2018 Turkish Parliamentary Election Results' Relationship with Interconnectedness of Deputy Characteristics and City Demographics

Omay Ece Ayyıldız, Ali Güzel, Berkant Kartop, Nazlıcan Turan, Hasan Kadir Uçmazoğlu, Elif Yıldırır

Abstract

Elections have always been a significant research area since forecasting is important for both public and political spheres. In this project, while analyzing the 2018 deputy election results, it is intended to understand effects of the features and underlying patterns in an election which are actually significant but unspoken-of. Main aim is to prove that the education and age status of a city is interconnected with elected deputies' profession, age, place of birth and lastly the political party the deputy represents. Such features, the dataset for the project, has been gathered from three different sources; TBMM database for deputy data, YSK and Biruni University's databases for city demographics. Analyzing such datasets, it is being questioned whether the well-educated cities elect deputies that have high profile professions, are they open for younger generations to have a role in assembly and do they value for their deputies to be fellow townsmen; and whether such answers are opposite for less educated cities. For such analysis, the network of the deputies is going to have an important role, representation of the demographically similar cities' deputies being similar by weighting edges with cosine similarity which shows the hidden relations and similarities within nodes, would support the hypothesis. In addition to that, to analyze the elected parties' target cities, the network of the cities based on their demographics is expected to show the cluster of cities that support the similar profile of deputy; in such part, it is also expected for cities of the same states or the cities that are near the sea to have similar centralities.
***Keywords**: TBMM, deputy, elections, network, demographics.*

1. Introduction

Since the acceptance of the multi-party system in 1946, The Republic of Turkey adopted different electoral systems. Currently the parliamentary elections of the country are held every five years alongside with the presidential elections on the same day and a system of proportional representation in the parliament of political parties with a national threshold of 10% is valid which still remains as a controversial topic in Turkish politics. Each 81 city has a deputy quota that varies according to the number of habitants of that particular city. For

instance, Istanbul with the population of nearly 16 million has 98 seats in the parliament while Tunceli, Turkish city with the least number of habitants with 83.645 people, has a quota of 2 deputies. It is known that the main race in the last 2018 elections was between The People's Alliance which secured 53.7 percent of the votes and the Nation Alliance which received 33.9 percent. *(The Grand National Assembly of Turkey).* Because many social scientists argue that Turkey is ruled by an authoritarian democracy, both the results of parliamentary and presidential elections are significantly affecting the country's future policies and since the next elections are going to be held 18 June 2023 which is getting closer and becoming a trend topic countrywise, revealing the underlying factors in the last elections was the main motivation of this project. It is known among the population of the country that voters' educational status and age distribution highly affects the party that they vote for as well as the city's geographical conditions such as closeness to the sea or geographical region significantly affects the parliamentary elections according to pretty common rumors in the country. Our project focuses on these indicated relations as well as some other possible less commonly talked/untaught relations such as the deputy's education, professions and fellow townsmen status. The results of the last parliamentary elections of Turkey which were held on 24 June 2018 as a part of the country's general elections with 59,354,840 voters together with the demographics datasets of its 81 cities are formed out of previously ready datasets in order to dig down deep into the factors affecting the outcome with network science and data science methods in this project. (The Grand National Assembly of Turkey). Two main networks, one with 81 cities demographics such as the gender, age and education distribution of the voters and other with the deputy's previously indicated characteristics are formed and analyzed.

2. Related Work

The first paper from the literature review is *"Tanzimat'tan Günümüze Türkiye'de Parlamentonun Sosyal Profili"* written by Özlem Ertem in 2010 for her sociology masters thesis in Istanbul University. Even though this paper deals with a large time scale starting from 1877, the section that we are most interested in is the section 3.6 (*"1980'den Günümüze TBMM'de Görev Yapan Milletvekillerinin Sosyal Profili")*, which goes over the social profiles of the senators starting from 1980. Subsections of this section include profession, age, educational status and sex distribution of the senators (Ertem, 2010).

