate (RG) has emerged as one of the first private social networks for researchers and developers. It adds many new features and activities to the normal Online Social Networks (OSNs). One of these services is the question/answer environment where users can share their queries and obtain answers. This service is considered very important for expertise sharing in RG. In this work, the question/answer platform of RG is studied. A web-crawler has been implemented in order to harvest the questions, authors' information and answers from RG. In this paper, the collected data has been utilized to generate a directed graph that connects users according to their questions and answers. The constructed graph is employed to extract some graph metrics, such as, node degree, centrality and connectivity. The results analysis shows that, users in RG do not care about sharing their expertise. In addition, it is observed that most of the nodes in the graph are disconnected, which means that, most of the questions have no answers. It is also observed that the number of triangles in the network graph is small which means that the network is not highly connected. Paper must have abstract.

Keywords: ResearchGate (RG), Social Network Analysis (SNA), Graph centrality, Question-answer platform

I. Introduction

With millions of subscribers, online social networks (OSNs) developed to be one of the most successful services on the Internet [1]. OSNs are general purpose networks. They can be defined as entertainment networks. Others can define them as forums. Others can use them to share its expertise, such as, photos, videos, files and discussions. Now days, many massive OSNs are online, such as, Twitter, Facebook and LinkedIn. However, the general purpose nature of these networks made them un-confidential for research and development expertise sharing.

ResearchGate (RG) has emerged as one of the first private social networks for researchers and developers. The registration process of RG requires a private domain Email address. This property increases the credibility of RG. Besides, it eliminates fake and anonymous accounts. RG was constructed to increase collaboration between researchers all over the world. It was launched in 2008. Nowadays, RG has more than three million registered researchers and developers [2].

RG is a special type of OSNs. It adds many new features and activities to the normal OSNs. First, it is private and constructed for scientist and researchers. Second, it has a build-in forum, which allows researchers to upload question and answers. Third, it allows developers and researchers to search for jobs. Fourth, it maximizes the collaboration between researchers by sharing publications, opinions and expertise. Finally, Blogs on users' pages have vanished and replaced with the question/answer forum.

In this work, the question/answer platform of RG has studied. To facilitate the current work, a web-crawler to harvest the questions, authors, answers and their writers have implemented. To this end, this information was utilized to generate a directed graph for users. The link between the users or nodes starts from the answerer and ends or points at the questioner. The generated graph has been studied using social network analysis methods to extract node degree, centrality, and connectivity of the graph.

This paper is organized as follows**. Section II** summarizes these paper contributions. **Section III** provides a brief review of the literature that has been completed in the area of social network analysis. **Section IV** presents related graph theory. **Section V**, explain the proposed system model. **Section VI** illustrates the experiment parameters and used data. **Section VII** shows extensive results. Finally, **section VIII** concluded the paper.

II. Contributions

The main contribution in this paper is to answer to the following questions:

1) Is the question/answer platform popular in RG? To answer this question all the questions, answers and the participated users in this platform were harvested and compared with the total number of users in RG.

2) What is the percentage of expertise sharing in RG? To answer this question, the number of questions and answers in the harvested data were measured.

3) Are user's selfishness? To answer this question, node degree and eccentricity have been utilized.
4) Is the connectivity in this network high? To answer this question cluster coefficient, centrality and triangles in the network were utilized.

III. Related Work

Studying real network as graphs inspired researchers over the years. They heavily studied social, information, technological, and biological graphs [3]. These studies gain researchers more insights of how these networks may evolve and how bugs, errors and diseases may separate. In [4], [5], it has been reported that the World Wide Web (WWW) follows an exponential degree distribution with more than 269 thousands nodes and around 1 and half million edges. This study provided that the WWW is a small world graph. In [6], the author studied the citation network as a directed-graph. The author computed node number and vertexes. However, the cluster coefficient and degree correlation were not computed. These results made it hard to predict node distribution in citation network.

In the measurement works [7], [8], authors attempted to generate Internet graph to study its properties. They have produced an undirected graph with 10 thousands nodes and 31 thousands edges. A 0.035 global cluster coefficient value has been computed. These values have been computed over the years again to show the development of the Internet.

