Lithuanian Journal of Statistics 2012, vol. 51, No 1, pp. 51-56
Lietuvos statistikos darbai 2012, 51 t., Nr. 1, 51–56 p.
www.statisticsjournal.lt

SOCIAL NETWORK OF FACULTIES ACCORDING TO STUDENT PREFERENCES IN TRANSITION TO HIGHER EDUCATION

Güneş Mutlu[1], Ahmet Mete Çilingirtürk[2]

[1] Department of Research and Development, Istanbul Kultur University

Address: Ataköy Yerleşkesi Bakırköy, Istanbul, Turkey, E-mail: gunesgurgen@gmail.com

[2] Department of Econometrics, Marmara University

Address: Bahçelievler Kampüsü Bahçelievler, Istanbul, Turkey, E-mail: a.cilingirturk@gmail.com

Received: March 2012 Revised: April 2012 Published: November 2012

Abstract. In social network analysis, the studies on weighted adjacency matrix of nodes are increasing day by day. In this paper, a method is proposed by including node properties to neighbourhood matrix, in order to see the structures of weighted adjacency matrix that defines the relationship between the nodes. In accordance with this proposal, the relationship between the faculties of Turkish universities is studied according to student preferences. Weighted adjacency matrix between faculties is composed based on the frequency of faculty preference of students. By using the properties of faculties, this matrix is multiplied by the adjacency matrix, calculated by Squared Euclidian Distance. The weighted adjacency matrix of the faculties is compared with the re-calculated weighted adjacency matrix. It is observed that the relations between faculties are turned out to be more meaningful in new weighted neighbourhood matrix which is multiplied by Squared Euclidean Distance.

Keywords: weighted social networks, adjacency matrix, social network analysis.

1. Introduction

While working on social networks, most models assume that the links between nodes are equal. An adjacency matrix consists of values 0 and 1. However, in real life, links are not equal in most social, biological, ecological and economic systems. [11], [4] Girvan and Newman [5] repeated their weighted network study conducted with 16 apes in 2004 taking into consideration whether apes groomed each other or not. In a non-weighted social network adjacency matrix, attention was paid only to whether apes groomed each other or not, and only the values of 0 and 1 were included in the matrix. In a weighted social network matrix, on the other hand, information on how many times the apes groomed each other during the observation period was also included. In addition, in the tree diagram formed with a non-weighted adjacency matrix, no societal structure was observed for apes. Conversely, clear structures appeared in the tree diagram created with a weighted adjacency matrix. It can be seen that female and male apes were assigned to different groups in terms of their grooming behaviour [7], [1]. In social networks, the weights of links can be the frequencies of connections between the nodes; in road or air transport network, it can be the distance between the nodes or the time spent on travelling from one node to the other in the distance between the two nodes [8], [10], [6]. In addition, as in the application of Cai et al. [2], correlation or distance matrixes between nodes can also be used as the weights of links.

In this study, the data obtained during transition to higher education in Turkey have been used. In Turkey, students graduating from high school take the exams organized by the Measurement, Selection and Placement Center (Ölçme, Seçme ve Yerleştirme Merkezi, ÖSYM) and aim to be placed in one of the universities with their exam scores. The minimum scores of departments are determined by the students who were placed in that department in the previous year. In the current year, students take into consideration these scores and apply to ÖSYM by adding the universities/departments they want to be placed to their preference lists. Each student has the right to make 30 choices. In this application, these preference lists were utilized. Here departments in universities combine to form faculties. In each faculty, which represents natural clusters, there is a different number of departments.

In this paper, while evaluating the weights of the links between clusters (faculties), emphasis will be put on the fact that the frequency of being seen together in preference lists did not reflect the real link weights completely due to the difference in node numbers within clusters; therefore, a different weight calculation method is recommended. In natural clusters formed within systems, it can be clearly seen in the study of Wang et al. [9] on multi-weighted links that the evaluation of the links and weights both between and within clusters is essential.

