Using complex networks to analyze the Brazilian Chamber of Deputies

Ana Caroline Medeiros Brito

Dissertação de Mestrado do Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional (PPG-CCMC)
Ana Caroline Medeiros Brito

Using complex networks to analyze the Brazilian Chamber of Deputies

Master dissertation submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP, in partial fulfillment of the requirements for the degree of the Master Program in Computer Science and Computational Mathematics. FINAL VERSION

Concentration Area: Computer Science and Computational Mathematics

Advisor: Prof. Dr. Diego Raphael Amancio

USP – São Carlos
October 2020
Usando redes complexas para analisar a Câmara dos deputados brasileira

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestra em Ciências – Ciências de Computação e Matemática Computacional. VERSÃO REVISADA

Área de Concentração: Ciências de Computação e Matemática Computacional
Orientador: Prof. Dr. Diego Raphael Amancio

USP – São Carlos
Outubro de 2020
Aos meus pais, Edilene e Brito,
que, apesar da distância, sempre se fizeram presentes.
Gostaria de agradecer aos meus pais, que sempre estiveram dispostos a me escutar e me guiar quando tudo parecia confuso. Vocês são minha base e minha fonte de inspiração para continuar buscando meus sonhos. Sou eternamente grata por tudo que fizeram por mim até hoje.

Agradeço também ao meu orientador, professor Diego, por todos os ensinamentos e todas as oportunidades oferecidas. A Filipi Nascimento e ao professor Luciano Costa, que também me orientaram com cuidado durante esse mestrado. A Henrique Arruda, por ser sempre paciente comigo e tirar minhas dúvidas de iniciante. Aos demais colegas que conheci em minhas visitas ao IFSC, sempre atenciosos e bem humorados.

Aos colegas do NILC, em especial, Ana Carolina, Edresson Casanova, Henrico Brum, João Paulo, Jorge Valverde, Laura Cruz, Márcio Lima, Marco Cabezudo, Rafael Anchieta, Rafael Mineiro, Raphael, Renata Ramish, Rogério Figueiredo, Roney Lira, Sidney Leal e Xiana. Obrigada por terem me acolhido tão bem. Saudades das nossas conversas acompanhadas de bolos, pavês, churros e doces peruanos. Ir ao laboratório se tornou divertido.

Aos amigos da graduação, que continuam me acompanhando mesmo agora que estamos trilhando caminhos distintos, Esther Bárbara, Fernanda Isabel, João Pedro Holanda, Karyd’ja Souza, Lucas Torres, Luís Claúdio, Luis Tertulino, Luísa Rocha, Raquel Oliveira, Pitágoras Alves, Priscila Bessa e Ronaldo Figueiredo.

Aos meus amigos, Morena Gil, Luan Lédio e Sofia Aravanis. Guardo muito carinho e amor por vocês desde o Ensino Médio. Apesar da distância, sempre que nos reencontramos é como se nada tivesse mudado.

Ao meu primo e amigo, Elton Cardoso, por todas as conversas e todos os cafés que tornaram a quarentena mais suportável. Obrigada por ter revisado esse manuscrito, mesmo não sendo sobre sua área de pesquisa.

Ao meu amigo e companheiro, Luan Martins. Sou muito grata por todo carinho e paciência que teve comigo até hoje. Sem dúvidas, você foi uma pessoa muito importante ao longo desses dois anos e sou feliz por ter compartilhado todos esses dias contigo.

Finalmente, gostaria de agradecer ao ICMC. A toda equipe de profissionais incríveis e atenciosos, sempre prontos para atender bem os alunos. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.
“Há um tempo em que é preciso abandonar as roupas usadas, que já tem a forma do nosso corpo, e esquecer os nossos caminhos, que nos levam sempre aos mesmos lugares. É o tempo da travessia: e, se não ousarmos fazê-la, teremos ficado, para sempre, à margem de nós mesmos.”