A third type of networks have been studied in [9], the author of this study generated a software-classes directed-graph. With 1377 nodes and 2213 edges, the author studied the properties of this small graph. The author reported a mean node-node distance with 1.5 hops. In addition, global clustering coefficient and degree correlation coefficient were computed. This study showed that software classes don't follow the small world phenomenon.

Furthermore, in the works of [10], [11], authors generated a small undirected graph based on harvested data to simulate P2P network. A graph of 880 nodes and 1296 edges has been implemented and studied. Authors found that the average shortest path in this graph is 4.2 hops and the global clustering coefficient is 0.012. Unfortunately, their implemented graph was too small to mimic a swarm of P2P networks.

In the work of Milgram [12], authors of the experiment have analyzed the average path length for social networks of people in the United States. The result of experiment showed that only 20 percent of the packages sent reached their target, which means, average chain length of 6.5. The experiment was repeated at more global scale (18 targets, 13 countries, 60K participants) by Dodds [13]. The result of the experiment showed that 384 messages reaching their target, yielding an average path length of 4.

Finally, authors of [14], [15] attempted to study human's neural networks as a directed graph. They have constructed a directed graph with 307 neural and 2359 edges between them. Subsequently, they studied the properties of this graph. They reported a global clustering coefficient of 0.18 and an average shortest path value of 3.97. These values demonstrated that human's neural network is a small world graph.

In this work, the author attempted to construct a directed graph to study the properties of question-answer environment of RG network. The aim is to analyze this graph to study the properties of the question-answer environment and how it is evolving. In recent years, question answering environments have been studied for yahoo [16] and java forums [17]. In yahoo answers forum; authors analyzed the forum categories and cluster them according to content characteristics and patterns of interaction among the users, and they found that some users focus on some specific topics, while others participate in different categories. They also characterized the entropy of users' interests. The conclusion of their work was to predict, within a given category, whether a particular answer will be chosen as the best answer by the asker. In java forums [17], authors investigated the structure and algorithms in expertise networks in online communities. They tested a set of network-based ranking algorithms, including PageRank and HITS, on the large size social network in order to identify users with high expertise. They also identified the rules for governing the question-answer dynamic in the network. Their algorithms can be applied to evaluate not only Java forum, but also design and implementation of online expertise-sharing communities.

However, these studies did not demonstrate how question-answering environment is behaving or if they follow the small world phenomenon or not. In our work we attempted to answer these quires.

IV. Related Graph Theory

In this section, some background of graph theoretical properties that have been utilized in this study will be presented.

IV.1. Graph Types

There are two types of graphs: directed and undirected graphs. Directed graphs edge is an ordered pair, and the ordered pair represents the direction of the edge that links two vertices [18], [19]. In undirected graphs, there are no directions associated with the edge, and so, the edge is unordered pair. Symmetric relationships between objects can be represented by undirected graphs. In directed graphs in-degree and out-degree of each node is not equal, while it is equal in directed graphs. If the directed graph was represented in matrix, then the matrix will not become a symmetric graph. But in undirected graphs the matrix will always becomes a symmetric graph.

IV.2. Node Degree and Degree Distributions

Node degree is defined as the number of edges that connect the node to other nodes. In undirected graph each node has a degree. However, in directed graph each node has in-degree and out-degree. In-degree is defined as the number of edges end in the node and out-degree is defined as the number of edges start from the node. In the other hand, degree distribution is the function that defines how node degrees behave. Node degree can follow any distribution function. The most famous node distribution function is the power-law function.

IV.3. Power-law Degree Distribution

A power-law is a functional relationship between two quantities, where one quantity varies as a power of another [20].

IV.4. Transitivity or Clustering

Transitivity or clustering in social networks analysis means that, the friend of your friend is likely also to be your friend, and therefore, a triangle is constructed as shown in figure 1. So, if vertex A is connected to vertex B and vertex B to vertex C, then there is a heightened probability that vertex A will also be connected to vertex C. Clustering coefficient is divided into local and global values.

