Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi 2020 8(1) 179–188

Journal of Social Sciences of Mus Alparslan University

anemon

Derginin ana sayfası: http://dergipark.gov.tr/anemon

Araştırma Makalesi ● Research Article

Determination of Influential Countries by Cultural and Geographical Parameters*

Kültürel ve Coğrafi Parametrelere Göre Lider Ülkelerin Tespiti

İhsan Tuğal,[a,*] Ali Karcı [b]

[a] Muş Alparslan University, Faculty of Engineering and Architecture, Department of Computer Engineering, 49250, Muş/Turkey. ORCID: 0000-0003-1898-9438

[b] İnönü University, Faculty of Engineering, Department of Computer Engineering, 44280, Malatya/Turkey. ORCID: 0000-0002-8489-8617

MAKALE BİLGİSİ

Makale Geçmişi:

Başvuru tarihi: 10 Mayıs 2019

Düzeltme tarihi: 22 Ağustos 2019

Kabul tarihi: 03 Eylül 2019

Anahtar Kelimeler:

Ülke Birliktelikleri

Sosyal Hesaplama

Çizge Entropi

Lider Ülkeler

Sosyal Ağ

ÖZ

Ülkelerin kültürel ve coğrafi durumu, ülkeler arasındaki ilişkileri belirleyen en önemli faktörlerdir. Bu çalışmada çizge entropi kullanılarak, sosyal ağ analizi yöntemleri ile lider ülkeler tespit edildi. Ülke benzerlikleri kültürel ve coğrafi parametreler olan din, dil, bölge, kıta bilgileri ile ölçüldü. Benzerlikler kullanılarak ülkelerin ilişki yapısı çizge yapısında çıkarıldı. Sosyal ağ yapısında merkezi ülkelerin tespiti için geleneksel merkezilik ölçümleri yerine Shannon ve Karcı entropi kullanıldı ve elde edilen sonuçlara göre bu iki entropi türü kıyaslandı. Karcı entropi de Shannon'dan farklı olarak bu iki değerin ölçümdeki etkisinin alfa ile ayarlanabildiği gösterildi. Bu çalışmada kullanılan parametrelerle elde edilen sonuçlar ve önerilen yöntemler ile ülkelerin liderliğinin ve etkinliğinin belirlenmesi uluslararası ilişkilerde yeni analizler ve bakış açıları ortaya çıkaracaktır.

ARTICLE INFO

Article history:

Received 10 May 2019

Received in revised form 22 August 2019

Accepted 03 September 2019

Keywords:

Country Associations

Social Computing

Graph Entropy

Influential Countries

Social Network

ABSTRACT

The cultural and geographical situation of countries cover the most important factors determining the relations between countries. In this study, a new method was proposed to identify the influential countries. Similarities were calculated by religion, language, zone, and landmass information of countries. This data was transformed into a graph structure. Central countries were determined by Shannon and Karcı entropy over country similarities. The number of similarities and similarity ratios of each country with other countries were used when identifying the influential countries. The difference of Karcı entropy from Shannon, that the effect of these two values can be adjusted by the alpha. With the parameters used, the determination of countries' leadership will reveal new analysis and perspectives in international relations.

1. Introduction

In his book Clash of Civilizations, Huntington mentions that people's cultural and religious identities will be the primary source of conflict in the Post-Cold War World(Huntington, 2001). Geographical features of the states and their position on the world determines the foreign policy they follow and the effect on the formation of relations. The differences between civilizations constitute the main conflict areas of global politics. In this context, social network analysis was applied by using the cultural and geographical information of the countries. The effect of these parameters on the

*This study is derived from the first author's doctoral dissertation titled "Analysis and application of social networks using graph entropy" under the supervision of Ali KARCI in İnönü University, Graduate School of Natural and Applied Sieciences, Department of Computer Engineering.

** Sorumlu yazar/*Corresponding author.*
e-posta: i.tugal@alparslan.edu.tr

http://dx.doi.org/10.18506/anemon.563211

relations of countries was shown mathematically. The influential countries in the structure formed by these parameters were determined by the proposed method based on entropy. General definitions and concepts were simply explained. The aim is that a researcher from the social sciences can understand it. Because the results of the analysis of the data set were presented to the attention of social scientists.

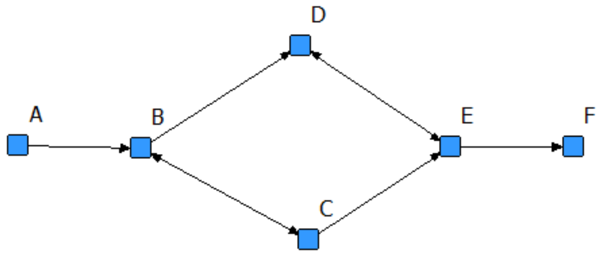
Social network analysis is the measure of relationships and flows between connected entities. Social networks contain a great deal of information about the behavior of the entities and their relationships. Thus, various information can be obtained by analyzing these behaviors and relations. One of the most important purpose for study social network analysis is the identification of hidden organizational structures and important elements that are affecting the system (Chen, Lü, Shang, Zhang, & Zhou, 2012; Shetty & Adibi, 2006). The identification of influential nodes is essential to explore social networks, biological networks, chemical networks, commercial networks, Internet, ecological networks, traffic networks. In this study the efficiency of the countries were calculated by using Shannon and Karcı entropy over country similarities. Religion, language, landmass, and zone information of 193 countries from the "Flags" dataset were used (Lichman, n.d.). The analyses have been carried out by using entropy within the social network structure. The most influential countries and hidden clusters were identified by using the specific parameters of the countries.

This paper is organized as follows. Section 2 begins with a brief overview of the social network analysis and the social network structures. Section 3 describes the graph entropy with country similarities. Section 4 describes the results of the proposed algorithm is implemented to the selected data of 10 random countries. In Section 5, the analysis results of the all Flags dataset are shown. Conclusions and discussions are presented in Section 6.