(Fernando Pessoa)
Redes complexas são amplamente utilizadas para modelar indivíduos e suas interações sociais. Neste trabalho, utilizamos essa representação para investigar dados de votações de legisladores brasileiros na Câmara dos Deputados (1991 a 2019). Ao gerarmos redes de deputados como vértices conectados pela similaridade de seus comportamentos votando proposições, desejamos caracterizar o comportamento de entidades políticas: coesão de partidos políticos, coalizões e suas dinâmicas de formação e dissolução. Para esse fim, definimos um framework que consiste em procedimentos para a geração das redes de votação por ano e por janelas sobrepostas de três meses; remoção de arestas pouco significativas; identificação de comunidades; e definição das medidas de *isolamento* e *fragmentação* para caracterização de grupos. A análise das redes anuais possibilitou uma visão mais geral do cenário político brasileiro. As redes mais recentes parecem menos coesas do que as mais antigas, indicando que os deputados estão, em geral, concordando menos entre si. As comunidades identificadas tendem a corresponder às estruturas de coalizão entre partidos. Além disso, o número efetivo de comunidades revela que há poucos grupos em relação a ampla quantidade de partidos políticos existentes no sistema brasileiro. A caracterização dos partidos políticos com o uso das medidas topológicas revelou alguns fenômenos interessantes, como mudanças de coalizão associadas a mudanças de mandato. A análise detalhada dos meses antes, durante e depois do impeachment de Dilma Rousseff revelou que ela não tinha apoio da Câmara dos Deputados bem antes de deixar seu cargo.

**Palavras-chave:** Redes complexas; Redes complexas sociais; Ciência política; partidos políticos brasileiros.
ABSTRACT

BRITO, A. C. M. Using complex networks to analyze the Brazilian Chamber of Deputies. 2020. 69 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2020.

Complex networks are widely used to model individuals and their social interactions. In this research project, we use this representation to explore the voting data of Brazilian legislators of the Chamber of Deputies (from 1991 to 2019). When generating co-voting networks consisting of deputies as nodes connected by their similarity in their voting propositions behaviors, we aim to characterize political entities’ behavior: political parties’ cohesion, coalitions and their dynamics of emergence and dissolution. To this end, we defined a framework consisting of procedures in order to generate co-voting networks by year and, alternatively, by overlapping windows of three months; removal of less meaningful edges; communities identification; and the definition of isolation and fragmentation measures for groups characterization. The yearly networks analysis allows an overview of the Brazilian political scenario. For example, the most recent networks seem less cohesive than those in the early years, indicating that the deputies are overall in less agreement. The identified communities tend to represent the coalitions established among political parties. Furthermore, the effective number of communities reveals that there are few groups in comparison to the wide number of political parties in the Brazilian system. The characterization of these groups by using topological measures, revealed some interesting phenomenons, such as coalitions changes correlated with terms changes. The detailed analysis of the months before, during, and after the Rousseff’s impeachment revealed that she did not have support of the Chamber of Deputies months before she left office.

Keywords: Complex networks; Social complex networks; Political science; Brazilian political parties.
LIST OF FIGURES

Figure 1 – Four network models are presented. The darker the color, the higher the vertex degree. ......................................................... 25

Figure 2 – Disparity filter example. In (a), there is the original weighted network. In (b), there is the $\alpha_{ij}$ values calculated following the Equation 2.7. When $\alpha$ is set to 0.8, the dashed lines are edges removed. ........................................ 29

Figure 3 – Number of propositions by year. .......................................................... 36

Figure 4 – Radar map view of the Brazilian political parties. ............................... 41

Figure 5 – Number of seats distributed among political parties in the Chamber of Deputies changing over the years. .............................................. 52

Figure 6 – Average shortest path lengths over the years. The values ranges approximately between 2 and 5. ................................................................. 52

Figure 7 – Comparison between the modularity values of the communities and the political parties. ................................................................. 53

Figure 8 – NMI between communities and political parties. ................................... 53

Figure 9 – Comparison between diversity and size of the communities and political parties partitions. .............................................................. 54

Figure 10 – Visualization of the co-voting networks evolving over time. ................. 55

Figure 11 – Average shortest path length between a political party and all the others. The background color indicates the political party in the presidency in the given period. In particular, the gray area corresponds to the Dilma Rousseff’s process of impeachment. ......................................................... 57

Figure 12 – Isolation and fragmentation of the political parties. The background color indicates the political party in presidency in the given period. In particular, the gray window corresponds to the Dilma Rousseff’s process of impeachment. 58

Figure 13 – Political parties distribution among communities (2000, 2002, and 2003). . 59

Figure 14 – Political parties distribution among communities (2011, 2012, 2015, and 2016). 60

Figure 15 – Political parties distribution among communities in 2018 and 2019. ........ 61
## CONTENTS