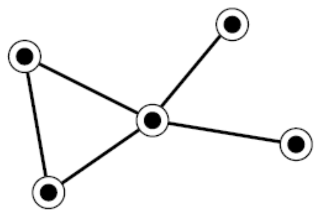

Fig.1 An example for open and closed triplets.

IV.4.1. The Global Clustering Coefficient

The global clustering coefficient is based on triplets of nodes. A triplet is three nodes that are connected by either two (open triplet) or three (closed triplet) undirected ties or edges. A triangle consists of three closed triplets, one centered on each of the nodes. The global clustering coefficient is the number of closed triplets (or 3 x triangles) over the total number of triplets (both open and closed). The measure of global clustering coefficient gives an indication of the clustering in the whole network (global), and can be applied to both undirected and directed networks (often called transitivity [21]. However, it cannot be applied to weighted networks. Global clustering coefficient C can be computed as in equation (1).

$$C = \frac{1}{n}\sum_{i} C_i \qquad (1)$$

Where Ci is the local value proposed by Watts and Strongatz in [14]. Where a (connected triple) means a single vertex with edges running to an unordered pair of others.

Also, C can be used to measure the fraction of triples that have their third edge filled in to complete the triangle. In contrast, C is the mean probability that two vertices that are network neighbors of the same other vertex will themselves is neighbors (closed triplet). This notion helps us, in this paper, to assess the RG users mutual interests.

IV.4.2. Local Clustering Coefficient

The main idea behind calculating the cluster coefficient of a node in a graph is based on the numbers of triangles that a node participates in over the total number of triangles that it can participate in. The local clustering coefficient is based on ego network density or local density [22], [23], and [14]. Therefore, the outcome ranges between 0 and 1. 0 if no ties exist between the neighbors and 1 if all possible ties exist. This metric is used in this paper, in order to find the connectivity strength for the RG community users, by finding the number of users who are a member in a triangle in the graph. If this number is high, it means that the users have a strong connectivity in their expertise. Otherwise, If this number is not high, then it indicates that users do not have much mutual interests, as shown in results section, section VII.

IV.5. Centrality

Graph centrality is the methods and algorithms that can be utilized to find the core or central nodes of the graph. Many methods and algorithms have been implemented to rank nodes in a graph. Eigenvector, closeness and betweenness are the most common methods to calculate the centrality of nodes. In the following section, four different methods will be presented. These metrics are used in order to evaluate our constructed graph (which represents the RG activity), as explained in section V, system model. For example, to find the number of users that has major impact on RG users activity, e.g.; ask/answer questions. This will be clear in Section VII, results.

IV.5.1. Betweenness and Closeness Centrality

Betweenness of a node is defined as total number of paths that flow through this node. Closeness of a node is defined as the total number of shortest paths that flux through the node [24]. Closeness is used in centrality estimation more than betweenness since it is hard to compute all the paths that pass through a node.

IV.5.2. Eigenvector Centrality

Eigenvector centrality is a measure of the importance of a node in a network. It assigns relative scores to all nodes in the network based on the principle that connections to high-scoring nodes contribute more to the score of the node in question than equal connections to low-scoring nodes. Google's PageRank is a variant of the Eigenvector centrality measure. Adjacency matrix can be used to find eigenvector centrality [24].

IV.5.3. Eccentricity

The distance between two vertices in a graph is the number of edges in a shortest path connecting them (also called a graph geodesic). This is also known as the geodesic distance. There may be more than one shortest path between two vertices. If there is no path connecting the two vertices, i.e., if they belong to different connected components, then conventionally the distance is defined as infinite.

In the case of a directed graph the distance d(u,v) between two vertices u and v is defined as the length of a shortest path from u to v consisting of arcs, provided at least one such path exists. In contrast with the case of undirected graphs, d(u,v) does not necessarily coincide with d(u,v), and it might be the case that one is defined while the other is not.

The eccentricity ϵ (v) of a vertex v is the greatest geodesic distance between v and any other vertex. It can be thought of as how far a node is from the node most distant from it in the graph.