2. Social Network Analysis

Social networks are commonly used by many researchers. The frequent use of social networks has caused these structures to be used for the data analysis in very different areas. Their analysis and examination have led to the discovery of many new findings. The simplest definition of the social network analysis is to examine social structures and make inferences from the actor's relations within the network. Identification, interpretation, and inferences are usually acquired from the relationship of the actors. It examines the effects of nodes and the links that come from the relationships of the nodes (Gürsakal, 2009). Social network analysis is interested in the relations formed within the network structure. Relational data indicates the relationship between nodes and the node components. From these links, data are being analyzed using various analysis and measurement methods and significant information is obtained (Durland & Fredericks, 2005; Wolfe, 1997). The social network is the mathematical graph structure that is defined as nodes (actor, player), and the ties, edges or links that connect them (Karagöz & R, 2013; Scott, 2000; Wasserman & Galaskiewicz, 2012).

The mathematical representation of social networks can be in the form of different definitions and models. One of these notations is the graph structure that is used in this study. Figure 1 shows that individuals or entities are considered as nodes and links as edges (S P Borgatti, Everett, & Freeman, 2002). The adjacency matrix is a matrix in which the row (i) and column (j) represent a node and the cell indicated by (ij) represents the connection between i and j nodes. If there are n nodes in the specified network, the adjacency matrix becomes a square matrix with $n * n$ dimension as shown in Figure 1.



$$Nodes(\boldsymbol{V}) = \{A, B, C, D, E, F\}$$

$$Edges\ (\boldsymbol{E}) = \{(A, B), (B, C), (B, D), (C, B), (C, E), (D, E), (E, D), (E, F)\}$$

Rows	Columns: A	B	C	D	E	F
A	0	1	0	0	0	0
B	0	0	1	1	0	0
C	0	1	0	0	1	0
D	0	0	0	0	1	0
E	0	0	0	1	0	1
F	0	0	0	0	0	0

Figure 1. Directed graph with adjacency matrix

In fact, people and beings are connected by various bonds. There are many connection types. For example, a marriage and business relationship between families (Padgett, 2010), a structure created by e-mail communication or instant messaging applications is actually a link and a social network. We are all in any local community and we are only a few "links" away from anywhere in the world. We need to look at how we construct these networks and how it affects our lives. When these structures are examined, it has been observed that they exist in every area of life. They are detailed, complex, and common. How does it occur? How does it develop over time? How does it work? What might be the effects of the changes? We must provide enough answers to these aforementioned questions. The best part of these structures is that they can be expressed mathematically, and their calculations are simple as the graph is suitable for the data structure (Gürsakal, 2009).

3. Graph Entropy with Country Similarities

Entropy is a quantitative balance indicator and a measure of the irregularity which is generally expressed with the second law of the thermodynamics. According to this law entropy either remains constant or increases. It expresses the uncertainty or change within the system. It is mainly used to analyze energy transfer within an isolated system. When the disorder is big, the entropy of the system is high (Raymond, 2005). Systems that do not allow outside intervention can be defined as irregular systems. The entropy of these systems is continuously increasing. The law of entropy means that without additional energy and without a certain cost, there will be no continuity of process with the same energy. After Physic, Shannon entropy (Shannon, 1951) is used to measure uncertainty in communication. This measure is based on the probabilistic uncertainty, so it is begun to use in many different disciplines, such as social network.

Usually centrality methods like degree centrality, closeness centrality, betweenness centrality are used to understand the importance of the nodes. Centrality measures makes visible these invisible patterns of graphs (Stephen P. Borgatti, 2005; Freeman, 1978). But in this paper entropy centrality methods used for solutions.

Recently, many different entropy calculation methods have been proposed and implemented in the graphs. Fei and Deng (Fei & Deng, 2017) addressed the problem of how to identify influential nodes in complex networks by using relative entropy, which combines the advantages of existing centrality measures. Peng et al. (Peng, Yang, Cao, Yu, & Xie, 2017) characterized the features of mobile social networks and presented an evaluation model to quantify influence by entropy. Rashevsky is the first researcher who measured the structural complexity of the graphs by using the graph entropy (Rashevsky, 1955). He analyzed the complexity of organic molecules within the structure where the nodes are physically indistinguishable from the atoms and the edges which consists of chemical bonds. The definition is based on the fact that the graph consisting of n nodes are divided into k equivalent classes according to the rank of the nodes. Trucco used this definition in (Ernesto, 1956) and the role of the group of a graph in determining sets of the equivalent points were shown. The two nodes are assumed to be equivalent if they have the same orbit and can keep the association of the graph when they change their place. Mowshowitz showed the properties of some graph operations using the same measures (Mowshowitz, 1968). Namely, he interpreted the entropy as the graph's structural information. Entropy can be used to analyze the structural complexity of the social network systems. Also, it can be used to obtain information in the social networks, in order to understand the importance of the nodes, to look at the quantities of relations between nodes and to identify communities. Improvements can be made with entropy when optimization is needed in the calculations (Fei & Deng, 2017; Nie, Guo, Zhao, & Lu, 2016; Peng et al., 2017; Zhang, Li, & Deng, 2018).

Let $p = (p_1, p_2, \dots, p_n)$ be the probabilities of the nodes in each case for a random event with n states. The entropy of node v is calculated as $E(v)$,

$$E(v) = -\sum_{i=1}^{n} p_i log p_i,\ 0 \leq p_i \leq 1 \tag{1}$$

It was aimed to conduct social network analysis using landmass, zone, language, religion data which is thought to be suitable for finding similarity among 30 properties belonging to 193 countries data set. When we convert this data into a graph, we have a network with 10133 undirected edges.

The dataset which is used in this application was taken from a study done in 1986. Today some of these countries have been divided into different countries or renamed. However, in this study analyzes were made on existing data without considering these changes. It was thought that those selected factors would be more effective in defining country relations. Their similarities were calculated. These parameters can be reduced and multiplied to look at results from different perspectives. The coefficient of the parameters can be increased to rise the effect of the parameters.

- **Religion:** It can be seen that religions even sects are very influential on the formation of country associations.
- **Language:** Language cohesion is a kinship mark and one of the most influential factors in the relationship between countries.
- **Landmass and Zone:** Countries that are neighbors and in the same geographical location have a forced togetherness. Because changes in the country will affect the neighbors more than other.

