***Zargari H.Y.*[1], *Khalili F.P.*[2], *Janani N.E.*[2], *Sami B.T.*[2]**

[1]Institute of Information Technology of ANAS, Baku, Azerbaijan

[2]Ardabil Branch-Islamic Azad University, Ardabil, Iran

[1,2]hamid_zargari@ardabiltelecom.ir

EXPERIMENTAL EVALUATION OF ACADEMIC PERFORMANCE OF STUDENTS AND THEIR CENTRALITIES IN MOBILE SOCIAL NETWORK

We collected required social information about a group of undergraduate students in university from their mobile social network for preparation the social graph and their centralities. Their educational performance compared with four centralities in order to evaluate the relation between educational performance and social centralities. Finally the correlation coefficients were calculated and discussed.

***Key words**: mobile social network, degree centrality, closeness centrality, betweenness centrality, eigenvector centrality, educational performance.*

1. Introduction

Social Network Analysis (SNA) is a well known topic for sociologist in recent decades. It is a field that deals with interdependent social actors and the flow of resource along the relational linkages between actors. Among many methodologies in social network analysis, graph theory is used where nodes are supposed as actors and edges as ties.

The centralities are measures to find importance and high level role of actors in the social networks. There are different centralities in social networks and all of them have independent concepts. In 1950s when network analysis was primarily defined [1], network centrality and its role in different environments have been widely focused and we look at this field of knowledge from mobile social network view of point.

Centrality in social network is shown to be is related to educational performance for adolescents [2]. A web base study compares the educational performance and network centrality in a group of students [3]. In an approach the relationship between network centrality and academic performance among a group of 47 PhD students from emails.

It shows that a reversed U-shaped curve appears between network centrality (both betweenness and degree) and students' academic performance where the x-axis is centrality and y-axis is academic performance [4].

In this study, centralities in a mobile social network are compared to educational performance of students in a group of electrical engineering students. We wanted to find the effect of centrality in a mobile social network on students' marks.

Because of using mobile communication concepts in this paper, we should review some terminologies in the following section. After definition the used material and methods, the later sections are dedicated to the results of analysis in different forms of centralities. Finally the results are discussed.

2. Billing in mobile communication

In this study we work on CDR (Call Detail Records) files to retrieve the connection details of subscribers in a Mobile Switching Center (MSC). CDR files are made by every switching system to prepare the billing information of subscribers such as call duration, originating and terminating numbers (A-numbers and B-numbers), furthermore, original and final physical coordination. CDR files are closed periodically and the next file is opened to continue the data warehousing. In such studies the MSISDN (Mobile Station Integrated Service Digital Network Number) can be assumed as nodes or vertices [5].

CDR files are transformed to matrices by time stamps and then can be transformed again to a sequence of graphs in different time period.

3. Materials

Centralities in a mobile social network of 18 single and female 22-23 years old university students in a class have been calculated. Due to occupation of some of male students we focused only on female students who have not other concerns such as occupation or marriage in sampling group. Using a prepared program we retrieved the required information from billing log files and collected them in tables. We collected this information in seven days in an Excel sheet.

They made a matrix with 107×107 elements. They are retrieved from all calling logs in mobile switching center in snowball method. We collected 2,805 calls and after simplifying the rest number of actors were 107 actors. For preserving the privacy rights, all the information is gathered from volunteers and all the phone numbers were encrypted.

The average marks of students during 7 terms (3.5 years) are collected as criteria of students' performance. We also compared them by their other scientific activities and the lecturer's opinion to have more real estimations.

The information was applied to the Ucinet software [6], to find the centralities as in Table 1.

Table 1

One Sample Line of Prepared Table from Ucinet

ID	OutDg	InDg	OutBon	InBon	Out2Ste	In2Ste	OutARD	InARD	Between
A1	0	0.03	0	0.54	0	0.05	0	0.04	0

Used abbreviations in the Table 1 are as follow:

OutDg: Outgoing Degree centrality of nodes in social digraph (directed graph).

IntDg: Incoming Degree centrality of nodes.

OutBon: Outgoing Bonacich (Eigen Vector) centrality of nodes [7].

InBon: Incoming Bonacich centrality of nodes.

Out2Ste: Outgoing Two Step Closeness centrality of nodes (k-steps indicates how many steps should be traversed to other actors).

In2Ste: Incoming Two Step Closeness centrality of nodes.