The radius r of a graph is the minimum eccentricity of any vertex, equation (2).

$$r = \min_{v \in V} \epsilon(v) \qquad (2)$$

The diameter d of a graph is the maximum eccentricity of any vertex in the graph. That is, d it is the greatest distance between any pair of vertices, equation (3).

$$\mathrm{d} = \max_{v \epsilon V} \epsilon(v) \qquad (3)$$

To find the diameter of a graph, first find the shortest path between each pair of vertices. The greatest length of any of these paths is the diameter of the graph.

V. System Model

In this proposed model, graph is used in order to formulate the problem. In RG system, there is a group of research users who are signed up into social network. In this research community, users can ask/answer questions. Also, a user can ask or answer multiple questions. A user is modeled as a node, and the question answer is modeled as a directed edge, such that the edge incoming to the user node who asked the question and the edge is outgoing from the user who answered the question. Fig. 2 shows a simple example on how to construct the graph for the RG activity. There are three users, U1, U2, and U3. Edge 1 (e1), and edge 2 indicate that U1 answered two questions that asked by U2. Edge 3 indicates that U2 answered a question which is asked by U1. Also, U1 answered a question asked by U3. Notice that U3 did not answer any question. U4 asked a question and no one answered, also U4 did not answer any question.

Another example, for special case scenario, when there are some nodes in the graph and no edges exist, this means that nobody answered questions.

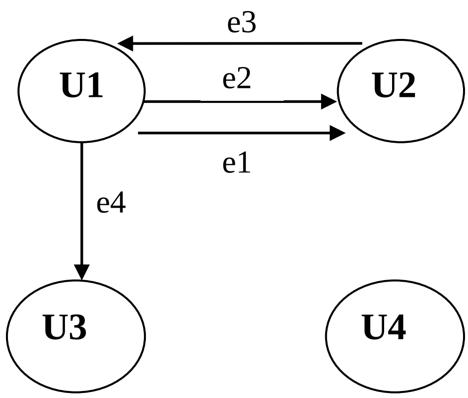

Fig.2 Simple example on the proposed model graph construction.

VI. Experiment

A crawler that collects questions and answers from RG platform was implemented. The crawler harvested many information of each question, such that, the authors name, the date of the question, number of answers and the name of answerers. These data have been utilized to generate a directed graph that connects the subscribers through question-answer platform. A connection occurs between a node 'A' and a node 'B' if node A asked a question and node B answered the question. The direction in this graph starts from the answerer to the questioner. A total of 82,682 questions have been collected, and 506765 answers have been harvested with an average of 6.1 answers for a question. 59579 subscribers asked or answered at least one question. The constructed graph consisted of 59,579 nodes with 506,765 links (edges) between these nodes (users). The average node degree (for outgoing edges) is **8.5**. This number indicates the average number of answered question by a user. Also, the average number of answers for a question is **6.1** (from the collected data). There is a high variation between the number of questions and the number of answers in the platform (The experiment was conducted in Nov 2013). Table 1 shows a summary of the collected data.

TABLE I

SUMMARY OF THE COLLECTED DATA

Number of questions	82682
Number of Answers	506765
Number of users	59579
Average number of questions	1.387
Average number of answers for a question	6.1
Average number of answers from a user	8.5

VII. Results

The results are divided into three main parts. The first part is node degree distribution. The second part is the centrality study of the graph. The final part is the connectivity study of this graph. These parts are as follows

Fig 3 shows the node degree distribution of the generated graph. It can be observed that the node degree distribution follows a power law distribution. In addition it can be observed that the highest node degree is approximately 300. This number is small comparing with other graphs such as AS or IP graphs.