In this study entropy calculations were performed using the similarity values of the edges. Similarity can be measured by many different methods (Linyuan & Zhou, 2011; Tuğal, Kaya, & Tuncer, 2013). The similarity ratios between the nodes can be measured by using common properties of nodes. The number of similar features of the two nodes were calculated. If there are more common features between two nodes, the similarity of these nodes is greater than others. Similarity ratio was obtained by dividing sum of common similar properties number between two nodes to the total defined parameter coefficient used for similarity. More weight coefficient can be given to some of the parameters. If it is thought that there are some more influential parameters for the relationship, we may increase the effect of parameter by increasing the weight.

The similarity calculations were performed by using Equation (2) or Equation (3). The ω denotes the parameter coefficient defined for each parameter. The $s(i,j)$ is the similarity value of nodes for a parameter. If a parameter value is similar for i node and j node, its value is 1. The $s(i,j)$ values for each parameter are multiplied by established parameter coefficients and summed then it was divided by the weight coefficient sums to see the effect of all parameters.

$$S_{ij} = \frac{(\omega_1 s(i,j) + \omega_2 s(i,j) + \cdots + \omega_n s(i,j))}{\omega_1 + \omega_2 + \cdots + \omega_n} \tag{2}$$

$$S_{ij} = \frac{\omega_1 s(i,j)_{(\boldsymbol{relig.})} + \omega_2 s(i,j)_{(\boldsymbol{lang})} + \omega_3 s(i,j)_{(\boldsymbol{land})} + \omega_4 s(i,j)_{(\boldsymbol{zone})}}{\omega_1 + \omega_2 + \omega_3 + \omega_4} \tag{3}$$

In weighted networks, s_{ij} , represents the similarity value of the edge between nodes i and j and probability vector is p_{ij}. The $s_{ij} > 0$ and $\Gamma(i)$ is the neighbors of node i,

$$p_{ij} = \frac{s_{ij}}{\sum_{j\in\Gamma(i)} s_{ij}} \tag{4}$$

$$E(v_i) = -\sum_{j=1}^{n} p_{ij} log p_{ij} \tag{5}$$

$$E(v_i) = \sum_{j=1}^{d_i} |(-p_{ij})^{\alpha} log p_{ij}| , \quad 0<\alpha \tag{6}$$

After each node (i'th node) entropy is calculated by using Equation (5) then the total entropy of the graph is calculated. The base of the logarithm is 2. Here, we used the Karcı entropy (Karcı, 2018; İ. Tuğal & Karcı, 2019) in Equation (6) that using α exponent.

4. Application with Sample Data

The general factors that connect humans are kinship, neighborhood, common past, common political ideas. The same situation affects the relations of the countries. The countries cooperate with each other. They come together and take collective decisions on many subjects. The European Union, The Arabic Association, NATO and many other similar organizations continue to operate today. There are many factors that form this kind of unities. It can be easily seen that common geography, religion and language are influential in the formation of these organizations (İhsan Tuğal & Karcı, 2016). From this, it was thought that social network analysis methods can be used to reach influential countries and countries' unity map.

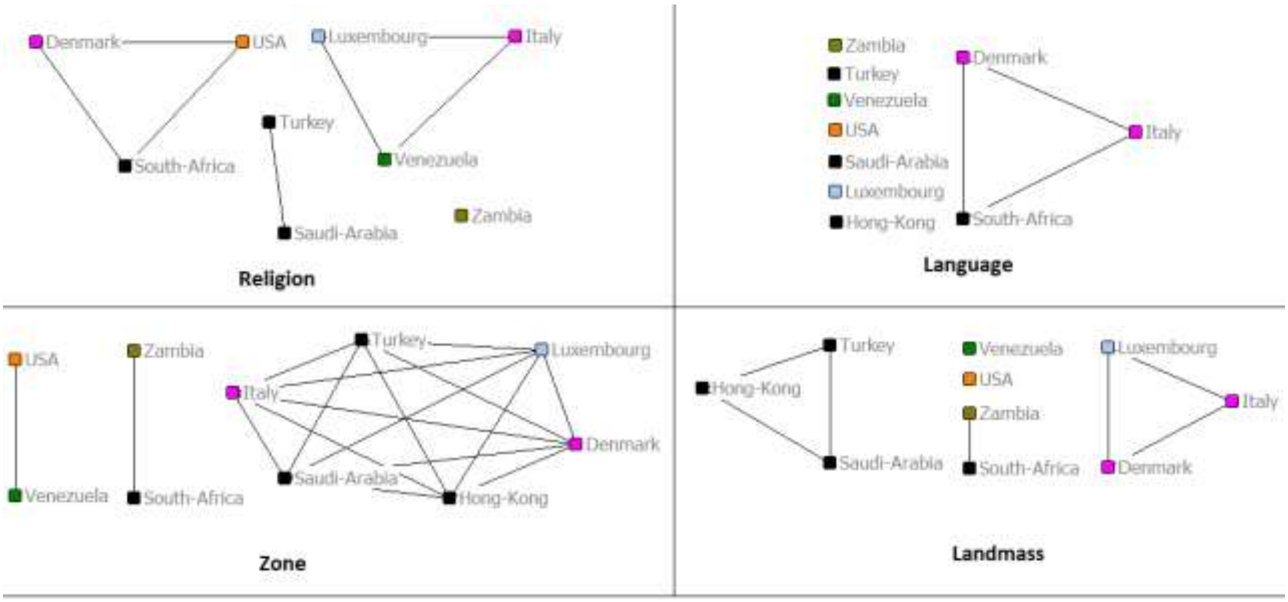


Figure 2. The link structure of the 10 countries given in Table 1.

This example illustrates how the proposed algorithm works. The data of 10 countries was selected.