1 INTRODUCTION ................................................................. 19
   1.1 Goals ................................................................. 20
   1.2 Structure of the document ......................................... 20

2 THEORETICAL FOUNDATION .............................................. 23
   2.1 Complex network definition ...................................... 23
   2.2 Network representation and topological features .............. 24
       2.2.1 Paths and walks ........................................... 26
       2.2.2 Entropy and diversity ...................................... 26
       2.2.3 Modularity ................................................... 27
   2.3 Community detection ............................................... 27
   2.4 Edge pruning ........................................................ 28
   2.5 Chapter summary .................................................. 28

3 PREVIOUS CONCEPTS ........................................................ 31
   3.1 Brazilian presidential system ..................................... 31
       3.1.1 Propositions ................................................ 32
       3.1.2 Political parties ............................................ 32
       3.1.3 Party cohesion and coalitions ............................. 33
       3.1.4 Impeachment ................................................ 33
   3.2 Our dataset .......................................................... 35
   3.3 Chapter summary ................................................... 36

4 RELATED WORKS .............................................................. 39
   4.1 Political science .................................................... 39
       4.1.1 Poole and Rosenthal (1985) ............................... 39
       4.1.2 Leite and Trento (2018) ................................... 40
   4.2 Network science ..................................................... 41
       4.2.1 Maso et al. (2014) .......................................... 42
       4.2.2 Cherepnhalkoski et al. (2016) ........................... 42
       4.2.3 Levorato and Frota (2016) ................................. 43
       4.2.4 Ferreira, Matos and Almeira (2018) ...................... 44
   4.3 Chapter summary ..................................................... 45
Complex networks have great potential for applications, in particular, for those involving individuals and their interactions, named as social networks. They can be used to model dynamics such as individuals influence (TANG et al., 2009), members characterization in online social networks (KUMAR; NOVAK; TOMKINS, 2010), opinion formation (JALILI, 2013), scientists dynamics evolving over time (VIANA; AMANCIO; COSTA, 2013), and social groups identification using temporal interactions series (SEKARA; STOPCZYNSKI; LEHMANN, 2016).

Whereas this wide diversity of possibilities, this Master’s research is interested in data mining using complex network to analyze the voting data of the deputies in the National Congress lower house of Brazil, also known as Chamber of Deputies. The Brazilian system of governance is classified as a democracy since there are periodic popular elections to choose representatives. The behavior of the Congress members is a recurring topic of study in Political science. In general, the representatives are distributed among political parties, so concepts such as cohesion and coalitions are usual to characterize the profile of these groups in the government. Ideally, each party represents a different ideological position, but issues ranging from bureaucratic obstacles to members’ private interests make the understanding this system not trivial.

There are a few studies exploring voting data using complex networks, in particular, we are interested in co-voting networks where the legislators are connected by the similarity of their voting patterns. When using this type of network, topological metrics may go alongside Political science concepts such as cohesion and coalition (CHEREPNALKOSKI et al., 2016). The identification of communities, i.e., of groups of individuals densely connected internally, reveals the political parties divided into the government and opposition (MASO et al., 2014). Therefore, we consider complex networks appropriate to analyze voting data. In this research, we used network mining techniques to analyze political events such as changes of term and impeachments.
We choose to use voting propositions of Brazilian deputies because of several facts. First, the Brazilian system of governance is a fairly recent democracy under the constant effects of unethical practices common among legislators and maladministration. Brazil have recently undergone political transformations, such as Rousseff’s impeachment in 2016. Brazil is a vast and diverse country in ethnic and cultural aspects. Therefore, it presents a plurality of ideas that probably reflect a large number of political parties defending different perceptions.

1.1 Goals

The main goal of this Master’s project was to analyze the Brazilian national legislative according to its voting patterns using complex networks. We also incorporated the temporal aspect when generating the networks, which revealed the evolving dynamics hidden in the data. Our primary purpose is detailed below:

- To characterize groups of individuals using topological metrics extracted from the networks;
- To compare the communities and political parties partitions;
- Identification of emergence and dissolution of coalitions.

We aimed to characterize the Brazilian political parties, cohesion, coalition dynamics, and the impacts of terms changes in the organization of this system. We identified distinct groups of deputies that, in general, do not correspond to the Brazilian political parties. The analysis of the government strength during Rousseff’s term revealed PT’s primary coalition dissolution before impeachment. We hope the methodology applied in this research could also be useful for analyzing data from other countries or even other scopes where the interactions between groups of individuals are the object of study.