OutARD: Outgoing Average Reciprocal Distance of nodes.

InARD: Incoming Average Reciprocal Distance of nodes.

Between: Betweenness centrality of nodes.

Using NetDraw [8] we draw the graph based on different centralities. Figure 1 (left) shows a sample social graph made up of both female and male students in the same class. Prior to considering the female students, we collected information from 30 female and male students for our study. Here we show only the mixed graph to be compared with the female graph (Figure 1 - right). We focus only on the graph of female students.

The square formed nodes are first considered female students that are supposed as first group for snowball method. Some of them have been deleted because of simplification of graph. It should be mentioned that the primarily graph has been simplified by a few times deleting the pendants and isolates. It is done until making an unchangeable graph, though some of our original actors are deleted. So there are only a few square formed nodes.

The graphs in figure are prepared from the logs in a normal working week. There is not any special event or annual vacation in that time period. It illustrates that the social graph has not big components and the number of small groups are more.

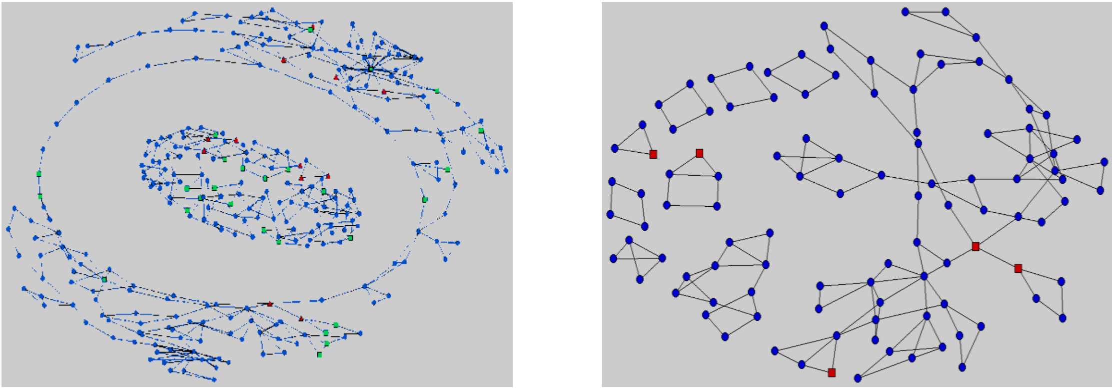

Figure 1. Graph of all students comprising female and male students by spring embedded layout (left) and the social graph of female students after deleting the pendants and isolates by spring embedded layout (right)

4. Results

4.1. Degree centrality

Degree centrality is the number of links that are connected to a node. If ties have directions, then we can categorize them to indegree and outdegree. Indegree is the number of links hat are terminated in the node and outdegree is the number of originating ties from the node.

For a graph $G=(V, E)$ with n vertices, the degree centrality $C_D(v)$ for vertex v is:

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

where $\deg(v)$ is the degree of the vertex v.

Using Netdraw, we prepared the social graph that shows the degree centrality of nodes up to their node size as in figure 2 (left).

It is obviously shown that none of our focused groups have high values in centrality.

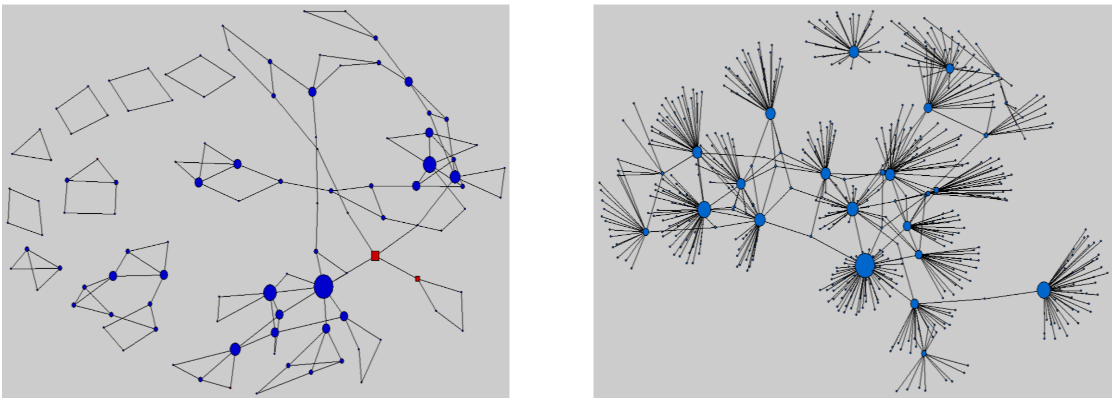

Figure 2. Graph of female students, size of nodes depicts the degree centrality (left) and graph of female students in 2 months, size of nodes depicts the degree centrality (right)