In the first part of this section, in parallel to our own observations, Ertem (2010) states that the majority of the senators are either doctors, lawyers or engineers. Right after these two professions, dentists and pharmacists occupy the majority of the profession distribution, which makes the health sector the biggest portion of the sector based distribution (Ertem, 2010). In addition to these, Ertem (2010) states that the community leaders were a rising profession group at that time, which is inline with our findings for recent election results.

When it comes to the age distribution of the senators; Ertem (2010) states that, even though there was an increase in amount of relatively young senators due to the influence of 1980 coup, after the 1991 elections and the ones following that the age groups 41 - 50 and 51 -60 started regaining their positions in terms of being the majority. This gradual increase in the portion of the relatively older groups led them to eventually being the majority in the recent elections, which is again inline with the observations done by us.

According to Ertem's (2010) findings; educational status of the senators only went up, taking the 85% bachelor's degree rate of 1983 to 94% in 2007, which is almost equal to our findings. Additionally; Ertem (2010) states that the majority of the elementary school graduates are from central right wing parties, and she points out that the majority of these senators are from the East and South-East Anatolia. She points out that the elementary school graduate senator portion coming from these regions never fell below the 40% rate. Just like our observations, Ertem (2010) mentions that the majority of the senators who have a graduate/masters degree are coming from the social sciences department of their university. Ertem (2010) also states that the majority of the senators who went through a technical education (engineering, architecture, technicians..) consists of engineers.

Another paper we focused on is *"Türkiye Seçmen Araştırmaları 20' Seçmen Grupları ve Eğilimleri"* which is a research about how the different elements such as sex, age and education of the voter effects their voting prefrences in the 2020 and 2018 elections.

In the research they analyzed the education levels of the electors and found that the difference in education level of the electors created a difference in their voting preference. According to the results, voters with lower education levels were more likely to vote for deputies from AKP and the voters with higher education levels were more biased to vote for deputies from CHP. (TEAMA, 2020).

The research has also stated that different age groups had different biases when it comes to the deputies they have chosen. The citizens in the age group 18-29 preferred the deputies from CHP and the age citizens in the age group 30-50 preferred the deputies from AKP. (TEAMA, 2020). At the last part of the profile analyzing section the researchers also stated that the citizens with a lower income were more likely to prefer a different deputy than the citizens with higher income. (TEAMA, 2020). This implies that the difference in the profession of citizens affects their voting preferences. These results state that the demography of the cities is one of the main factors for deputy election which is inline with our observations.

In order to make a literature review for justifying our network model; we will look into the *"A complex network approach to political analysis: Application to the Brazilian Chamber of Deputies"* paper, written by the authors Ana Caroline Medeiros Brito, Filipi Nascimento Silva, and Diego Raphael Amancio published in March 19, 2020. This paper introduces a "network-based methodology to study how political entities evolve over time". Similar to our research, networks on this paper are constructed from the voting data, which is the Brazilian Chamber of Deputies Voting data in their case. Nodes and edges of this network are represented by deputies and voting similarity among deputies, respectively. Authors of the paper believe that "the proposed framework could be used to complement the study of political dynamics and even applied in similar social networks where individuals are organized in a complex manner".

Methodology recruited by the authors is summarized in three steps. In the first step, which is network construction, authors state that they created "a network from voting patterns in a given period", with the condition of linking the deputies if they "vote in a similar fashion in several propositions". Right after the network construction step, they have proceeded with the network backbone extraction step, which involves the "removal of the weakest edges". By doing this, authors state that they can "create a network that can be treated by most of the traditional community detection methods". At the last part, by employing community detection methods, they identified the "groups of deputies displaying similar voting patterns".

There are several critical measurements to this network, these measurements are used to analyze networks obtained from voting patterns. These measurements are:

- **Distance between two parties A and B:** "Used to quantify the level of coalition between A and B"
- **Average distance dG:** "Average distance between all nodes in the network"
- **Political fragmentation for group A, F(A):** "Intragroup dispersion"

Although there are various papers analyzing political datasets via network science methods and papers of social scientists analyzing Turkish election results, there are not any significant work applied for Turkish election results or political datasets which includes utilization of networks. Last but not least, even though we are inspired by such work, we have combined these ideas and used some other methods that are described further in the following part.