2. Higher education preference network model

In the study, the preference lists of 1000 students who took the exams necessary for transition to higher education in 2010 and were placed in one of the undergraduate programs were examined. Faculties in the preference lists were divided into 7 geographical regions so as to allow for classification and make a better evaluation of the networks which would be drawn. In this paper, the analysis made for the Mediterranean Region is used and comparisons are made. From the preference lists, data belonging to 68 different faculties in the Mediterranean Region were obtained.

In the study, a non-weighted social network is being established with the adjacency matrix which shows whether two faculties are included in the same preference list or not.

$$a_{ij} = \begin{cases} 0, & \text{link between nodes i and j} \\ 1, & \text{no link between nodes i and j} \end{cases} \tag{1}$$

$$A_{ij} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \rightarrow \quad \text{Non-weighted adjacency matrix} \tag{2}$$

However, the weighted social network matrix which will be created with information on the frequency of two faculties appearing together in the preference lists will provide more detailed information.

$$a_{ij}^{w} = a_{ij} * w_{ij}; \quad w_{ij} = \text{The frequency of two faculties appearing together in the same list} \tag{3}$$

$$A_{ij}^{w} = \begin{pmatrix} a_{11}^{w} & a_{12}^{w} & \cdots & \cdots & a_{1n}^{w} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ a_{n1}^{w} & a_{n2}^{w} & \cdots & \cdots & a_{nn}^{w} \end{pmatrix} \quad \text{Weighted adjacency matrix} \tag{4}$$

The network which was structured based on the information on coexistence in preference lists of nodes only for the Mediterranean Region is given in Figure 1. The figure shows that the link between some of the nodes is drawn with thicker lines, compared to others. The thickness of the links seen in the figure is proportional to the coexistence in the preference lists. Some sample node couples with a strong link have been examined.

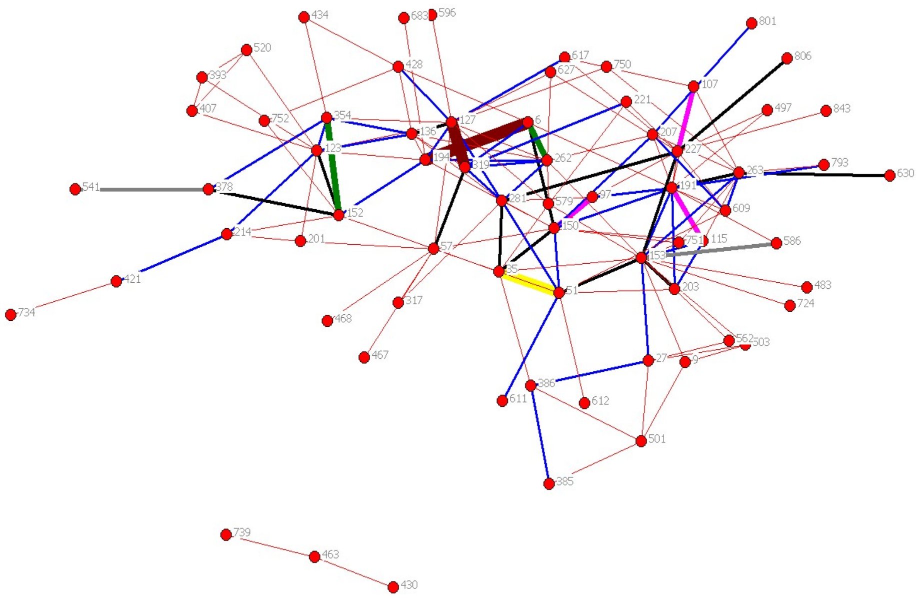

Figure 1. Network weighted with frequency for the Mediterranean Region

a_{ij} is the existence or absence of a link between two nodes (faculties). If there is a link between two nodes, $a_{ij} = 1$; otherwise, $a_{ij} = 0$. This information shows whether two faculties coexist in the same preference list. Figure 1 is drawn according to the calculated a^{w}_{ij} values. a^{w}_{ij} values show whether two faculties coexist in the same preference list; however, they also consist of information on the number of preference lists in which they coexist. a^{w}_{ij} not only shows whether there is a link between two faculties but also displays the level of the existing link value. The bigger the link value, the thicker the link line.