Finally, in systems of governance where the people elect the representatives, it is essential to check how these individuals are organized and which ideas are defending. In this way, the institutions are maintained, and the people may benefit from the public agents’ actions. One result of this research is to provide a systematic procedure to investigate these political dynamics, increasing the mass of people who can access this information.

1.2 Structure of the document

- Chapter 2 presents the theoretical foundation necessary to understand the work developed during this research;
1.2. Structure of the document

- Chapter 3 explains briefly the Brazilian system of governance and some political concepts essential as well as introduces the dataset;

- Chapter 4 includes related works from Political science and Network science;

- Chapter 5 presents the methodology used to generate and analyze the co-voting networks obtained from the Brazilian voting data;

- Chapter 6 presents and discuss the main results obtained during this research;

- Chapter 7 recalls the research’s goals and the corresponding final results alongside a brief discussion of limitations and future works.
This chapter presents the main theoretical foundation necessary to understand this research. Here the basic concepts related to complex networks are defined and explained, starting with the network definition and some well-known models such as Small-world and Scale-free. Then, we detail some topological metrics, for example, entropy and modularity. The final sections review community detection algorithms and edge pruning techniques.

2.1 Complex network definition

A network is a graph $G = \langle V, E \rangle$, where $V$ is a set of vertices, and $E$ is a set of edges. An edge is a connection between two vertices $\langle v_i, v_j \rangle$. It can be undirected or directed, indicating if the relation is mutual or not. When directed, the order matters to interpret the relation between the entities represented in the system. In the Network science terminology, vertices are usually called **nodes**, and edges are called **links**. In essence, a complex network can be defined as a network that differs statically from regular graphs such as random graphs or lattices (COSTA et al., 2007).

The increase of Network science’s sub-research fields is associated with its potential modeling real-world systems. The World Wide Web, roads, protein connections, and human relationships are just some examples of data that can be modeled as networks. Despite this effective relation with real-world data, there are synthetic networks that are also complex. Artificial or synthetic networks can be understood as an informal term to define networks generated based on a model or on synthetic data, in contrast to networks generated based on real-world data.

Networks are classified by the type of data represented in the system or by particular topological proprieties. Biological, Social networks, textual networks, information networks are examples of classifications by the type of information. To understand and characterize these
networks better, some classical topological models can be useful to comparison: random graphs, small-world, scale-free, and community structures (COSTA et al., 2007). A random graph, in general, is an random object constructed to match some specific properties of a real-world network, being used as an unbiased reference to verify if the network of interest displays non-trivial features. Figure 1a shows an example of a graph generated following the random model of Erdos and Rényi implemented using an igraph routine called Graph.Erdos_Renyi with parameters \( n=64 \) and \( p=0.2 \).

In the Small-world model, most of the vertices in the network are closely connected, i.e., the number of vertices connecting any pair of vertices is usually small. This property is related to the predominance of cycles of order three. That is, if vertex \( i \) is connected to \( j \) and \( k \), there is a high probability of \( j \) and \( k \) be connected. This property is more frequent in the Small-world model than in a random graph. Metabolic networks and words co-occurrence networks are examples of Small-world networks found in real-world (WAGNER; FELL, 2001; CANCHO; SOLÉ, 2001). The network in Figure 1b was generated using an igraph routine called Graph.Watts_Strogatz, with parameters \( dim=2, size=8, nei=2, p=0.2 \).

The Scale-free model is known for the preferential attachment propriety. In networks of this type, some vertices have a large number of connections (called hubs), while the majority has few connections (see Figure 1c). The World Wide Web is one of the most recurrent examples of a real-world network with Scale-free property (ALBERT; JEONG; BARABÁSI, 1999). Figure 1c shows an example of Scale-free network following the Barabási and Albert model, it was generated using an igraph routine called Graph.Barabasi with parameters \( n=64 \) and \( m=5 \).

Graphs with community structure have a specific topological characterization, as shown in Figure 1d. Some different definitions have been proposed in the literature. The one used in this work, according to Girvan and Newman (2002), defines communities as subsets of vertices more connected internally than with vertices of different communities. Co-authorship networks are examples of networks with community structure. The network in Figure 1d was generated following the algorithm proposed by Girvan and Newman (2002), for 64 vertices randomly distributed among four communities, where the probability \( p_{in} \) of an edge be placed between vertices in the same community is 0.6 and the probability \( p_{out} \) of an edge be placed between vertices in different communities is 0.01.