3. Methods and Datasets

The research for the dataset was made, considering the characteristics of selected deputies and city demographics. Deputy characteristics include profession, which political party they support, and which city they were born in, while city demographics include education, age, and gender distribution of the cities. Mentioned data were acquired from various sources such as published data from the High Election Board, official records of the Grand National Assembly of Turkey, and statistics from Biruni University. After the acquisition of the data, a system of relation and interconnectivity between those variables was developed and transformed the data according to that system by eliminating unrelated content and creating a new data frame manually.

In conclusion, two datasets were created, one being the deputy characteristics dataset (661 deputies, 6 features) and the other city demographics dataset (81 cities, 2 features). For cities dataset, cosine similarity of each city based on these 2 features to construct the network. After creating the matrix using networkx it was converted into a graph and then it is visualized by Gephi.

For deputy dataset one hot encoding was conducted since there were many different professions. However, after the process it was realized that the data expanded dramatically (+80 features - unique professions). So that, feature extraction was performed. The professions were limited to 5 categories, which were; 1.business person, 2.executive, 3.politician, 4.lawyer, 5.social scientist, 6.doctor, 7.engineer/natural scientist. Such extracted dataset was used in network construction, the other dataset with specific professions were used in exploratory analysis to show the professions in a more detailed way. However, such steps were still not enough for the network construction since there were still too many nodes and the network was hard to read. Also, to easily compare the voter and deputy networks, group_by method was used to group the deputies with their cities. Then, the percentage of each row (city) has been calculated to have the final dataframe.

While constructing the network for the deputy dataset, cosine similarity was used too. Such metric was selected since it calculates the similarity between each variable and is one of the most efficient metrics. Also, after constructing some other networks with other metrics it was seen that the cosine similarity was the most suitable one.After that, for such cosine similarity matrices, community detection is also pursued.

For exploratory analysis, the detection of the most/least educated cities were important. For such aim, the weighted averages of the education levels were taken. As a result, Gaziantep, Ankara and Adana were chosen to explore the data more.

3. Results

a. Network Construction

Our network is constructed based upon the voters dataset, which holds the data of the electors based on cities. Csv is read into DataFrame and then by using the Networkx module, we convert the DataFrame into Graph. Since we are interested in finding important connections based on the features of the dataset, calculating cosine similarity and distance, we convert our network in a form such that interconnections can be understood. The first form of the graph is visualized in this way: (node names represent the given numbers of the cities.)

0	ADANA	20	BURSA	40	İZMİR	60	NEVŞEHİR	80	ZONGULDAK
1	ADIYAMAN	21	ÇANAKKALE	41	KAHRAMANMARAŞ	61	NİĞDE		
2	AFYONKARAHİSAR	22	ÇANKIRI	42	KARABÜK	62	ORDU		
3	AĞRI	23	ÇORUM	43	KARAMAN	63	OSMANİYE		
4	AKSARAY	24	DENİZLİ	44	KARS	64	RİZE		
5	AMASYA	25	DİYARBAKIR	45	KASTAMONU	65	SAKARYA		
6	ANKARA	26	DÜZCE	46	KAYSERİ	66	SAMSUN		
7	ANTALYA	27	EDİRNE	47	KIRIKKALE	67	SİİRT		
8	ARDAHAN	28	ELAZIĞ	48	KIRKLARELİ	68	SİNOP		
9	ARTVİN	29	ERZİNCAN	49	KIRŞEHİR	69	SİVAS		
10	AYDIN	30	ERZURUM	50	KİLİS	70	ŞANLIURFA		
11	BALIKESİR	31	ESKİŞEHİR	51	KOCAELİ	71	ŞIRNAK		
12	BARTIN	32	GAZİANTEP	52	KONYA	72	TEKİRDAĞ		
13	BATMAN	33	GİRESUN	53	KÜTAHYA	73	TOKAT		
14	BAYBURT	34	GÜMÜŞHANE	54	MALATYA	74	TRABZON		
15	BİLECİK	35	HAKKARİ	55	MANİSA	75	TUNCELİ		
16	BİNGÖL	36	HATAY	56	MARDİN	76	UŞAK		
17	BİTLİS	37	IĞDIR	57	MERSİN	77	VAN		
18	BOLU	38	ISPARTA	58	MUĞLA	78	YALOVA		
19	BURDUR	39	İSTANBUL	59	MUŞ	79	YOZGAT		