Table 1. Node couples with a strong link in a frequency-weighted adjacency matrix for the Mediterranean Region

Node couple	Explanation
6–194	Mustafa Kemal University, Faculty of Education–Çukurova University, Faculty of Education
127–319	Mehmet Akif Ersoy University, Faculty of Education–Akdeniz University, Faculty of Education
35–51	Çukurova University, Faculty of Science and Letters–Mersin University, Faculty of Science and Letters
152–354	Çukurova University, Faculty of Economic and Administrative Sciences–Mersin University, Faculty of Economic and Administrative Sciences
6–262	Mustafa Kemal University, Faculty of Education–Mersin University, Faculty of Education

When the node couples given in Table 1 are examined, it can be seen that a faculty of education has the most frequent coexistence with other faculties of education, and a faculty of economic and administrative sciences has the most frequent coexistence with other faculties of economic and administrative sciences. It is not surprising that such faculties with several departments and a variety of minimum scores frequently appear in preference lists together.

However, when examined in terms of the strength of their coexistence, more information is needed related to nodes. Everybody knows that it is natural that faculties of medicine, dentistry, and pharmacy which have only one department that can be preferred and which have very close areas of profession are preferred together. Therefore, strong links between these and similar faculties should not be neglected.

At this stage, it was believed that the distance matrixes used in clustering analysis could be utilized. It was planned to use the adjacency matrix obtained from different characteristics of faculties other than appearing together in preference lists in calculating link weights.

One of the widely-used distance measures developed for quantitative data is Euclidean distance. Euclidean distance is found by taking the square root of the squares of differences between *i* and *j* objects in each dimension. The Euclidean distance between two points in a *p* dimension space, x_i and x_j, is calculated as follows: [3]

$$d_{ij}(x_i, x_j) = \left(\sum_{k=1}^{p} |x_{ik} - x_{jk}|^2\right)^{1/2} \quad (5)$$

The distance obtained by taking the square root of Euclidean distance is called squared Euclidean distance.

Some of the reasons that could explain the preference for faculties were determined, and an adjacency matrix squared Euclidean distance was calculated and formed between faculties based on these characteristics. The variables used in the calculation of the adjacency matrix and their descriptive statistics are given in Table 2. Due to the differences in the mean and standard deviation, the data were standardized between 0 and 1 before the adjacency matrix was calculated.

Table 2. Descriptive statistics of node characteristics

Variable	Mean	ss
Quota	73.7	41.1
Minimum score	373.3	80.5
Maximum score	421.9	65.2
Average rank in preference	12	3.3
Vacant quota	1.8	8.4
Number of male students	33.2	22.1
Number of female students	38.7	23.6

Starting from here, the adjacency matrix is formed as follows:

$$A_{ij}^{d} = \begin{pmatrix} a_{11}^{d} & a_{12}^{d} & \cdots & \cdots & a_{1n}^{d} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ a_{n1}^{d} & a_{n2}^{d} & \cdots & \cdots & a_{nn}^{d} \end{pmatrix} \quad (6)$$

The values in A^d_{ij} matrix formed with the calculation of the distance between two faculties by using the variables which are given in Table 2 show the distance between the i^{th} and the j^{th} faculty. The smaller the a^d_{ij} value, the closer the two faculties in terms of the relevant variables.

When both matrixes were calculated, the adjacency matrix weighted by frequency was treated with a weighting procedure again with the adjacency matrix calculated by using the characteristics of nodes. The matrix recommended for the weighted adjacency matrix to be used with the networks that have interbedded social structures is as follows:

Adjacency matrix weighted with Euclidean distance

$$A^{fd}_{ij} = \begin{pmatrix} (a^{fr}_{11} \times a^{d}_{11}) & (a^{fr}_{12} \times a^{d}_{12}) & \cdots\cdots & \cdots & (a^{fr}_{1n} \times a^{d}_{1n}) \\ \vdots & \vdots & \cdots\cdots & \cdots & \vdots \\ (a^{fr}_{n1} \times a^{d}_{n1}) & (a^{fr}_{n2} \times a^{d}_{n2}) & \cdots\cdots & \cdots & (a^{fr}_{nn} \times a^{d}_{nn}) \end{pmatrix} \quad (7)$$

Each cell of the adjacency matrix obtained from the characteristics of faculties was multiplied by the same cell in the adjacency matrix of coexistence frequencies of faculties, and a new adjacency matrix was formed. The network drawn with this new matrix can be seen in Figure 2.