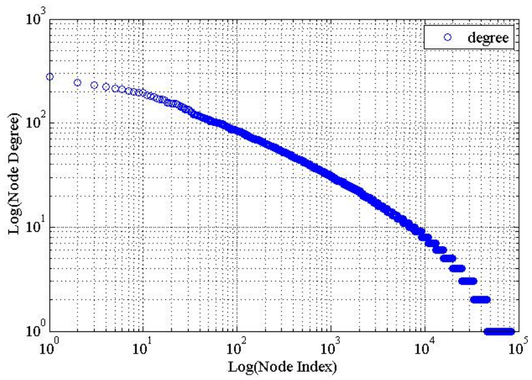

Fig3. The node degree distribution of the generated graph

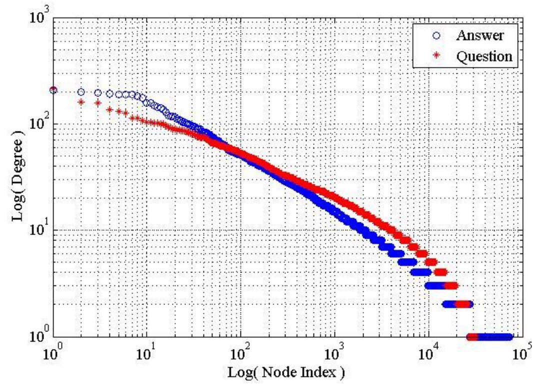

Fig4. The in/out degree distribution of the nodes

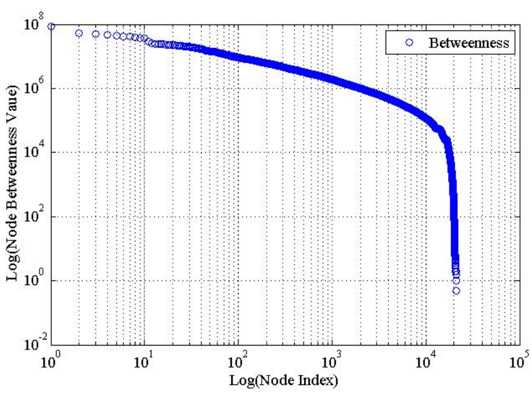

Fig5. The betweenness centrality values of the nodes in the graph

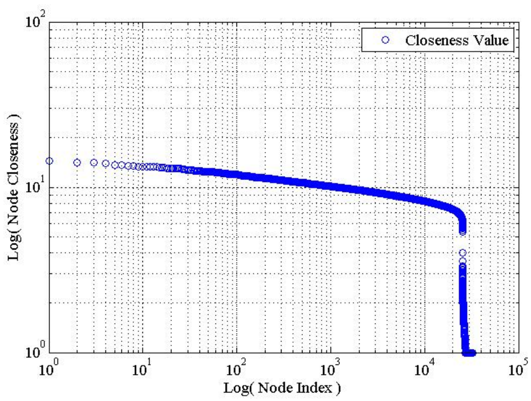

Fig6. Closeness Value distribution

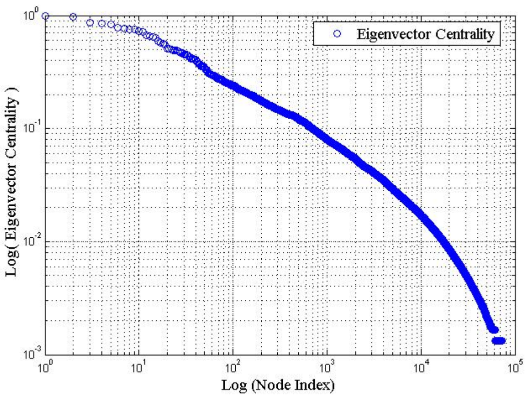

Fig7. The Eigenvector for the graph

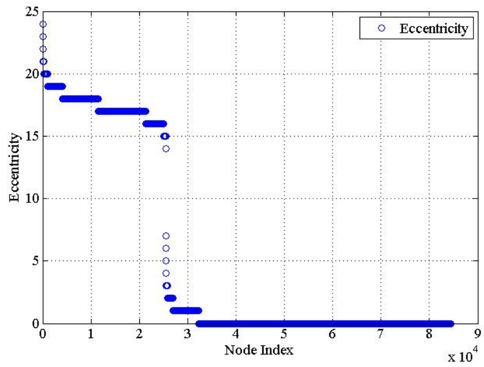

Fig8. The eccentricity value of nodes in the graph

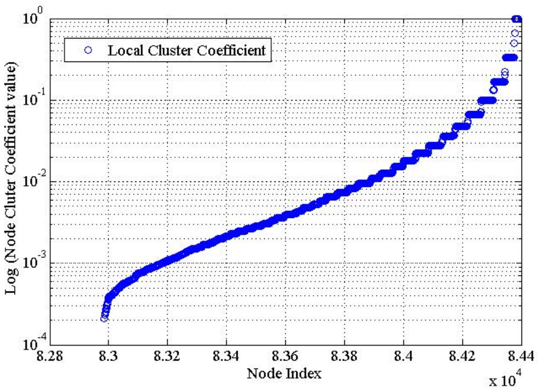

Fig9. The local cluster coefficient of the nodes

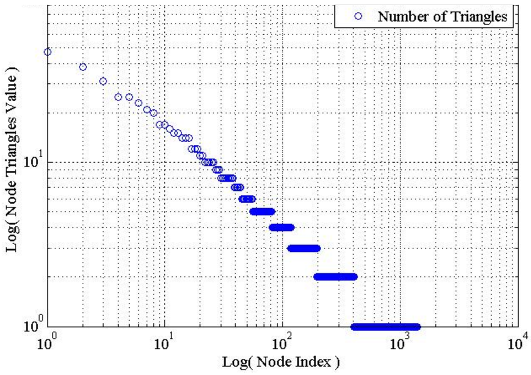

Fig10. The relation between nodes and the number of triangles they construct with their

Fig4 shows the in/out degree distribution of the nodes. The out degree represents the total number of answers that a user participated in the network. The in-degree represented the total number of answers a user got for his questions. This figure shows that the answer and the question degree distributions follow a power law distribution as in the node degree distribution. This means that, only few users participated with a massive number of questions or answers in the network. On the other hand, the other user only shared a few questions or opinions. Moreover it can be observed that the total number of questions shared in the system from a user exceeds the number of answers or opinions shard from other users. The slope of the answer curve descends faster than the slope of the question slope.

Fig5 shows the betweenness centrality values of the nodes in the graph. It can be observed that the number of nodes that have a betweenness value greater than 0 are approximately 20,000 nodes only. In addition, it can be observed that the slop of the graph descending smoothly following a power-law for a proximately 11,000 nodes. However, the slope after 11,000 descends very fast to values less than 1. This figure shows that less than 11,000 users have significant impact on the network. These nodes have high betweenness