Journal of Statistical and Econometric Methods, Vol. 9, No. 4, 2020, 1-13
ISSN: 2241-0384 (print), 2241-0376 (online)
Scientific Press International Limited

Hidden Power of Communication: Understanding Organizational Structure in Bank Branches: Ziraat Bank Example

Dr. İlker MET[1], Ayfer ERKOÇ[2], Himmet AKSOY[3] and Hacer Özge ÖZER[4]

Abstract

By all means, having a well-functioning organizational communication channel is a must to maintain the corporate culture and align with the organizational strategy and goals. In order to improve communication levels in an organization, it is crucial to know the communication habits and levels of employees. Organizational Network Analysis (ONA) is a structured way to visualize how communication, information, and decisions flow through an organization, and the nature of communication patterns among employees. "A picture is worth a thousand words" is one of the most commonly used phrases. On the other hand, a graph is worth much more than that. "A visual representation of data, in the form of graphs, helps human resources professionals gain actionable insights and make better data driven decisions based on them". The role of the team leader is important in increasing communication between the teams .In this study, the dimensions of organizational communication between the managers and employees in 1.642 banking branches has been analysed with the ONA. Then 12 metrics were calculated. Using the correlation analysis, the highest relationship among these metrics was selected. Branches with similar characteristics were grouped by cluster analysis. It is seen that the cluster with the best characteristics is also the branches with high performance.

JEL classification numbers: D85, C15, C81, J24, J21
Keywords: Organizational Network Analyses, Fruchterman-Reingold Algorithm, Clustering, Correlation Analyses, ONA Metrics. R Programming, People Analytics.

[1] Ziraat Bank, Ankara, Turkey.
[2] Ziraat Bank, Ankara, Turkey.
[3] Ziraat Bank, Ankara, Turkey.
[4] Ziraat Bank, Ankara, Turkey.

Article Info: *Received:* May 24, 2020. *Revised:* June 11, 2020.
Published online: August 30, 2020.

1. Introduction

In recent years, there has been an increased interest in network research in physical and social sciences. For researchers, networks theory has become significant, providing explanations of social phenomena in a wide variety of disciplines, from psychology to the economy [1].

Social network analysis has been used since the mid-1930s to advance research in social and behavioural sciences. Sociometry (sociograms, sociomatrices), graph theory, dyads, triads, subgroups, and block models, reflecting substantive concerns such as reciprocity, structural balance, transitivity, clusterability, and structural equivalence, made their appearances and are quickly adopted by the relatively small number of "network analysts." It is easy to trace the evolution of network theories and ideas from professors to students, from one generation to the next. The field of network analysis is even analyzed as a network [2]. The ONA puts together multiple data sources to help companies understand how, where, and which teams work together. Which teams need to communicate more to achieve company goals? The Network analysis also simplifies data-driven decision making for process improvement. Beside network analysis gave idea to the managers about performance indicators, studies show that relationships also affect the organizational commitment [3].

ONA is a growing trend in the field of People Analytics in Human Research studies. New developments in social sciences highlight social networks as phenomena that can motivate people and change lives [4]. One of the factors that make social networks so popular is the academic studies. Besides graph methods, various metrics make the network easier to understand and interpret. Studying with graphic metrics through statistical models makes measurement easier. Learning how to map social connections helps managers to use real power in their companies and renews their formal organization to allow the development of informal organizations [5].

Network data have been obtained via surveys and questionnaires, archives, observation, diaries, electronic traces, and experiments. Further research on data quality is needed. Beyond improved samples and further investigation of the informant accuracy/reliability issue, this should cover common indices of network structure, address the consequences of sampling portions of a network, and examine the robustness of indicators of network structure and position to both random and non-random errors of the measurement [6]. In recent years, it is preferred to use the log records and electronic traces to analyses the relationship with the data. The main reason to shift preferences to this way, stop interrupting the employee from dividing with questionnaire methods, prevents biased answers and understands objectively [7].Organizational network analysis (ONA) can be a powerful means of making invisible patterns of information flow and collaboration in strategically important groups visible [8].