The network drawing of the matrix weighted with the adjacency matrix obtained by using different characteristics of nodes can be seen in Figure 2. First of all, the comparison between the two figures shows that no change occurred in the structure of the network after reweighting. The only difference that occurred between the networks was in the link weights between the nodes. Some node couples with a strong link in the network obtained with a reweighted matrix are given in Table 3.

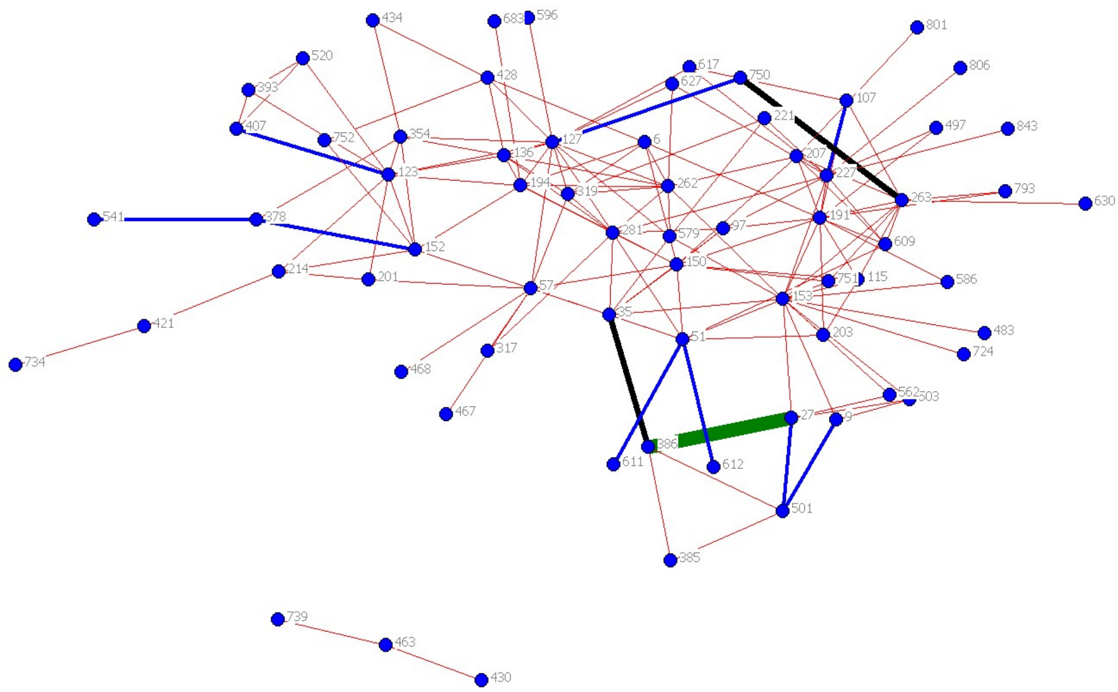

Figure 2. Network reweighted with Euclidean distance for the Mediterranean Region

As can be seen in Table 3, strong links occurred between different faculties in the network that were obtained with a reweighted matrix. Here the link between the faculties with no variety in department and minimum score such as medicine, dentistry and pharmacy appeared to be stronger. In this reweighted network, the strength of the links in the faculties of economic and administrative sciences became somewhat weaker but still remained strong. However, it is clear that the strong link between such faculties as medicine, dentistry and pharmacy, which are well-known for their coexistence in the preference lists of students, could not be revealed with the adjacency matrix which was obtained with the number of coexistence in preference lists.