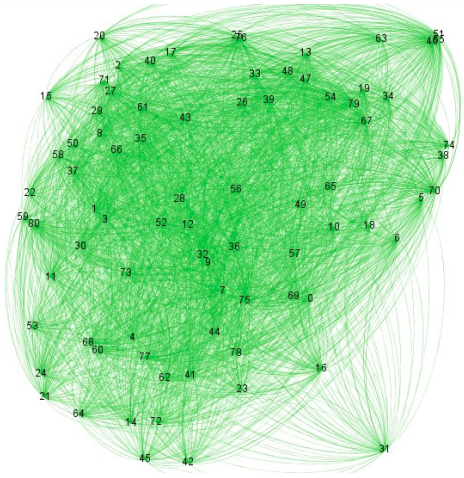

Figure 1.

It can be seen that this network is huge and it consists of some redundant properties. In order to get rid of the redundancies and transform the network to a more efficient one, disparity filtering is used. Disparity filtering gets rid of the edges that are not relevant or have little information. Also the edges are scaled in order to calculate distances and better understand the relevant data. The next phase of the network is visualized :

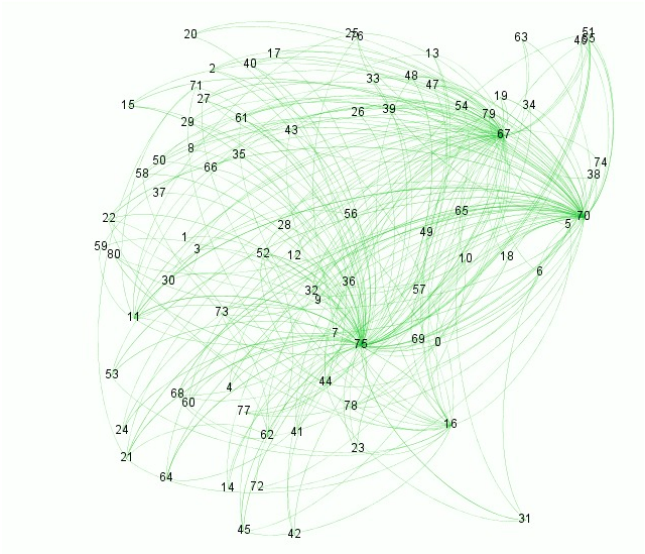

Figure 2.

Using the same methods the network of deputies of each city have been constructed. In order to decrease the number of nodes and not cause an overcrowding network (of 660x660 cosine similarity matrix), group_by method is used to sum each city's deputy characteristics.

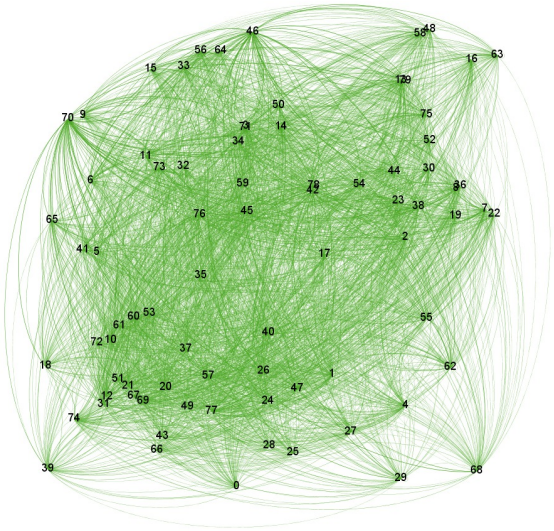

Figure 3.

The filtering also have been conducted to network in order to eliminate the low-degree edges, without eliminating any nodes.