ONA is an application of social network analysis, a method that is typically focused on connections between individuals, to an organizational entity. It is a descriptive, empirical research method for mapping and measuring relationships between people, groups, and organizations with the resources, knowledge and tasks that are used to perform work. The key feature distinguishing network theory and measurement from traditional data analytic methods is the use of structural or relational variables analyzed using techniques based on graph theoretic methods [9]. The resulting insights can help managers understand critical performance factors such as how information diffuses among individuals and influences the speed, quality, and accuracy of organizational decisions [10]. ONA techniques can provide empirical data to plan for and justify the allocation of resources as well as aid decision-making by revealing links between information networks and process performance [11].

For ONA the base requirement is metadata from digital communication systems such as e-mail, and calendar appointments (MS Exchange, GSuite, etc.), chat, and collaboration (Skype, Slack, etc.). Before the analysis of the metadata, anonymize (hash) all email addresses, and remove all communication content, so only metadata - such as sent/received timestamps - is collected. Then start work to map the anonymized email addresses and the demographic information, such as team, department, region, etc.

There are special tools for mapping and analyzing networks such as Gephi and Cytoscape. R Programming is also preferred because of its capabilities for calculating detailed statistics and metrics.

2. Research Method

This paper focused on the following research questions:

- Where an organization is most siloed?
- Which branches are collaborating well?
- Where the bank can invest optimally to improve performance through enhanced collaboration?
-

For this purpose analysis steps are designed as follows:

1. Collecting data,
2. Creating maps for each branch with R programming,
3. Calculation of ONA metrics,
4. Cluster analysis using metrics,
5. Interpretation of the results.

2.1 Collecting Data

3-months communication data is used in the analysis. Data consist of e-mail, Skype, and VoIP phone record. In the network analysis content isn’t important. So the data were anonymized trough algorithms and the metadata is transformed into a format only with whom and how many times they had the conversation.

2.2 Creating Maps

After data preparation, maps are created for each branch. In this section igraph package for R programming is used [12].

The strength of R in comparison to stand-alone network analysis software is threefold. In the first place, R enables reproducible research that is not possible with GUI applications. Secondly, the data analysis power of R provides robust tools for manipulating data to prepare it for network analysis. Finally, there is an ever-growing range of packages designed to make R a complete network analysis “tool”. The main goals of the igraph library are to provide a set of data types and functions for, pain-free implementation of graph algorithms, fast handling of large graphs, with millions of vertices and edges, allowing rapid prototyping via high-level languages like R.

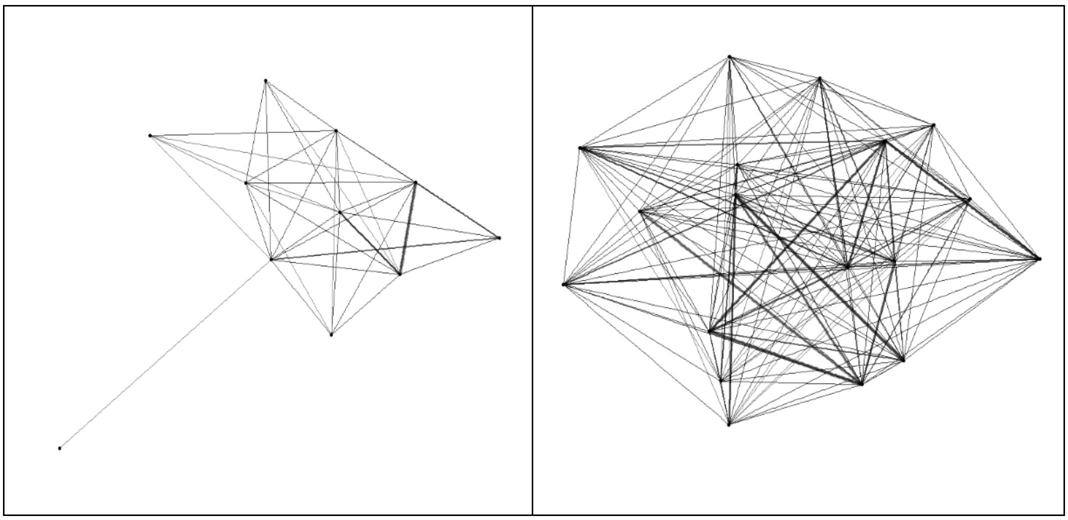

Figure 1: Branch’s Organizational Network Examples

As seen in Figure 1, branch network graph on the left shows that there are employees with disconnected communication. The ONA map gives clues. Not all employees in the branch are in relation. So communication density is weaker than the right branch. Branch 2 (on the right) outperforms well than Branch1 (on the left), which is visible in the ONA map for Branch 1 where there are weak connections and outliers.