The curve in figure 3 (left) shows the average marks against outdegree centrality. It shows that high values of centrality are performed by middle marks. On the other word, the more powerful and weak students have little centrality scores. In the Figure 3 – right, the average marks of students were compared with their in degree centrality.

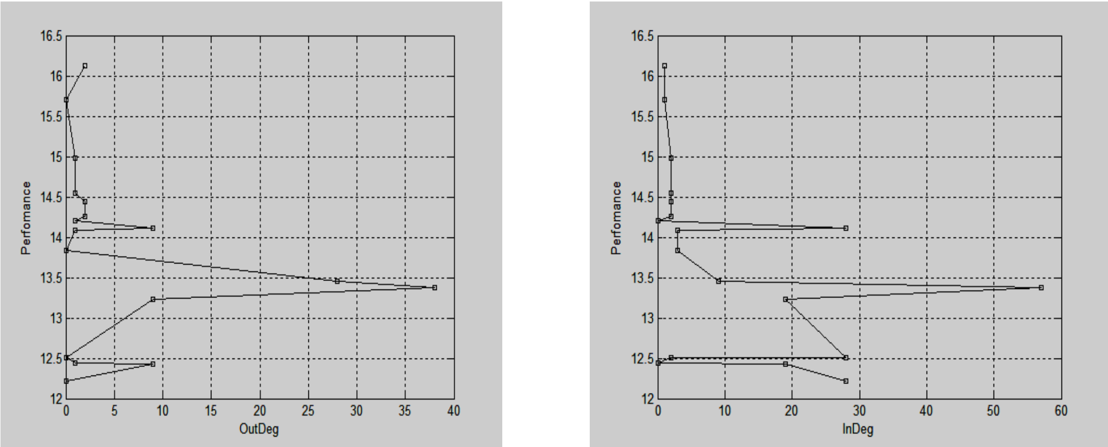

Figure 3. Educational performance against outdegree centrality (left) and Educational performance against indegree centrality (right)

The same result for Figure 3- right is obtained, but low marks have a few better performances.

4.2. Betweenness centrality

Betweenness is another centrality of a node into a graph. Vertices that appear on many shortest paths between other vertices have higher betweenness than the others. For a graph $G=(V,E)$ with n vertices, the betweenness centrality $C_B(v)$ for vertex v is:

$$C_B(v)=\sum_{s\neq v\neq t\in V}\frac{\sigma_{st}(v)}{\sigma_{st}}$$

where σ_{st} is the number of shortest paths from s to t, and $\sigma_{st}(v)$ is the number of shortest paths from s to t that pass through a vertex v.

In calculating betweenness and closeness centralities of all vertices in a graph, it is assumed that graphs are not digraph.

Figure 4 shows the social graph, but the size of nodes is dedicated by their betweenness centrality. Only one member of our initially focused group has high score in betweenness centrality (square shaped nodes).