Table 1. 10 country data from Flags dataset

Name	Religion	Language	Zone	Landmass
Zambia	Ethnic	Others	SE	Africa
Turkey	Muslim	Japanese/Turkish/ Finnish/Magyar	NE	Asia
Venezuela	Catholic	Spanish	NW	S. America
USA	Other Christian	English	NW	N. America
Saudi-Arabia	Muslim	Arabic	NE	Asia
South-Africa	Other Christian	Other Indo-European	SE	Africa
Luxembourg	Catholic	German	NE	Europe
Italy	Catholic	Other Indo-European	NE	Europe
Hong-Kong	Buddhist	Chinese	NE	Asia
Denmark	Other Christian	Other Indo-European	NE	Europe

Table 2. Similarity count matrix of 10 countries

	Zambia	Turkey	Venezuela	USA	Saudi-Arabia	South-Africa	Luxembourg	Italy	Hong-Kong	Denmark
Zambia	4	0	0	0	0	2	0	0	0	0
Turkey	0	4	0	0	3	0	1	1	2	1
Venezuela	0	0	4	1	0	0	1	1	0	0
USA	0	0	1	4	0	1	0	0	0	1
Saudi-Arabia	0	3	0	0	4	0	1	1	2	1
South-Africa	2	0	0	1	0	4	0	1	0	2
Luxembourg	0	1	1	0	1	0	4	3	1	2
Italy	0	1	1	0	1	1	3	4	1	3
Hong-Kong	0	2	0	0	2	0	1	1	4	1
Denmark	0	1	0	1	1	2	2	3	1	4

Landmass, zone, language, religion information of the mentioned countries was taken as shown in Table 1. A graph was created by using this information. The similarities and entropy values of the countries were computed. The most effective countries and clusters were identified. Referring to Figure 2, for example, between Saudi Arabia and Turkey formed edges for religion, zone and landmass. The number of similarities is 3. According to the parameters, the similarity counts of the countries are shown in Table 2.

If desired, the effect of some parameters to the network can be increased. For example, when the effect of religion is doubled, the matrix of the similarity counts becomes as in Table 3.

Table 3. Similarity count matrix (The effect of religion is 2)

	Zambia	**Turkey**	**Venezuela**	**USA**	**Saudi-Arabia**	**South-Africa**	**Luxembourg**	**Italy**	**Hong-Kong**	**Denmark**
Zambia	5	0	0	0	0	2	0	0	0	0
Turkey	0	5	0	0	4	0	1	1	2	1
Venezuela	0	0	5	1	0	0	2	2	0	0
USA	0	0	1	5	0	2	0	0	0	2
Saudi-Arabia	0	4	0	0	5	0	1	1	2	1
South-Africa	2	0	0	2	0	5	0	1	0	3
Luxembourg	0	1	2	0	1	0	5	4	1	2
Italy	0	1	2	0	1	1	4	5	1	3
Hong-Kong	0	2	0	0	2	0	1	1	5	1
Denmark	0	1	0	2	1	3	2	3	1	5

If the similarity values of the nodes are calculated according to the Equation (3), the similarity matrix of the nodes becomes as in Table 4.

The network of the countries in Table 4 is illustrated in Figure 3. The most central nodes are Denmark and Italy. South-Africa is the only country that connects Zambia to the network

Table 4. 10 country similarity matrix

	Zambia	**Turkey**	**Venezuela**	**USA**	**Saudi-Arabia**	**South-Africa**	**Luxembourg**	**Italy**	**Hong-Kong**	**Denmark**
Zambia	1	0	0	0	0	0,4	0	0	0	0
Turkey	0	1	0	0	0,8	0	0,2	0,2	0,4	0,2
Venezuela	0	0	1	0,2	0	0	0,4	0,4	0	0
USA	0	0	0,2	1	0	0,4	0	0	0	0,4
Saudi-Arabia	0	0,8	0	0	1	0	0,2	0,2	0,4	0,2
South-Africa	0,4	0	0	0,4	0	1	0	0,2	0	0,6
Luxembourg	0	0,2	0,4	0	0,2	0	1	0,8	0,2	0,4
Italy	0	0,2	0,4	0	0,2	0,2	0,8	1	0,2	0,6
Hong-Kong	0	0,4	0	0	0,4	0	0,2	0,2	1	0,2
Denmark	0	0,2	0	0,4	0,2	0,6	0,4	0,6	0,2	1

In the example, each node removed from graph one by one and the total entropy value calculated again with similarity values as described in proposed algorithm.

Algorithm

1- For each parameter

Identify n nodes

Create nxn adjacency matrix from the $nx3$ edge list of the association according to the parameter

If the parametric effect of the network is to be increased, multiply the matrix by the desired coefficient

2- Compute the similarity value between nodes

for $i = 1{:}n$

for $j = 1{:}n$

Similarity (i, j) = *(compute the similarity value of nodes i and j according to Equation (2))*

end

3- for $k = 1{:}n$

remove node k from network

4- Compute Entropy

for $i = 1{:}n - 1$

for $j = 1{:}n - 1$

Entropy (i, j)= *(compute entropy according to Equation (5) and Equation (6) by using Equation (4))*

5- Calculate total entropy

6- Go to step 3 if $k \neq n$

7- Rank obtained entropy values from small to high

8- Identify the most influential nodes.

The obtained total entropy values are shown in Table 5. The total entropy value decreases when the most effective node in the network was removed from the network. According to the results obtained, the most effective nodes of the network are Italy and Denmark. The most ineffective node is Zambia. It can be said that Turkey and Saudi Arabia have similar values and are in the same community.

Table 5. The entropy values of network after removing one of the countries

Country Number	**Country Name**	**Total Entropy**
1	Zambia	20,1276630381946
2	Turkey	17,9202873088361
3	Venezuela	18,7597842091975
4	USA	18,6850106294451
5	Saudi-Arabia	17,9202873088361
6	South-Africa	17,4817975856653
7	Luxembourg	17,2241493360873
8	Italy	16,7609053413772
9	Hong-Kong	17,9000108029849
10	Denmark	16,5305934323261