The Fruchterman-Reingold Algorithm is a force-directed layout algorithm. The idea of a force-directed layout algorithm is to consider a force between any two nodes. In this algorithm, the nodes are represented by steel rings, and the edges are springs between them. The attractive force is analogous to the spring force and the repulsive force is analogous to the electrical force. The basic idea is to minimize the energy of the system by moving the nodes and changing the forces between them. In this algorithm, the sum of the force vectors determines which direction a node should move. The step width, which is a constant, determines how far a node moves in a single step. When the energy of the system is minimized, the nodes stop moving and the system reaches its equilibrium state [13].

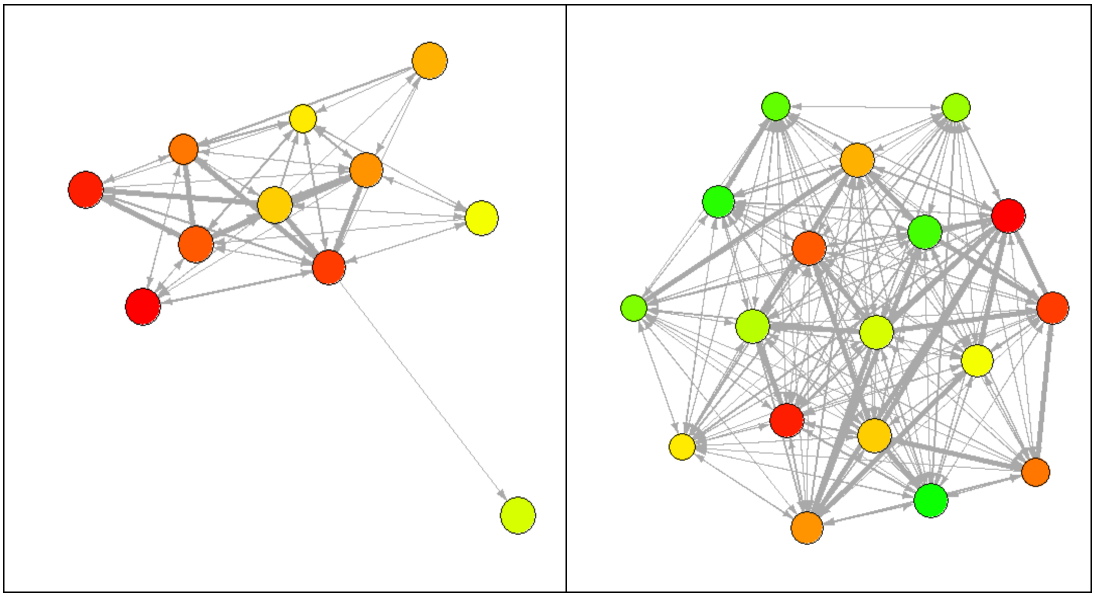

Figure 2: Branch's Organizational Network Examples with Fruchterman-Reingold Algorithm

Fruchterman-Reingold Algorithm can be created network maps that shows weak/strong relationships. The same branches' in Figure 1 ONA maps are created by this algorithm. As can be seen easily in Figure 2, relations are weaker in the Branch2 (on left). Branch2 (on the right side) outperforms well than Branch1 (on the left) which is visible in the ONA map for Branch 1 where there are weak connections and outliers.

2.3 Network Statistics

Networks are commonly used to study complex systems and statistics which facilitate the interpretation of network. The metrics used in the study are listed below:

- **Degree**

It means the number of direct links with other actors. It can be analyzed in different ways as in degree, out degree and total degree.Degree centrality (degree in and degree out), betweenness centrality and closeness centrality are commonly used centrality measures [14].

- **Betweenness**

It measures the presence of an actor in the network and shows whether the actor is the bridge between other actors or not. High centrality means that the actor is a mediator in the group.

- **Closeness**

The closeness indicates the average distance between an actor and all other actors in the network.