Table 3. Node couples with a strong link in the adjacency matrix reweighted with Euclidean distance for the Mediterranean Region

Node couple	Explanation
27–386	Çukurova University, Faculty of Dentistry–Çukurova University, Faculty of Medicine
35–386	Çukurova University, Faculty of Science and Letters–Mersin University, Faculty of Medicine
263–750	Mustafa Kemal University, Faculty of Agriculture–Çukurova University, Adana Health College
9–501	Mersin University, Faculty of Pharmacy–Mersin University, Faculty of Medicine
27–501	Çukurova University, Faculty of Dentistry–Mersin University, Faculty of Medicine

The strength of the link between the nodes in the social network that occurred in the course of the analysis of the adjacency matrix weighted with frequency and the strength of the link between the nodes in the social network that occurred with the adjacency matrix reweighted with Euclidean distance appeared to be different. When this dissimilarity was evaluated, it became clear that the links between two nodes obtained with the adjacency matrix weighted with frequency were less strong than those obtained with the adjacency matrix reweighted with Euclidean distance. At this point, the relation between the faculties which appeared to have a strong link due to the multitude of departments was subject to a "correction" operation using node characteristics. Thus, it was ensured that faculties with a single department but strong links occur.

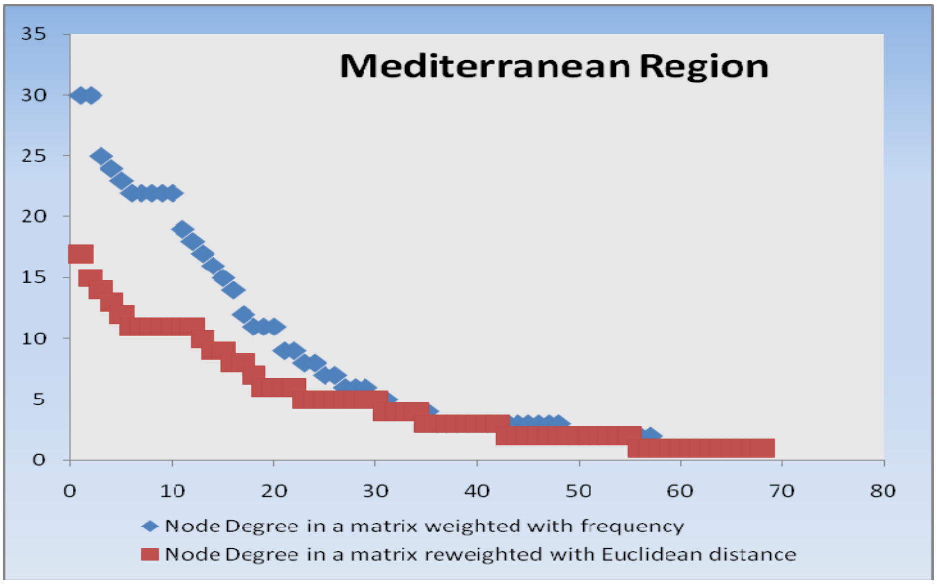

Figure 3. Ranks of the nodes of two weighted matrixes for the Mediterranean Region

When the ranks of the nodes of two matrixes are examined for the Mediterranean Region, it can be seen that they both comply with the laws of force. Therefore, it can be said that they are appropriate for the model irrespective of the scale. The fact that no significant change occurred in rank distribution makes it clear that reweighting was concentrated in correlation weights only.

3. Conclusions

The strength of the link between nodes in social networks is an issue which is as important as the existence of that link. The more information these network weights between nodes include, the more realistic the evaluations and the resulting strong and weak links will be. When social networks are examined, it can be seen that the strong links obtained with a weighted adjacency matrix are between university–faculty nodes that include different departments, which are in abundance in our country. The evaluation based only on these findings and revealing the structure of higher education institutions in Turkey at university–faculty level with network analysis will not be an accurate approach.