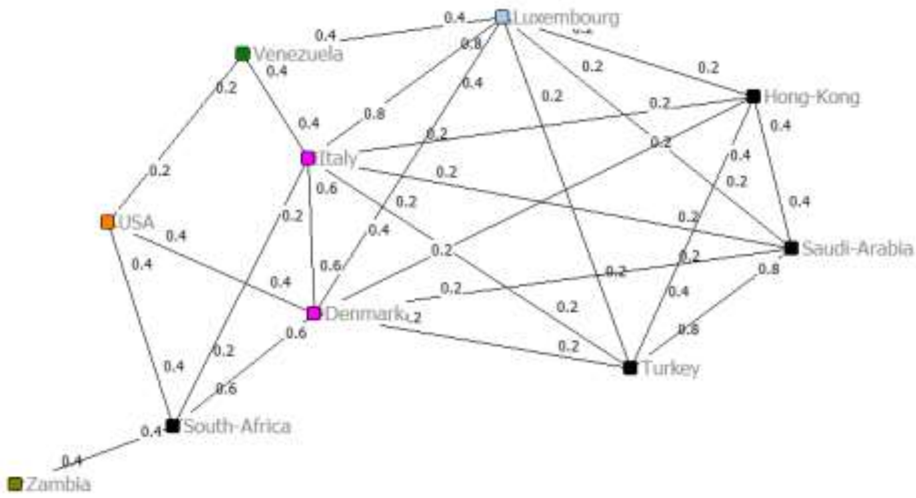


Figure 3. Weighted countries graph

In Table 6, the results obtained with Shannon entropy were compared with the results which are obtained from Karcı entropy. The α coefficient selected between 0.1 and 1.2 makes the difference between the entropy values of the nodes even more apparent. When the coefficient values are changed, the importance of some countries increase and some of them decrease. The cause of this change is reducing the effect of edge's weight and increasing the effect of node's degree. When the coefficient α is equal to 0.1, the effect of the edge weights is very weak. If α value is chosen bigger than 1.3, the inverse calculation starts to take place. Overhead nodes in the rankings begin to come down. As the α coefficient decreases, the situation reverses. According to the link density or flow of the graph, this threshold value changes. In the Flags graph, this threshold is 1.3. Karcı entropy can be used to follow the effect of node degrees and edge weights to the network on the weighted networks. With the change of α, it was seen that there was a change in the order of the effective countries. In Shannon entropy, South-Africa is more effective than Hong-Kong, Turkey and Saudi Arabia because it has more weighted edges with effective nodes. It's the only node that connects Zambia to the network. But when the α coefficient used by the Karcı entropy decrease, the rank of importance of South-Africa drops. Because in the Karcı entropy the degree of the nodes was taken into consideration more than. Because of the similar degree and linked nodes, the entropy values of Turkey and Saudi Arabia are equal. When compared with Hong Kong, their neighbors and their edge weights with these neighbors are similar. But in Karcı entropy, they look more influential than Hong Kong because the weight of link to each other is bigger.

Table 6. Comparison of Shannon and Karcı entropy

Karcı	**α=1,2**	**Karcı**	**α=1,1**	**Shannon**	**α=1**	**Karcı**	**α=0,9**	**Karcı**	**α=0,8**	**Karcı**	**α=0,7**
Denmark	12,06	Denmark	14,09	Denmark	16,53	Denmark	19,48	Denmark	23,07	Denmark	27,45
Italy	12,22	Italy	14,28	Italy	16,76	Italy	19,75	Italy	23,36	Italy	27,75
South-Africa	12,28	Luxembourg	14,62	Luxembourg	17,22	Luxembourg	20,38	Luxembourg	24,23	Luxembourg	28,92
Luxembourg	12,46	South-Africa	14,62	South-Africa	17,48	South-Africa	21	Hong-Kong	25,32	Hong-Kong	30,31
Turkey	12,85	Turkey	15,15	Hong-Kong	17,9	Hong-Kong	21,24	South-Africa	25,33	Turkey	30,4
Saudi-Arabia	12,85	Saudi-Arabia	15,15	Turkey	17,92	Turkey	21,29	Turkey	25,39	Saudi-Arabia	30,4
Hong-Kong	12,87	Hong-Kong	15,15	Saudi-Arabia	17,92	Saudi-Arabia	21,29	Saudi-Arabia	25,39	South-Africa	30,68
USA	13,23	USA	15,69	USA	18,69	USA	22,36	USA	26,88	USA	32,47
Venezuela	13,29	Venezuela	15,75	Venezuela	18,76	Venezuela	22,44	Venezuela	26,97	Venezuela	32,57
Zambia	14,11	Zambia	16,82	Zambia	20,13	Zambia	24,19	Zambia	29,21	Zambia	35,41

Karcı	**α=0,6**	**Karcı**	**α=0,5**	**Karcı**	**α=0,4**	**Karcı**	**α=0,3**	**Karcı**	**α=0,2**	**Karcı**	**α=0,1**
Denmark	32,8	Denmark	39,38	Denmark	47,48	Italy	57,51	Italy	69,66	Italy	84,69
Italy	33,1	Italy	39,64	Italy	47,65	Denmark	57,51	Denmark	69,96	Denmark	85,45
Luxembourg	34,66	Luxembourg	41,73	Luxembourg	50,45	Luxembourg	61,24	Luxembourg	74,63	Luxembourg	91,3
Hong-Kong	36,43	Hong-Kong	43,99	Hong-Kong	53,33	Hong-Kong	64,92	Turkey	79,29	Turkey	97,11
Turkey	36,55	Turkey	44,11	Turkey	53,44	Turkey	64,97	Saudi-Arabia	79,29	Saudi-Arabia	97,11
Saudi-Arabia	36,55	Saudi-Arabia	44,11	Saudi-Arabia	53,44	Saudi-Arabia	64,97	Hong-Kong	79,36	Hong-Kong	97,39
South-Africa	37,33	South-Africa	45,61	South-Africa	55,96	South-Africa	68,92	South-Africa	85,21	South-Africa	105,7
USA	39,41	USA	48,05	USA	58,85	USA	72,4	Venezuela	89,37	Venezuela	110,8
Venezuela	39,5	Venezuela	48,13	Venezuela	58,91	Venezuela	72,4	USA	89,46	USA	111
Zambia	43,11	Zambia	52,72	Zambia	64,73	Zambia	79,81	Zambia	98,78	Zambia	122,7

The results in Table 6 show that more accurate measurements can be made by Karcı entropy. And this entropy can be used to adjust the importance of node degree for analysis. In this way, the graph can be analyzed with a different perspective.