- **Coreness**

Coreness measures the importance of actors in the network structure and shows which of the actors are in the centre.

- **Eigen Centrality**

Eigenvector centrality (Eigen centrality or prestige score) is a measure of the influence of a node in a network.

- **Page Rank**

This approach appears as Google's Page Rank technology. The high level of this metric indicates that the node is the most communicated person in the network.

- **Transitivity**

Transitivity is the overall probability for the network to have adjacent nodes interconnected, thus revealing the existence of tightly connected communities (or clusters, subgroups, cliques). A social preference to be friends with your friends' friends [15].

- **Reciprocity**

Reciprocity can be explained by a tendency for relationships. **The metric** is a measure of the likelihood of vertices in a directed **network** to be mutually linked [16].

- **Ego Size**

These functions find the vertices not farther than a given limit from another fixed vertex. These are called the neighbourhood of the vertex.

- **Assortativity**

Assortativity is often operationalized as a correlation between two nodes. Positive values of r indicate a correlation between nodes of similar degrees, while negative values indicate relationships between nodes of different degrees [17].

- **Authority Score**

Authority score is a natural generalization of eigenvector centrality. Generally speaking the higher the authority score of a domain, the more trusted it is.

- **Hub Score**

Hub is a natural generalization of eigenvector centrality. The idea of a hub score is that a good hub points good authorities. Hubs and authorities are a natural generalization of eigenvector centrality [18].

Finally edge density per branch is calculated. The density of a graph is the ratio of the number of edges and the number of possible edges. The 12 metrics described above are calculated for 1.642 branches. Due to the high number of branches, cluster analysis is performed to identify the branch managers with the weakest communication. Before proceeding to cluster analysis, statistics with the highest correlation are determined related variables are used in cluster analysis. Other metrics are used as decision support points.

2.4 Correlation Tests

For correlation analyses, the Pearson correlation is used. Pearson correlation can be used only when x and y are from the normal distribution. So all metrics except 0-1 value are scaled to normality.

Table 1: Correlation Test Results

Row	Column	Core	P-value
eigen centrality	authority score	0.96	0.00
ego size	coreness	0.95	0.00
coreness	assortativity	0.77	0.00
ego size	assortativity	0.77	0.00
page size	closeness	0.70	0.00

Table 1 shows the variables with the highest correlation and the p-value is < 5% shows that the correlation between the metrics is significant. In cluster analyses, eigen centrality, authority score, ego size, and coreness metrics are used.

2.5 Cluster Analysis

Firstly, it is tested whether the data set is suitable for clustering or not. In this section Hopkins statistic is used. The statistic is calculated as the mean k-nearest neighbor distance in the real data-set divided by the sum of the mean k-nearest neighbor distances in the real and across all the simulated data-sets. A value greater than 0.5 indicates clustered data [19]. The Hopkins stat is calculated as 0.99. So the dataset is suitable for classification.

Clusters are created by R programming and cluster maps are shown below:

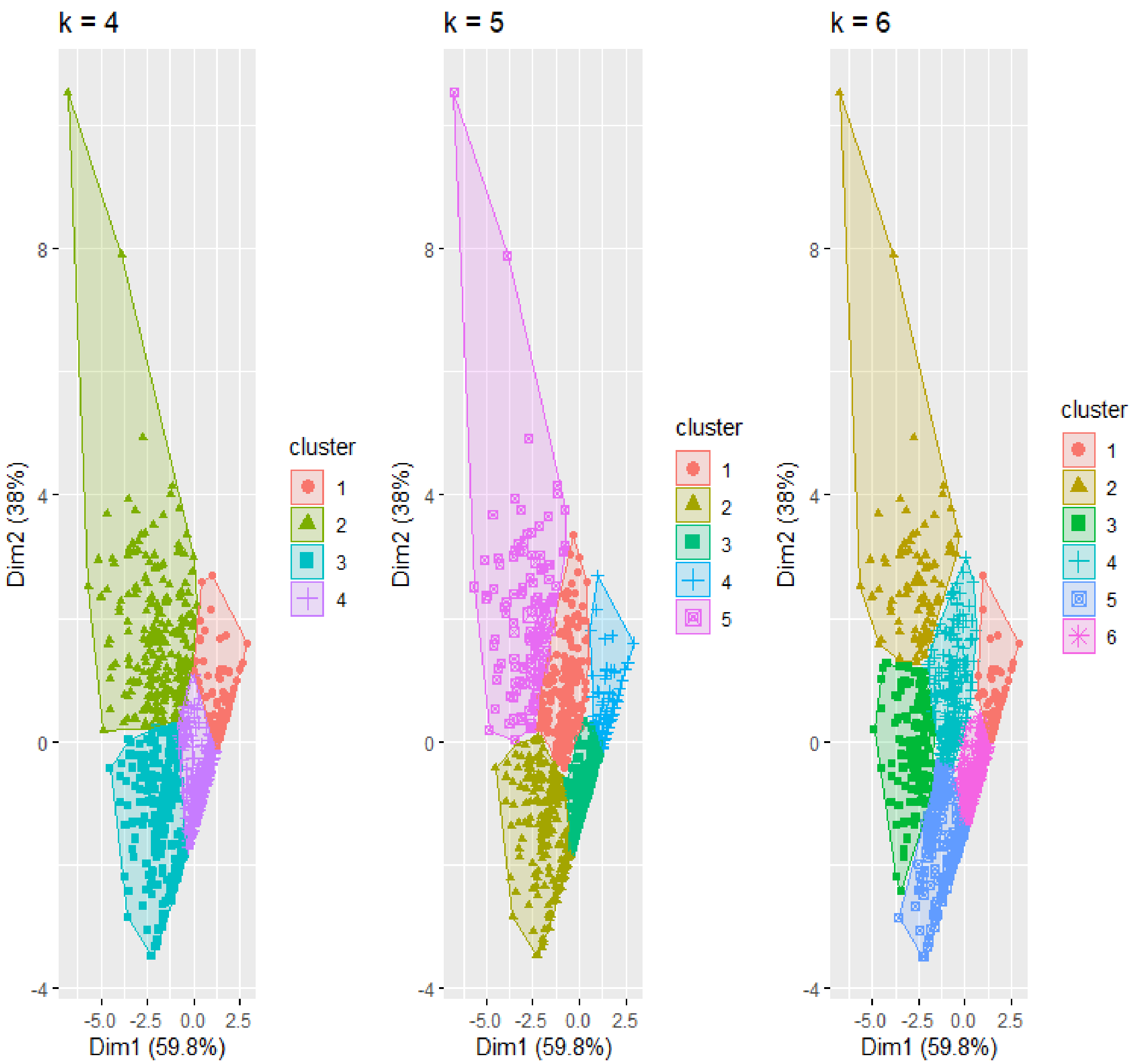

Figure 3: Cluster Maps

To decide the number of cluster mean of silhouette value and BSS (between_SS)/TSS (Total_SS) ratio are used.