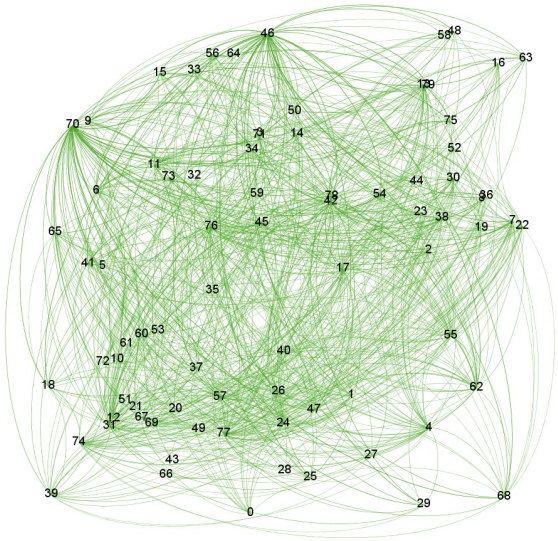

Figure 4.

b.Network Statistics

	Avg. Degree	Avg Weighted Degree	Density	Modularity	Avg. Clustering Coefficient	(Nodes, Edges)
Deputy Network	79.775	22.389	1.01	0.09	0.996	(80, 3191)
Voter Network	22.825	11.855	0.289	0.09	0.483	(80,913)

From the statistics above, it can be said that the deputies have more connections with each other even though the group_by method is used and node number is decreased. So that, the deputy network is more dense and has a higher clustering coefficient.

c. Exploratory Data Analysis

The analyze results on profession distribution for one of the least educated cities (Gaziantep) and one of the most educated cities (Ankara) are the figures below,

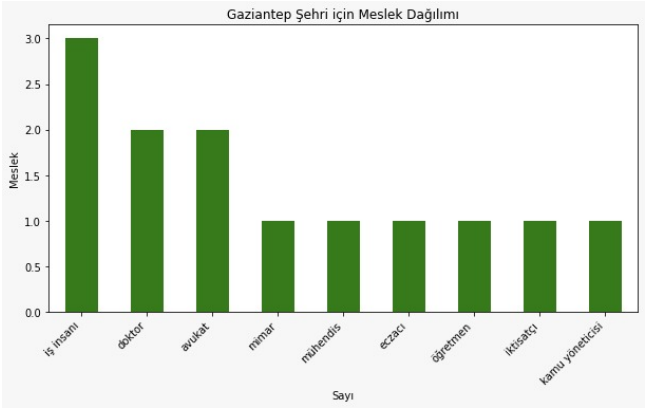

Figure 5.

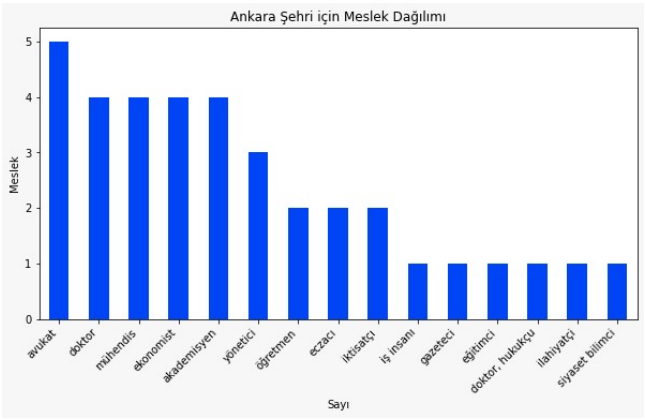

Figure 6.

The analysis findings about whether cities elect the deputies from their hometown was conducted,

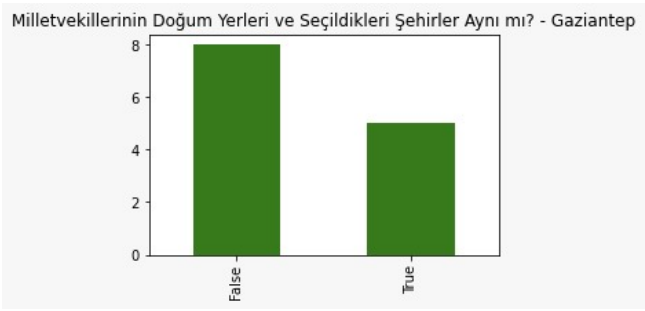

Figure 7.