5. Experimental Results

Using the above-mentioned method, the application was performed using 193 countries in the Flags data set. The statistical information of the parameters in the Flags data set is given in Table 7. As seen, the most common religions in these countries are Other Christian and Catholic. The most talked languages are Others and English. The African landmass has more countries. As the zone where the countries are most dense is NE.

The similarity between two countries was measured according to common parameters and parameter coefficients. Religion's parameter coefficient was given as 2. The number of similar parameters of the two countries were calculated. Similarity ratio was obtained by dividing sum of common similar parameter number between two countries to the total defined parameter coefficient used for similarity. The s_{ij} similarity ratio between two countries was found.

Table 7. The statistical information of the Flags dataset

RELIGION			LANGUAGE		
Values as number	Religion Name	Number of Countries	Values as number	Language Name	Number of Countries
0	Catholic	39	**1**	English	43
1	Other Christian	60	**2**	Spanish	20
			3	French	17
2	Muslim	36	**4**	German	6
3	Buddhist	8	**5**	Slavic	4
4	Hindu	4	**6**	Other Indo-European	30
5	Ethnic	27			
6	Marxist	15	**7**	Chinese	4
7	Others	4	**8**	Arabic	19
	Total	193	**9**	Japanese/ Turkish/ Finnish/ Magyar	4
			10	Others	46
				Total	193

LANDMASS			ZONE		
Values as number	Landmass Name	Number of Countries	Values as number	Zone Name	Number of Countries
1	N. America	31	**1**	NE	91
2	S. America	16	**2**	SE	29
3	Europe	35	**3**	SW	15
4	Africa	52	**4**	NW	58
5	Asia	39		Total	193
6	Oceania	20			
	Total	193			

Based on these similarity values, the entropy of each country was calculated. The sum of the obtained values gives the total entropy. For learning the influence of the country, total entropy was calculated again after removing the country from the graph. In this way the change of the graph was measured. Decrease in the value of the total entropy is actually a measure of the effectiveness of the node. After the total entropy values obtained by removing the countries one by one were ranked from small to big. It can be said that the upper countries are the most effective nodes of the graph.

With Shannon entropy and Karcı entropy, the most influential and less influential countries were found according to religion, language, landmass and zone parameters. The results are given in Table 8. If the entropy value decrease too much when the node is removed from network, it says that the node is important for the system. The rate of change of the total entropy value actually indicates the influence of the node on the system.

Table 8. The most influential and less influential countries

Sequence No	**Most Effective Countries**	**α=1**	**Most Effective Countries**	**α=0,5**	**Most Effective Countries**	**α=0,1**
1	Ethiopia	1236,8	Cameroon	12476,71	Cameroon	84146,3
2	Cameroon	1236,85	Ethiopia	12477,47	Ethiopia	84167,47
3	Sao-Tome	1237,06	Sao-Tome	12486,58	Sao-Tome	84261,97
4	Philippines	1237,18	Marianas	12488,75	Marianas	84280,1
5	Marianas	1237,18	Micronesia	12488,75	Micronesia	84280,1
6	Micronesia	1237,18	Philippines	12492,61	Philippines	84340,21
7	Malta	1237,39	Guam	12497,47	Guam	84352,72
8	Guam	1237,5	Kiribati	12497,47	Kiribati	84352,72
9	Kiribati	1237,5	Malta	12499,36	Cyprus	84382,57
10	Cyprus	1237,54	Cyprus	12499,48	Denmark	84382,57
11	Denmark	1237,54	Denmark	12499,48	Greece	84382,57
12	Greece	1237,54	Greece	12499,48	Netherlands	84382,57
13	Netherlands	1237,54	Netherlands	12499,48	Norway	84382,57
14	Norway	1237,54	Norway	12499,48	Sweden	84382,57
15	Sweden	1237,54	Sweden	12499,48	Malta	84405,61
…………………...						
179	Malawi	1239,32	Rwanda	12555,36	Cuba	84986,44
180	Mozambique	1239,32	Tanzania	12555,36	Papua-New-Guinea	84992,16
181	Rwanda	1239,32	Zaire	12555,36	American-Samoa	84993,66
182	Tanzania	1239,32	Zambia	12555,36	Cook-Islands	84993,66
183	Zaire	1239,32	Zimbabwe	12555,36	Niue	84993,66
184	Zambia	1239,32	Papua-New-Guinea	12555,51	Western-Samoa	84993,66
185	Zimbabwe	1239,32	French-Polynesia	12561,27	French-Polynesia	85059,62
186	Brazil	1239,43	Brazil	12565,15	Brazil	85085,53
187	Argentina	1240,13	Argentina	12583,87	Argentina	85262,04
188	Bolivia	1240,13	Bolivia	12583,87	Bolivia	85262,04
189	Chile	1240,13	Chile	12583,87	Chile	85262,04
190	Ecuador	1240,13	Ecuador	12583,87	Ecuador	85262,04
191	Parguay	1240,13	Parguay	12583,87	Parguay	85262,04
192	Peru	1240,13	Peru	12583,87	Peru	85262,04
193	Uruguay	1240,13	Uruguay	12583,87	Uruguay	85262,04