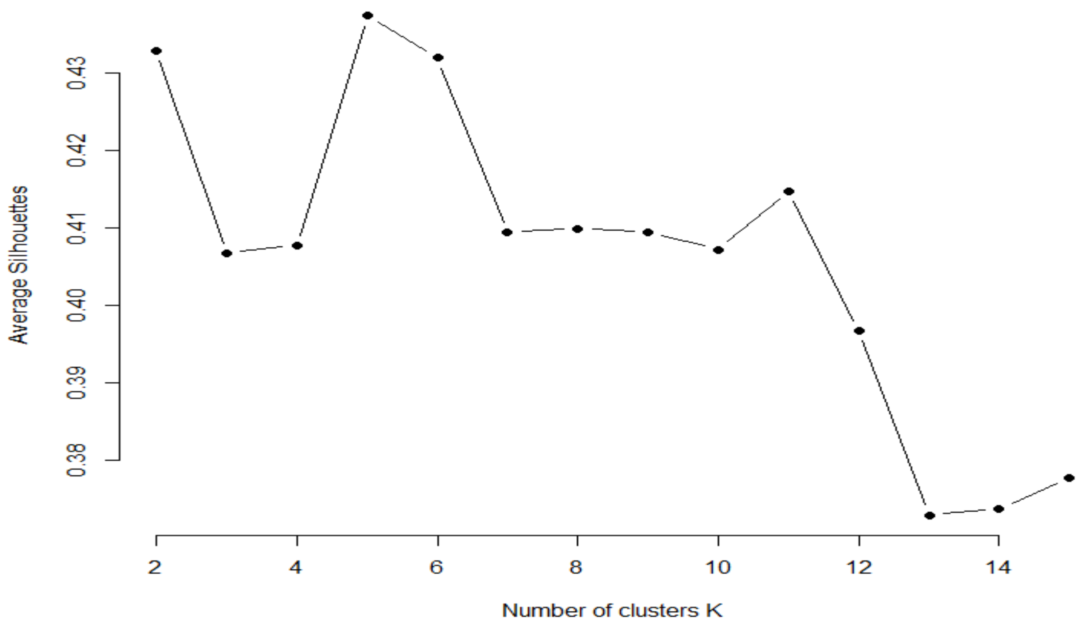

Figure 4: Average Silhouettes Values for Number of Clusters

The average silhouette approach measures the quality of a clustering. In oth er words, it determines how well each object lies within its cluster. A high average silhouette width indicates a good clustering.

In clustering, the goal is usually to get high similarity within each group and low similarity between each group [20]. For this purpose between_SS / total_SS is calculated. It's a measure of the goodness of the classification k-means has found. SS stands for the sum of squares, so it's the usual decomposition of deviance in deviance "between" and deviance "within." Ideally, a clustering that has the properties of internal cohesion and external separation, BSS/TSS ratio should approach 1.

Measure of the goodness of the classification (between_SS / total_SS) for 6 cluster is 81. 5% cluster is 77.2%.

Considering the number of branches, silhouette, and between_SS / total_SS, data is separated into 6 clusters.

Table 2: Cluster Summaries

Clusters	Count of manager	*Branch density	Hubs-score	Transitivity	Reciprocity	Page-rank	Eigen-centrality	Ego-size	Authority-score	Assortativity	Betweenness	Closeness	Coreness
1	254	0,90	0,97	0,96	0,94	0,15	0,90	10,15	0,89	-0,10	1,27	0,11	15,01
2	103	0,88	0,95	0,96	0,93	0,12	0,78	10,18	0,76	-0,10	0,88	0,11	14,88
3	377	0,97	0,98	0,99	0,98	0,26	1,00	5,37	0,99	-0,26	0,26	0,26	8,19
4	538	0,94	0,98	0,98	0,97	0,17	1,00	10,03	0,99	-0,11	1,15	0,11	15,91
5	255	0,90	0,99	0,97	0,95	0,12	0,98	15,08	0,97	-0,07	2,81	0,07	22,90
6	115	0,86	0,96	0,95	0,92	0,09	0,86	16,1	0,85	-0,04	3,26	0,06	22,63
Means	1642	0,93	0,98	0,97	0,96	0,17	0,96	10,2	0,95	-0,13	1,36	0,14	15,49

*Branch density shows managed branch's edge density.

When the clusters summarizing in Table. 2, cluster2, and cluster6 have almost lowest ratios. Cluster four is the best cluster and can be said that these branches are managed well and there is no communication problem in the branch.

3. Conclusions

Researches show that to strengthen communication, organizational network analysis can be used to understand organizational structure and hidden communication layers in the firm. This study proposed to bring specific suggestions from organizational network perspectives to human resources professionals that rebalance the organization's attention to employees and to their relationships.

Through organizational network analysis.

- Confidential information flow can be provided.
- Key persons in the organization can be identified.
- Information flow between the unit and people in the organization can be accelerated.
- Innovative processes can be improved.
- Motivation, sustainability, and corporate loyalty can be increased.
- The training needs can be determined, and cost-time loss can be reduced.
- Maximizing relationship management as well as accurate performance measurement.
- Can be used to plan communication flow in chaotic times.
- Can be used to HR's role in fostering a high-performance culture.
- Considerations in creating and managing organizational culture.

In the study, 1.642 branch managers, considering branch actors in the organizational structure of the Bank, are examined. Twelve ONA metrics are used in the analyses and correlation tests are determined which metrics are the most important. The most correlated metrics are eigen centrality, ego size, coreness, page size, authority score, assortativity (correlation rates>=0,7). Combining network analysis with statistical methods makes it easier to interpret the results. Most correlated metrics are used to identify clusters and to determine poorer communication. Cluster 2 and cluster 6 has the lowest average values. These clusters at the same time have poor performance. Then, to improve the performance in the branch, the necessary precautions are taken. Some suggestions for relationship management to human resources professionals are given below:

- Tasks that employees will perform together can be organized.
- Sports events can be organized.
- Mass organizations such as food and picnic can be organized properly.
- Positive morale events can be organized.
- Ensuring the use of social media within the network.
- Making small presentation meetings on social issues.
- Communication training and conferences can be organized.