values. These nodes are only 13% of the nodes in our graph. After crawling the data from the question-answer environment, we attempted to study the geographical distribution of participant. We utilized the names of the participants in the crawled data and generated a seeding list to a second crawler. The second crawlers harvested the users' pages from RG. These pages contain a database of information of each user. From these information, we filtered the name of institute or affiliation. We generated a new seeding list with these names and searched Google for their web-site addresses. From the received addresses, we utilized the name of countries that can be found in the end of the URL address. Unfortunately, 27% of the total number of participators in the question-answer environment did not include a well-know name of an institute. We eliminated these users from our analysis. We observed from the collected data the following. First, most of the participators (askers and answerers) are from India (60% total). Second, Middle East users have a high contribution in asking question and less contribution in answering (30% asking, 7% answering). Third, American institute participated with the highest contribution ratio in answering (28%). Finally, we noticed that users from Far East have the lowest contribution ratio in question asking (2%) and in answering 1%.

Fig 6 shows the second centrality metric of our graph, closeness. It can be observed from the figure that the closeness graph decreases continuously until it reaches a discontinuous point. Subsequently, it falls to zero. These nodes are the leaves of our graph. It can be noticed that the number of nodes with small closeness value is small comparing with the other nodes.

Fig 7 shows the final eigenvector centrality metric in our study. As it can be observed, this figure is a power-law distribution, and it is obviously noticed from the three centrality metrics, few nodes have high centrality value. Results obtained from Figures 5 and 6 indicate that only few RG users are highly active in asking and answering questions.