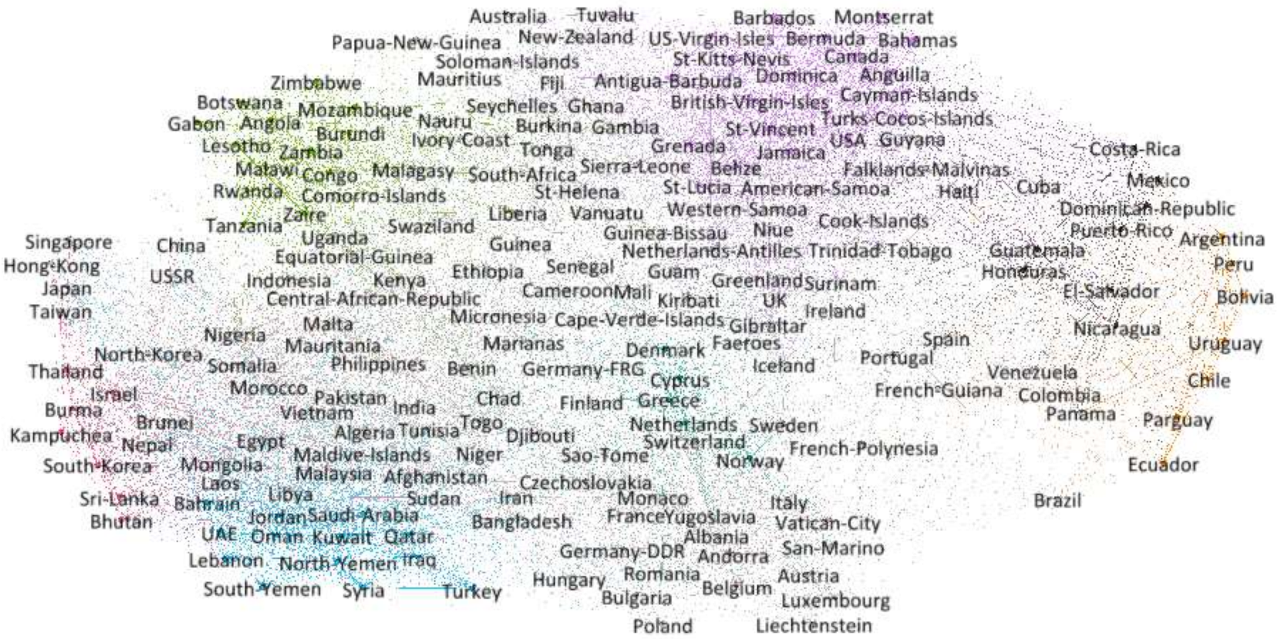


Figure 4. Unity and clusters of 193 countries

The clusters of the countries are shown in Figure 4. The influence of geographical factors can be seen in the formation of clusters. Moreover, the use of two geographical parameters in graph was made this situation more evident. Geographically close countries influence each other on language and religion. This fact was shown again.

It can be seen the correctness of our entropy-based method from Figure 5 that show the links of the most influential countries and Figure 6 that show the links of the less influential countries (M. Bastian, S. Heymann, 2009).

Figure 5. The links of the most influential countries

In Karcı entropy, when the small value was given to α, Malta and Philippines begin to descend down to the order of ranking as seen in Table 8. The order of Sao Tome has not changed, but it showed a tendency to fall into lower. It can be seen that they have links with the most ineffective countries from Figure 6. The effect of being associated with the ineffective nodes was measured by the Karcı entropy.

According to Shannon entropy, Ethiopia is the most influential country. Ethiopia has links to almost all the countries of the world except Latin countries. Ethiopia has features that have the most common values in most of the world as shown in Table 7. It’s features are Africa (landmass), NE (zone), Others (language) and Other Christian (religion).

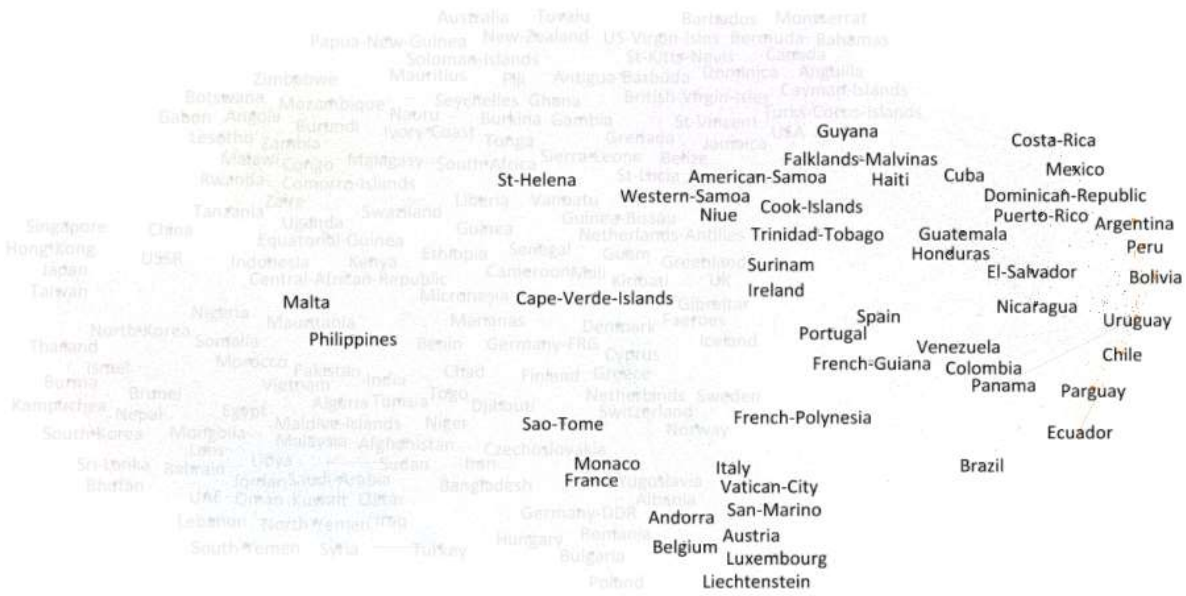


Figure 6. The links of the less influential countries

The less influential countries are Latin countries. Uruguay was the most ineffective country with "S. America", "SW”, “Spanish" and "Catholic" properties. Because there are small number of countries with geographical properties like Uruguay. The reason for the emergence of Latin countries as less influential countries is the zone and language parameters.

6. Conclusion

In this study, it has been shown that the relation map of countries can be extracted with social network analysis methods by choosing appropriate parameters. Useful information is obtained through the links and common parameters. The similarity index was calculated by religion, language, landmass, and zone information of countries. The most effective nodes of a system can be found by looking similarities and changes in the system. One of the methods for measuring change and uncertainty is entropy. From the Flags dataset, the most important countries were found by using graph entropy. The world's most influential countries and country clusters were determined by using their common geographical and cultural properties.

Shannon entropy is generally used in graph applications. In addition, we implemented the Karcı entropy. The results show that Karcı entropy can be used to adjust effect of the node degree and edge weight on the weighted networks.