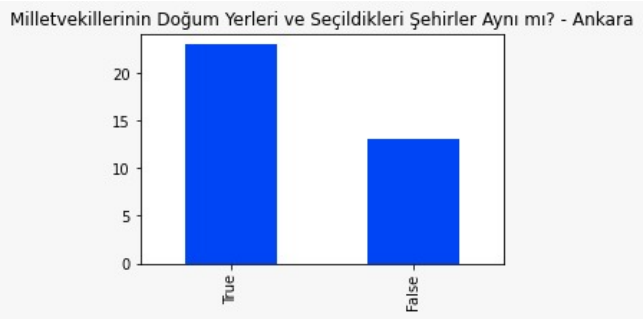

Figure 8.

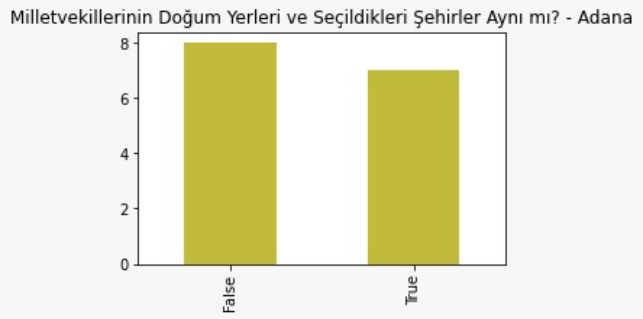

Figure 9.

Then, a city-base analysis about the age distribution of the deputies of Ankara and Antep have been conducted,

Figure 10.

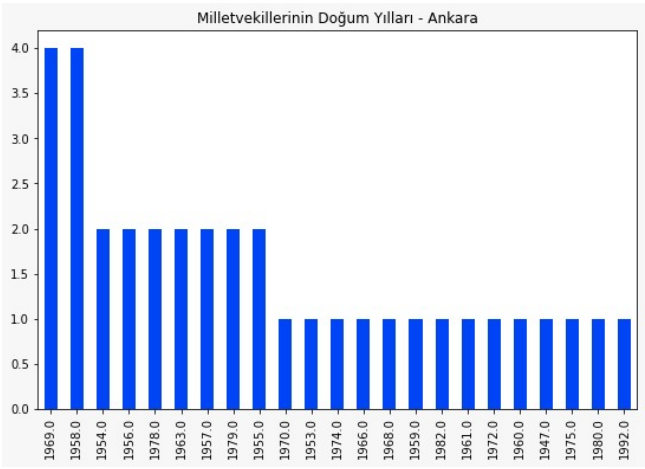

Figure 11.

Lastly, a final visualization of the professions of all deputies from AKP, CHP and MHP was conducted.

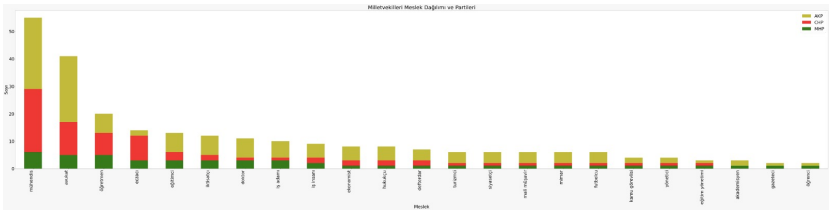

Figure 12.

4. Discussion

Using the datasets explained, various analysis was conducted before and after constructing the network. Most importantly, from the voters' network (see Figure 2.), it is seen that nodes of the cities in similar regions are closer to each other. For instance, the cities such as Siirt, Batman and Malatya are in similar places; such cities are geographically close to each other and such closeness continues in the network of closeness related to education and age distribution. However, some cities have been closer to some other groups unexpectedly. For instance, unexpected cities such as İzmir, Antalya have places near some eastern cities. Such condition was also the case for the deputies' network. This time, in most of the cases cities' demographic similarities showed interconnectedness with deputies' characteristics. (see Figure 4.) To prove such hypothesis community detection has been pursued, however, the method was not suitable since the degrees of the networks were high. Filtering also would require eliminating a high number of cities, such process would reduce the reliability.