On the other hand, strong links obtained with a reweighted matrix with Euclidean distance showed some differences as several characteristics of faculties were taken into consideration, in addition to being due to the variable which can be called “abundance in number”. When these strong links are examined, it can be seen that the university–faculty nodes which are well-known with their strong links according to their coexistence in preference lists come to the forefront. Due to this “pioneering” position of frequency abundance, it was witnessed that this new recommendation was successful in revealing the nodes with strong links which were excluded.

This study was conducted only on one sector, which is an aspect open to development. Further studies can deal with comparisons in different sectors or areas using an *adjacency matrix reweighted with Euclidean distance.*

References

1. Albert, Reka, Barabasi, Albert-Laszlo, “Statistical Mechanics of Complex Networks”, *Reviews of Modern Physics*, Vol. 74, No. 1, Ocak 2002.
2. Cai, S., Zhou, Y., Zhou, T., Zhou P., “Hierarchical Organization and Disasortative Mixing of Correlation-Based Weighted Financial Networks”, *International Journal of Modern Physics*, Vol. 21, No. 3, 2010.
3. Çilingirtürk, A.Mete, *İstatistiksel Karar Almada Veri Analizi*, 1. Basım, Seçkin Yayıncılık, Ankara, 2011.
4. Duncan, J., Watts, Steven, H., Strogatz, “Collective Dynamics of ‘Small-World’ Networks”, *Nature*, Vol.393, Haziran 1998.
5. Girvan, M., Newman, M. E. J., “Community Structure in Social and Biological Networks”, *Proc. Natl. Acad. Sci*., USA, Vol. 99, 2002.
6. Gürsakal, Necmi, *Sosyal Ağ Analizi*, Dora Yayıncılık, 1. Basım, Bursa, 2009.
7. Newman, M. E. J., “Analysis of Weighted Networks”, *Phys. Rev. E*., Vol.70-5, 2004
8. Newman, M. E. J., *Networks: An Introduction*, OXFORD University Press, 1. Basım, New York, 2010.
9. Wang, H., Boven, E. Van, Krishnakumar, A., Hosseini, M., Hooff, H. Van, Takema, T., Baken, N., Mıeghem, P. Van, “Multi-Weighted Monetary Transaction Network”, *Advance in Complex Systems*, Vol. 14, No. 5, Mayıs, 2011.
10. Wasserman, Stanley, Faust, Katherine, *Social Network Analysis: Methods and Applications*, Cambridge University Press, 13, Basım, New York, 2005.
11. Yook, S. H., Jeong, H., Barabasi, A. L., Tu, Y., “Weighted Evolving Networks”, *The American Physical Society*, Vol. 86, No. 25, 2001.

SOCIALINIS FAKULTETŲ TINKLAS PAGAL STUDENTŲ PASIRINKIMUS, SIEKIANT AUKŠTESNIO IŠSILAVINIMO

Güneş Mutlu, Ahmet Mete Çilingirtürk

Santrauka. Pastaruoju metu socialinių tinklų analizėje vis daugėja svorinės viršūnių gretimumo matricos tyrimų. Šiame straipsnyje siūloma į gretimumo matricą įtraukti viršūnių savybes, siekiant atskleisti svorinės gretimumo matricos, apibrėžiančios ryšius tarp viršūnių, struktūras. Taip pat, remiantis studentų pasirinkimais, tiriami sąryšiai tarp fakultetų Turkijos universitetuose. Svorinė fakultetų gretimumo matrica sudaroma remiantis fakultetų dažniais studentų pasirinkimų sąrašuose. Naudojantis fakultetų savybėmis, ši matrica dauginama iš gretimumo matricos, sudarytos skaičiuojant Euklido atstumo kvadratą. Svorinė fakultetų gretimumo matrica lyginama su perskaičiuota svorine gretimumo matrica. Pastebėta, kad prasmingesni ryšiai tarp fakultetų gaunami naujoje, iš Euklido atstumo kvadrato padaugintoje svorinėje gretimumo matricoje.

Reikšminiai žodžiai: svorinis socialinis tinklas, gretimumo matrica, socialinio tinklo analizė.