Relations between countries were affected by the specified parameters. But it was shown that these are not the greatest parameters that guide politics and economics around the world. Because the influential countries which are leading the world today did not come out in the first order when we used these cultural and geographical parameters.

The world is in constant change, transformation and remodeling. With the right parameters, this change and transformation can be measured by social network analysis. Thus, the economic, social, and military relations to be made with countries can be reconstructed. Developing meaningful policies and sustaining cultural collaborations are up to the right decisions. The correctness of these decisions must also be based on certain calculations. Methods that based on the certain calculations may contribute to making the right decisions for the future. The results obtained from this study, which was focused on methodology more, can be interpreted in terms of social sciences in another study.

References

Borgatti, S P, Everett, M. G., & Freeman, L. C. (2002). UCINET 6 For Windows: Software for Social Network Analysis, Analytic Technologies, Harvard, MA. *Analytic Technologies*.

Borgatti, Stephen P. (2005). Centrality and network flow. *Social Networks*, *27*(1), 55–71. https://doi.org/10.1016/j.socnet.2004.11.008

Chen, D., Lü, L., Shang, M. S., Zhang, Y. C., & Zhou, T. (2012). Identifying influential nodes in complex networks. *Physica A: Statistical Mechanics and Its Applications*, *391*(4), 1777–1787. https://doi.org/10.1016/j.physa.2011.09.017

Durland, M. M., & Fredericks, K. A. (2005). An introduction to social network analysis. *New Directions for Evaluation*, (107), 5–13. https://doi.org/10.1002/ev.157

Ernesto, T. (1956). A note on the information content of graphs. *The Bulletin of Mathematical Biophysics*, *18*(2), 129–135.

Fei, L., & Deng, Y. (2017). A new method to identify influential nodes based on relative entropy. *Chaos, Solitons and Fractals*, *104*, 257–267. https://doi.org/10.1016/j.chaos.2017.08.010

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215–239. https://doi.org/10.1016/0378-8733(78)90021-7

Gürsakal, N. (2009). *Sosyal Ağ Analizi*. Dora Yayınları.

Huntington, S. P. (2001). *Medeniyetler Çatışması*. Vadi Yayınları.

Karagöz, D., & R, Y. H. (2013). Examining research topics of doctoral theses written in the field of tourism by social network analysis. *Adıyaman University Journal of Social Sciences*, *15*.

Karcı, A. (2018). Notes on the published article "Fractional order entropy: New perspectives." *Optik - International Journal for Light and Electron Optics*, *171*, 107–108.

Lichman, M. (n.d.). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Retrieved January 4, 2018, from http://archive.ics.uci.edu/ml

Linyuan, L. L., & Zhou, T. (2011). Link prediction in complex networks: A survey. *Physica A: Statistical Mechanics and Its Applications*. https://doi.org/10.1016/j.physa.2010.11.027

M. Bastian, S. Heymann, M. J. et al. (2009). Gephi: an open source software for exploring and manipulating networks. In *Proceedings of International AAAI Conference on Web and Social Media*. https://doi.org/10.1136/qshc.2004.010033

Mowshowitz, A. (1968). Entropy and the complexity of graphs: I. An index of the relative complexity of a graph. *The Bulletin of Mathematical Biophysics*, *30*(1), 175–204. https://doi.org/10.1007/BF02476948

Nie, T., Guo, Z., Zhao, K., & Lu, Z. M. (2016). Using mapping entropy to identify node centrality in complex networks. *Physica A: Statistical Mechanics and Its Applications*, *453*, 290–297. https://doi.org/10.1016/j.physa.2016.02.009

Padgett, J. F. (2010). Open Elite? Social Mobility, Marriage, and Family in Florence, 1282–1494*. *Renaissance Quarterly*, *63*(02), 357–411. https://doi.org/10.1086/655230

Peng, S., Yang, A., Cao, L., Yu, S., & Xie, D. (2017). Social influence modeling using information theory in mobile social networks. *Information Sciences*, *379*, 146–159. https://doi.org/10.1016/j.ins.2016.08.023

Rashevsky, N. (1955). Life, information theory, and topology. *The Bulletin of Mathematical Biophysics*, *17*(3), 229–235. https://doi.org/10.1007/BF02477860

Raymond, C. (2005). *Physical Chemistry for the Biosciences*.

Scott, J. P. (2000). *Social Network Analysis: A Handbook. SAGE Publications*.

Shannon, C. E. (1951). Prediction and Entropy of Printed English. *Bell System Technical Journal*, *30*(1), 50–64. https://doi.org/10.1002/j.1538-7305.1951.tb01366.x

Shetty, J., & Adibi, J. (2006). Discovering important nodes through graph entropy the case of Enron email database (pp. 74–81). https://doi.org/10.1145/1134271.1134282

Tuğal, İ., & Karcı, A. (2019). Comparisons of Karcı and Shannon entropies and their effects on centrality of social networks. *Physica A: Statistical Mechanics and Its Applications*,*523*. https://doi.org/10.1016/j.physa.2019.02.026

Tuğal, İ., & Karcı, A. (2016). Computing association between countries. In *International Artificial Intelligence and Data Processing Symposium (IDAP 2016)* (pp. 87–92). Retrieved from idapold.inonu.edu.tr

Tuğal, İ., Kaya, M., & Tuncer, T. (2013). Link prediction in disease and drug networks. In *6-th International Conference of Advanced Computer Systems and Networks: Design and Application (ACSN 2013)* (pp. 46–50).

Wasserman, S., & Galaskiewicz, J. (2012). *Advances in Social Network Analysis: Research in the Social and Behavioral Sciences*. *Advances in Social Network Analysis: Research in the Social and Behavioral Sciences*. https://doi.org/10.4135/9781452243528

Wolfe, A. W. (1997). Social Network Analysis: Methods and Applications. *American Ethnologist*, *24*(1), 219–220. https://doi.org/10.1525/ae.1997.24.1.219

Zhang, Q., Li, M., & Deng, Y. (2018). Measure the structure similarity of nodes in complex networks based on relative entropy. *Physica A: Statistical Mechanics and Its Applications*. https://doi.org/10.1016/j.physa.2017.09.042