Fig 8 shows the eccentricity value of nodes in the graph. As mentioned eccentricity of a node is the longest path between it and other nodes.

This means that when eccentricity value is 0 the node has no links to other nodes 'disconnected'. These disconnected nodes are nodes that asked questions and no one answered them. The number of these nodes is approximately 45,000. Moreover, the highest eccentricity value in the graph is the diameter of the graph. This means that the diameter of our graph is 24 (note that bigger diameter value is better, since this indicates many users questions are answered. However, less diameter value means more users questions are left with no answers). Finally, this figure shows that the number of nodes that have short paths to all other nodes is small '10,000' nodes. These nodes are the nodes that have significant effects in the betweenness values.

Fig 9 shows the local cluster coefficient of the nodes. It can be observed that some nodes have a local cluster coefficient of 1. Which mean that all their neighbors are neighbors. It also can observe that the values decrease in a smooth way to reach 0.0001.

Fig 10 shows the relation between nodes and the number of triangles they construct with their neighbors. It can be shown that only few nodes ‘less than 20’ participate in constructing more than 10 triangles with their neighbors. In addition it can be noticed that only one node participated in larger than 40 triangles in the graph. Moreover, only 400 nodes have participated in more than one triangle and approximately 1000 nodes participated in the generation of one triangle.

These results show that the graph is lousy and it is not highly connected graph. It can be observed that users participate in asking questions are more than users answering. This fact may reflect selfishness in the question/answer platform.

VIII. Conclusions

With more than two million subscribers, RG network is one of the most popular private social networks in the world. Researchers and developers can meet and share their expertise in this network. RG provide many functionalities and services for its subscribers. One of these services is the question/answer environment where users can share their queries and obtain answers. We believe that question/answer platform is the most important service to share expertise in RG. In this work, the question/answer platform of RG is studied. A web-crawler has been implemented in order to harvest the questions, authors’ information and answers from RG. The collected data has been utilized to generate a directed graph that connects users according to their questions and answers.

The results analysis shows that many users in RG do not care about sharing their expertise by not answer questions. The number of users participated in the question-answer platform is less than 85K comparing with 2 million subscribers, which means that, the sharing percentage is low. In addition, it is shown that the participants in this platform tend to ask more than answering. It is observed that most of the nodes in the graph are disconnected, which means that most of the questions have no answers. It also observed that the number of triangles in the network graph is small, which means that the network is not highly connected, such as AS and IP graphs. However, it is observed that a few number of users who shared many answers obtained a high centrality value. Nevertheless, this number is small comparing with the total number of users. Our recommendation is that, RG should use a new method to stimulate the users to participate in the question-answer environment.

References

[1] W. Wongyai and L. Charoenwatana, “Examining the network traffic of facebook homepage retrieval: An end user perspective,” *Computer Science and Software Engineering (JCSSE)*, 2012 International Joint Conference on, pp. 77 – 81, 2012.
http://dx.doi.org/10.1109/jcsse.2012.6261929

[2] “ResearchGate,” http://www.researchgate.net/.

[3] M. E. J. Newman, "The Structure and Functions of Complex Networks", SIAM Review, Vol. 45, No. 2, pp.167-256, 2003.
http://dx.doi.org/10.1137/s003614450342480

[4] R. AlBert,H. Jeong, and , A.L.Barabasi, Diameter of the World Wide Web, Nature, 401 (1999), pp. 130-131. http://dx.doi.org/10.1038/43601

[5] L. Barabasi, R. Albert, and H. Jeong, "Scale-free characteristics of random networks: The toplogy of the World Wide Web, Phys. A, 281 (2000), pp. 69-77.
http://dx.doi.org/10.1016/s0378-4371(00)00018-2

[6] S. Redner, How popular is your paper? An empirical study of the citation distribution. Eur. Phys. J. B, 4(1998), pp. 131-134.
http://dx.doi.org/10.1007/s100510050359