Analyzing the city demographics data, one of most educated and least edu-

cated cities were analyzed. The main goal was to prove that highly educated cities elect deputies that also have high level of education, and vice versa for low-level cities. As an educated city Ankara, the most common deputy professions were lawyer (= 'avukat'), doctor(= 'doktor'), engineer (= 'mühendis') and academics (= 'akademisyen'). Such professions were what was expected for an educated city to elect. So that, our data supported the hypothesis on educated cities electing the educated deputies. On the other hand, as a less educated city, Gaziantep was analyzed. City's elected deputy professions were mainly business people (= 'iş adamı') - 23%, lawyer and doctor. Business people having the lead on a less educated city was also expected, the dataset also provided evidence for the hypothesis on deputy profession-city's education interconnectedness. (see Figure 5 and 6.)

As a part of the hypothesis, whether cities elect the deputies from their hometown or not. In analyzed cities, deputies that are not from that town were selected. However, in a city that is well educated, like Ankara, it was seen that the hometown deputies were almost twice the number of the deputies from another town. Such finding was partly inconsistent with the dataset and can open new paths to new hypotheses about the subject. (see Figure 7, 8 and 9.)

As a city-base analysis, the age distribution of the deputies have been conducted. Well-educated cities being open to elect younger deputies was one of the hypotheses. Analyzing Ankara, it was found that the city elected younger deputies who were born in 1992 and 1980. (see Figure 10 and 11.)

Additionally, political parties' preferred deputy professions were analyzed. AKP, CHP and MHP's preferred professions showed similarity. The most preferred professions were again engineer, lawyer and teacher (= 'öğretmen'). (see Figure 12.)

5. Conclusion

To conclude, it can be said that, looking at the networks and the charts given in results, the hypothesis on city and deputy interconnectedness holds in many cases; voters of the geographically and demographically closer cities elect the deputies with similar characteristics. However, there exists some cases in our networks where such hypothesis does not hold.

It is also clear that we have extracted some common rules among all cities that gives us a clear boundary for the profile of the selected deputy and such boundaries can be important for the next elections; according to our deputy - date of birth data, the majority of the selected deputies are born in between the years 1950 - 1970, which tells us that the voters tend to vote for middle-aged

deputies. The reason for this seems to be tightly bound with the cultural values of Turkey. Turkish voters seem to trust the middle-aged congress persons for representing them and doing the right thing for their cities.

In terms of the field of education, our data indicates that voters tend to vote more for the candidates with social-science backgrounds like economics, law and education. Even though the number of candidates with non-social-science-backgrounds are low in comparison, it is still seen that candidates with social-science backgrounds win against them. This distribution difference between candidates with social-science-backgrounds and candidates with non-social-science-backgrounds might have a deeper meaning and logical reasons. When we look at the major problems of developing countries like Turkey, the problems are always centered around three big topics; which are law, economics and education. So during the pre-election phase, the promises given by the area experts by using the authority coming from their field of expertise might have excited the voters and encouraged them to vote for these candidates and expect them to fix these major problems.

References

Brito ACM, Silva FN, Amancio DR (2020) A complex network approach to political analysis: Application to the Brazilian Chamber of Deputies. PLoS ONE 15(3): e0229928. https

The Grand National Assembly of Turkey. TBMM Yabancı Dillerde Tanıtım Sitesi. Retrieved June 22, 2022,
from https://global.tbmm.gov.tr/index.php/EN/yd/icerik/14

Turkey election results - June 24 Presidential Parliamentary polls. (n.d.). Retrieved June 22, 2022, from https://www.dailysabah.com/election/june-24-2018-election-results

Seçmen Gruplari ve Eğilimleri- teamarastirma.com. Teama Araştırma. (n.d.). Retrieved June 26, 2022, from http://www.teamarastirma.com/wp-content/uploads/2021/04/TEAM-Se

Ertem, Özlem. "Tanzimattan Günümüze Türkiye'de Parlamentonun Sosyal Profili." İstanbul: İstanbul Üniversitesi (Yayımlanmamış yüksek lisans tezi) (2010).