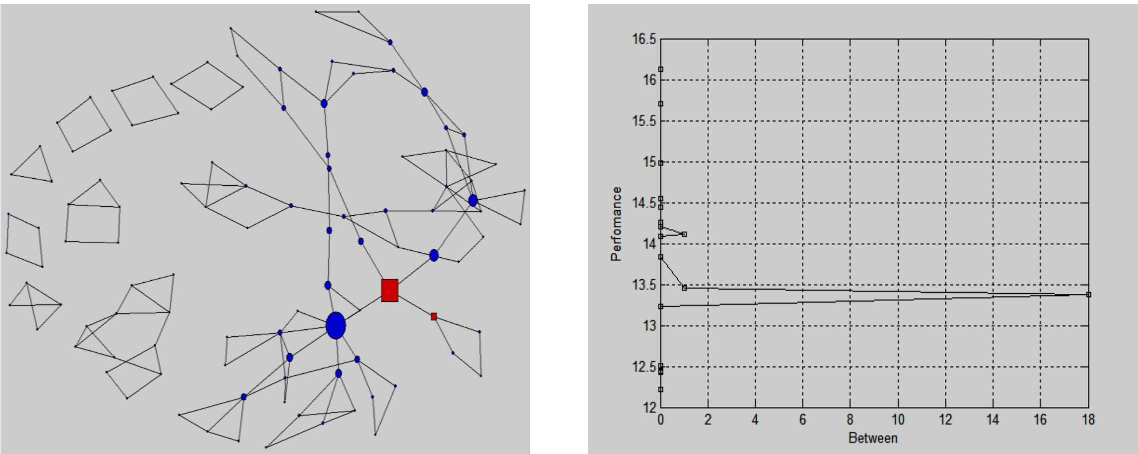

Figure 4. Graph of female students, size of nodes depicts the betweenness centrality (left) and Performance against betweenness centrality (right)

In figure 4-right we can see the performance against betweenness centrality where high values are dedicated to middle marks.

4.3. Closeness centrality

Closeness is a centrality scale within a graph. The node those have short geodesic distances to other vertices have higher closeness.

Closeness that is a measure of centrality in the network theory is defined as the mean geodesic distance (that is, the shortest path) between a vertex v and all other vertices reachable from it:

$$\frac{\sum_{t \in V \setminus v} d_G(v,t)}{n-1}$$

where $n \geq 2$ is the size of the network's 'connectivity component' V reachable from v and d_G is the geodesic distance.

Closeness also can be defined as reciprocal of this quantity. Communicated information in both ways is the same. The closeness $C_C(v)$ for a vertex v is the reciprocal of the sum of geodesic distances to all other vertices of V [9]:

$$C_C(v) = \frac{1}{\sum_{t \in V \setminus v} d_G(v,t)}.$$

Figure 5 shows the closeness centrality against performance where actors with middle and low marks have higher centrality. It is illustrated that actors who have middle marks obtain high values in closeness centrality.

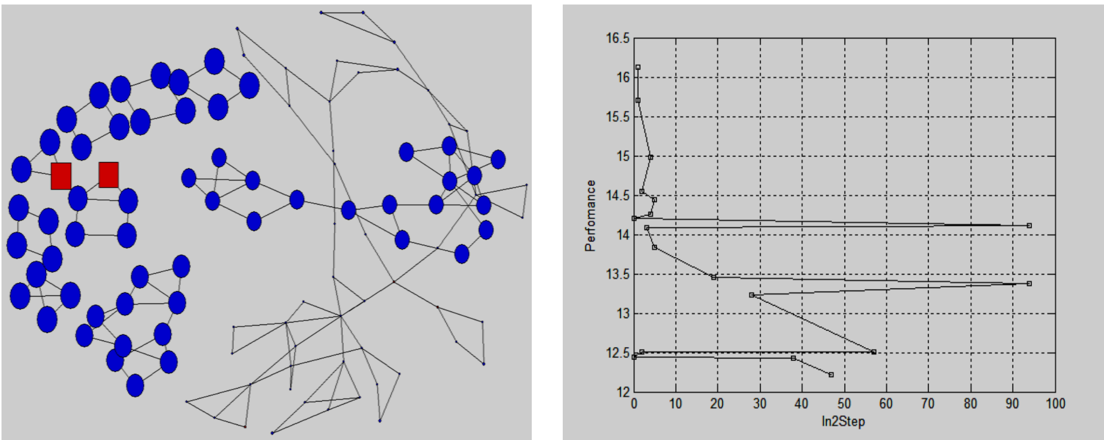

Figure 5. Closeness centrality by node size (left) and Performance against closeness centrality (right)

4.4. Eigenvector centrality

It is a relative score to all nodes in the network based on the principle that connections to high-scoring nodes low-scoring nodes. Let x_i denote the score of the ith node. Let $A_{i,j}$ be the adjacency matrix of the network.

Hence $A_{i,j}$=1 if the ith node is adjacent to the jth node, and $A_{i,j}$ = 0 otherwise. More generally, the entries in A can be real numbers representing connection strengths, as in a stochastic matrix. For the i^{th} node, let the centrality score be proportional to the sum of the scores of all nodes which are connected to it.