[7] S. N. Dorogovtsev and J. F. F. Mendes, Language as an evolving word web, Proc. Roy. Soc. London Ser. B, 268 (2001), pp. 2603-2606.
http://dx.doi.org/10.1098/rspb.2001.1824

[8] R. Ferrer I Cancho and R. V. Sole, The small world of human language, Proc. Roy. Soc. London Ser. B, 268 (2001), pp. 2261-2265.
http://dx.doi.org/10.1098/rspb.2001.1800

[9] S. Valverde, R. F. Cancho, and R. V. Sole, Scale-free networks from optimal design, Europhys. Lett, 60 (2002), pp. 512-517.
http://dx.doi.org/10.1209/epl/i2002-00248-2

[10] L. A. Adamic, R. M. Lukose, A. R. Puniyani, and B. A. Huberman, Search in power-law networks, Phys. Rev. E, 64 (2001), art. No. 046135.
http://dx.doi.org/10.1103/physreve.64.046135

[11] M. Ripeanu, I. Foster, and A. Iamnitchi, Mapping the Gnutella network: Properties of large-scale peer-to-peer systems and implications for system design, *IEEE Internet Comput*., 6(2002), pp. 50-57.
http://dx.doi.org/10.1109/4236.978369

[12] Jeffry Travers and Stanley Milgram. “An Experimental Study of Small World Problem”. In: Sociometry 32.4 (1969). Pp. 425-443. http://dx.doi.org/10.2307/2786545

[13] Roby Muhamad Peter Sheridan Dodds and Duncan J. Watts. “An Experimental Study of Search in Global Social Networks”. In: Science 301.5634 (Aug. 2003). PP. 827-629. http://dx.doi.org/10.1126/science.1081058

[14] D.J. Watts and S. H. Strohatz, Collective of “Small-world” networks, Nature, 393 (1998), pp. 440-442. http://dx.doi.org/10.1038/30918

[15] J. G. White, E. Southgate, J. N. Thompson, and S. Brenne, The structure of the nervous system of the nematode C. Elegance. Trans. Roy. Soc. London, 314 (1986), pp. 1-340. http://dx.doi.org/10.1098/rstb.1986.0056

[16] Lada A. Adamic, Jun Zhang, Eytan Bakshy, and Mark S. Ackerman, Knowledge Sharing and Yahoo Answers: Everyone Knows Something. WWW 2008: 665-674. http://dx.doi.org/10.1145/1367497.1367587

[17] Jun Zhang, Mark S. Ackerman, and Lada Adamic. 2007. Expertise networks in online communities: structure and algorithms. In Proceedings of the 16th *international conference on World Wide Web (WWW '07)*. ACM, New York, NY, USA, 221-230. DOI=10.1145/1242572.1242603 http://doi.acm.org/10.1145/1242572.1242603

[18] Shlomo Havlin and Reuven Cohen (2010). Complex Networks: Structure, Robustness and Function. Cambridge University Press. http://dx.doi.org/10.1017/cbo9780511780356

[19] J.A. Bondy and U.S.R. Murty. Graph Theory. Springer, GTM 244, 2008. ISBN: 978-1-84628-969-9

[20] Newman, M. E. J. (2005). "Power laws, Pareto distributions and Zipf's law". Contemporary Physics 46 (5): 323–351. http://dx.doi.org/10.1080/00107510500052444

[21] Stanley Wasserman, Katherine Faust : Social Network Analysis: Methods and Applications, Cambridge University Press, 1994, ISBN 1139788612, http://dx.doi.org/10.1017/cbo9780511815478

[22] John Scott: Social Network Analysis, SAGE Publications, 2000, ISBN 0761963391, 9780761963394.

[23] Uzzi, Brian and Jarrett Spiro (2005). "Collaboration and Creativity: The Small World Problem." American Journal of Sociology 111(2): 447-504. http://dx.doi.org/10.1086/432782

[24] Opsahl, T., Agneessens, F., Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. Social Networks 32, 245-251. http://dx.doi.org/10.1016/j.socnet.2010.03.006

Received: January 15, 2015; Published: February 17, 2015