References

[1] Borgatti, S.P., Mehra, A., Brass, D.J and Labianca, G. (2009). Network Analysis in the Social Sciences. Science, 323(5916), pp. 892-895. https://doi.org/10.1126/science.1165821

[2] Wasserman, S. , Scott, J. , & Carrington, P. J. (2005). Introduction. Models and Methods in Social Network Analysis, 1-7. https://doi.org/10.1017/cbo9780511811395.001

[3] Bozionelos, N. (2008). Intra-Organizational Network Resources: How They Relate to Career Success and Organizational Commitment, Personnel Review, Volume 37 No 3, pp. 249-263. https://doi.org/10.1108/00483480810862251. (2005). Models and Methods in Social Network Analysis. https://doi.org/10.1017/cbo9780511811395

[4] Westaby, J. D. (2012). Dynamic network theory: How social networks influence goal pursuit. https://doi.org/10.1037/13490-000

[5] Krackhardt, D. and Hanson, J. R. (1993). Informal Networks: The Company Behind the Chart. Harvard Business Review, From the July–August 1993 Issue.

[6] Marsden, P. V. (1990). Network Data and Measurement. Annual Review of Sociology, 16(1), pp. 435-463. https://doi.org/10.1146/annurev.so.16.080190.002251

[7] https://www.humanyze.com. Access date :18.03.2020

[8] https://www.robcross.org. Access date :01.03.2020

[9] Wasserman, S and Faust, K. (1994). Social Network Analysis: Methods and Applications, Cambridge, Cambridge University Press. pp. 825. Italian Political Science Review/ Rivista Italiana di Scienza Politica, 25(3), pp. 582-584. https://doi.org/10.1017/s0048840200023959

[10] Carley, K. M. (2002). Computational organization science: A new frontier. Proceedings of the National Academy of Sciences, 99(Supplement 3), pp. 7257-7262. https://doi.org/10.1073/pnas.082080599

[11] Merrill, J., Bakken, S., Rockoff, M., Gebbie, K. and Carley, K., (2007). Description of a Method to Support Public Health İnformation Management: Organizational Network Analysis. Journal of Biomedical Informatics, 40(4), pp.422-428. https://doi.org/10.1016/j.jbi.2006.09.004

[12] Csardi G, Nepusz T (2006). “The igraph software package for complex network research.” InterJournal, Complex Systems, 1695. http://igraph.org

[13] Fruchterman, T. M. & Reingold, E. M. (1991). Graph drawing by force-directed placement. Software: Practice and Experience, 21(11), 1129-1164. https://doi.org/10.1002/spe.4380211102

[14] Salman, C., (2018). A New Network Centrality Measure: Relative Edge Importance Method (Unpublished master’s thesis). University of Hacettepe, Ankara.

[15] https://www.stat.washington.edu/~pdhoff/courses/567/Notes/l16_transitivity.pdf. Access date: 24.03.2020

[16] Bo, J. (2015). Network Characteristics and Dynamics: Reciprocity, Competition and Information Dissemination. Doctoral Dissertations. pp.443.
[17] Newman, M. E. (2002). Assortative Mixing in Networks. Physical Review Letters, 89(20). https://doi.org/10.1103/physrevlett.89.208701
[18] Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. Journal of the ACM (JACM), 46(5), 604-632. https://doi.org/10.1145/324133.324140
[19] Banerjee, A., & Dave, R., Validating clusters using the Hopkins statistic. 2004 IEEE International Conference on Fuzzy Systems (IEEE Cat. No.04CH37542). https://doi.org/10.1109/fuzzy.2004.1375706
[20] Bolshakova, N., & Azuaje, F. (2003). Cluster validation techniques for genome expression data. Signal Processing, 83(4), 825-833. https://doi.org/10.1016/s0165-1684(02)00475-9.