Hence

$$x_i = \frac{1}{\lambda} \sum_{j \in M(i)} x_j = \frac{1}{\lambda} \sum_{j=1}^{N} A_{i,j} x_j$$

where $M(i)$ is the set of nodes that are connected to the i^{th} node, N is the total number of nodes and λ is a constant. However, the additional requirement that all the entries in the eigenvector are only the greatest eigenvalue results in the desired centrality measure.

The i^{th} component of the related eigenvector then gives the centrality score of the i^{th} node in the network. Power iteration is one of many eigenvalue algorithms that may be used to find this dominant eigenvector [10].

Figure 6 shows the eigenvector centralities by size of nodes. Figure 7 (left and right) depicts the educational performance against Bonacich [7] in and out centralities (Bonacich in and Bonacich out centralities).

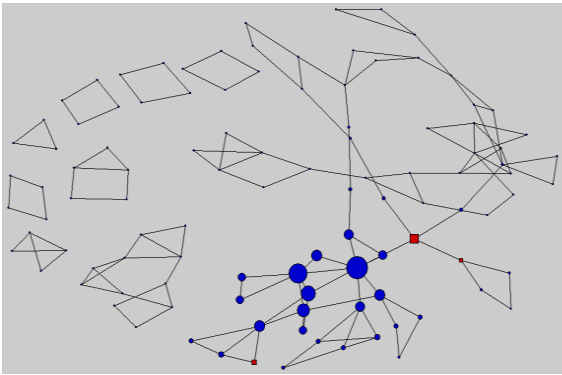

Figure 6. Eigenvector centrality with node size

Figure 7 shows the Bonacich centralities. They obviously depict that the very high and very low marks have not high values in Bonacich centrality.

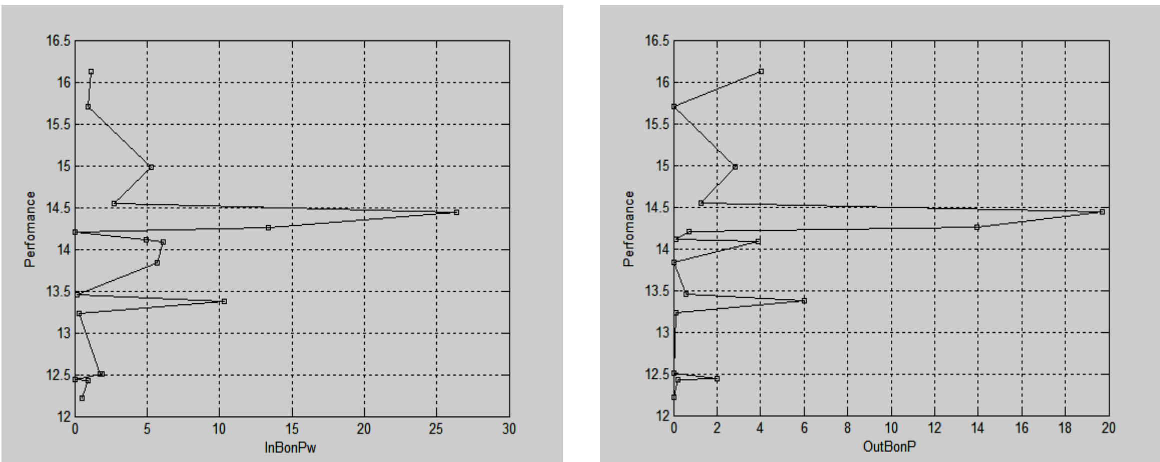

Figure 7. Performance against In Bonacich and Out Bonacich centralities

5. Conclusions

We wanted to find how the mobile social network centrality can affect the educational performance of students and for finding any relation between them we calculated the Pearson correlations. In this study the different centralities are considered as independent variables and students' marks as dependent variables. The Pearson correlations between the independent and dependent variables are shown in the Table 2. There is not any obvious correlation unless in InDeg and InARD.

Table 2

Pearson correlations between marks of students (dependent variable) and 9 different centralities (independent variable)

OutDg	InDg	OutBon	InBonPw	Out2Ste	In2Step	OutARD	InARD	Between
-0.16	-0.4261	-0.269	0.2141	-0.14	-0.3628	-0.134	-0.4129	-0.096

We could not find any direct or reverse proportional between centralities and educational performance. It is realized that almost in all types of centralities the best and the worst numbers have not high values. This result is obviously shown in Bonacich centrality.

The Bonacich centrality was shown that is related to educational performance [4] and our works emphasize the reversed U-shaped curve for Bonacich so that the high values of Bonacich centralities are dedicated to middle marks. In the mentioned paper as our work, the marks were supposed as independent variable and Bonacich centralities were considered as dependent variable.

Betweenness and indegree centralities seem to have reverse proportion to performance that should be considered in other studies.

6. References

1. Bavelas A., Communication patterns in problem-solving groups. Department of Economic and Social Science, Massachusetts Institute of Technology, 1950. grace.evergreen.edu/~arunc/texts/cybernetics
2. Calvó-Armengol A., Patacchini E., Zenou Y., Peer effects and social networks in education // Review of Economic Studies, 2009, v.76, no.10, pp. 1239-1267.
3. Yang H.-L., Tang J.-H., Effects of social network on student’s performance: a web-based forum study in Taiwan // Journal of Asynchronous Learning Networks, 2003, v.7, no.3, pp. 93-107.
4. Zhang Y., Rajabzadeh I., and Lauterbach R., Student network centrality and academic performance: evidence from united nations university. UNU-MERIT Working Paper Series, 2009, no.034, 31 p., http://www.merit.unu.edu/publications/wppdf/2009/wp2009-034.pdf
5. International Telecommunication Union: www.itu.int
6. Borgatti S.P., Everett M.G., and Freeman L.C. Ucinet for Windows: software for social network analysis. Harvard, MA: Analytic Technologies, 2002.
7. Bonacich P., Power and centrality: a family of measures // The American Journal of Sociology, 1987, v.92, no.5, pp. 1170-1182.
8. Borgatti S.P. NetDraw: Graph visualization software. Harvard: Analytic Technologies, 2002.
9. Sabidussi G., The centrality index of a graph // Psychometrika, 1966, v.31, no.4, pp. 581-603.
10. Newman M.E.J., The mathematics of networks. http://www-personal.umich. Edu /~mejn /papers/palgrave.pdf

UOT 004.351

Zərgəri Həmid Y.[1], Xəlili Fatimə P.[2], Canani Nemət E.[2], Sami Behzad T.[2]
[1]AMEA İnformasiya Texnologiyaları İnstitutu, Bakı, Azərbaycan
[2]Azad İslam Universiteti – Ərdəbil Universiteti, Ərdəbil, İran
[1,2]hamid_zargari@ardabiltelecom.ir

Tələbələrin təhsil məhsuldarlığına və onların mobil sosial şəbəkələrdə mərkəziliklərinə eksperimental yanaşma

Sosial qrafın qurulması və mərkəziliklərin hesablanması üçün tələb edilən sosial məlumatlar universitetin bakalavr səviyyəsi tələbələrinin bir qrupunun mobil sosial şəbəkəsindən toplanmışdır. Təhsil məhsuldarlığı ilə sosial mərkəzilik arasında əlaqəni qiymətləndirmək üçün tələbələrin təhsil məhsuldarlığı mərkəziliyin dörd növü ilə müqayisə edilmişdir. Nəticədə onlar arasında korrelyasiya ədədləri hesablanmış və analiz edilmişdir.

Açar sözlər: *mobil sosial şəbəkə, dərəcə üzrə mərkəzilik, vasitəçilik üzrə mərkəzilik, yaxınlıq üzrə mərkəzilik, məxsusi vektor üzrə mərkəzilik, təhsil məhsuldarlığı.*

УДК 004.351

Заргари Гамид Ю.[1], Халили Фатима П.[2], Джанани Немят Э.[2], Сами Бехзад Т.[2]
[1]Институт Информационных Технологий НАНА, Баку, Азербайджан
[2]Ардебильский Филиал – Исламский Университет Азад, Ардебиль, Иран
[1,2]*hamid_zargari@ardabiltelecom.ir*

Экспериментальный подход к производительности образования студентов и их центральности в мобильной социальной сети

Социальная информация, требуемая для построения социального графа и вычисления центральностей, была собрана из мобильной социальной сети группы студентов младших курсов университета. Для оценки отношения между производительностью образования и социальной центральностью было проведено сравнение производительности образования студентов с четырьмя видами центральности. В результате были вычислены и проанализированы корреляционные числа между ними.

Ключевые слова: *мобильная социальная сеть, центральность по степени, центральность по посредничеству, центральность по близости, центральность по собственному вектору, производительность образования.*