### 2.2 Network representation and topological features

One possible way to represent a network is by mapping it to an adjacency matrix. The vertices can be identified by integer values \( \{1, 2, \ldots, N\} \), given a matrix \( A \), the element \( a_{ij} \) represents a link between the vertices \( i \) and \( j \). If the value in this position is 1, then there is a link between them; otherwise, when the value is 0, there is no connection. For a weighted network, an additional matrix representing the weights can be used. Given a matrix \( W \), the element
2.2. Network representation and topological features

Figure 1 – Four network models are presented. The darker the color, the higher the vertex degree.

(a) Random model of Erdos and Rényi (ERD et al., 1959). Total of vertices: 64. Total of edges: 392.
(b) Small-world of Watts and Strogatz (WATTS; STROGATZ, 1998). Total of vertices: 64. Total of edges: 384.
(c) Scale-free of Barabási and Albert (BARABÁSI; ALBERT, 1999). Total of vertices: 64. Total of edges: 305.
(d) Network with community structures of Girvan and Newman (GIRVAN; NEWMAN, 2002). Total of vertices: 64. Total of edges: 321.

Source – Prepared by the author.

$w_{ij}$ represents the weight of the edge between the vertices $i$ and $j$. If the network modeled is undirected, both $A$ and $W$ are symmetric.

Two vertices are said neighbors or adjacent when there is an edge connecting them. The number of neighbors of a vertex is its degree,

$$k_i = \sum_{j=1}^{N} a_{ij}, \quad (2.1)$$

and, for weighted networks, the vertex strength,

$$s_i = \sum_{j=1}^{N} w_{ij}, \quad (2.2)$$
is more used. These metrics are examples of topological features that are useful to evaluate how the vertices (or nodes) and edges (or links) are arranged in the system.

### 2.2.1 Paths and walks

Path and walks are two related concepts that are sometimes mistaken. Both of them are important to understand some of the algorithms used in this work. In a graph, a walk is a sequence of edges (finite or infinite) connecting two vertices, being possible edges repetitions. A random walk is a walk defined by random steps, usually following some probability distribution of interest. A path is a specific type of walk composed necessarily by unique vertices and edges. A particular type of path, the shortest path is the shortest sequence of vertices connecting two vertices (source and target) in the graph (BONDY; MURTY et al., 1976). The average shortest path length of a network is defined as

\[
L = \frac{1}{N(N-1)} \sum_{i \neq j} l(v_i, v_j),
\]  

where \( l \) is the shortest path length between the vertices \( v_i \) and \( v_j \).

When two vertices are connected, the path length can be defined as a function of the number of edges. For weighted graphs, the length can be the sum of the edge weights, to remain meaningful, the values must be distances (not similarities). Some networks have more than one component, which implies in disconnected vertices. The path length, when there is no path, can be set as infinite.

### 2.2.2 Entropy and diversity

Entropy measures the order or disorder in the system. It is a general concept in thermodynamics, statistics, and information theory. For a probability distribution \( P \) of a random variable \( X \), this metric is defined as

\[
H(X) = - \sum_{i} P(x_i) \ln P(x_i).
\]  

Diversity is calculated as

\[
D(X) = e^{H(X)}.
\]  

This concept, common in Ecology, is interpreted as the effective number of distinct species in a group (CHAO; CHIU; JOST, 2016). Alternatively, when applied in Network science, can be used with discrete vertex attributes and topological metrics, for instance, labels and degree distributions (WANG et al., 2006). Given \( R \), defined as the number of different values that \( X \) assumes, if the \( x_i \) occurrences appear equally distributed among the vertices, then \( D(X) \) is equal to \( R \). Otherwise, if the distribution is biased, \( D(X) \) tends to be substantially lower than \( R \).
2.3. Community detection

2.2.3 Modularity

Modularity is a metric used to evaluate the quality of vertex clusters (or vertex partitions). For a graph represented by its adjacency matrix $A$, its modularity is defined as

$$Q = \frac{1}{2m} \sum_{ij} \left[ a_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j),$$

(2.6)

where $m = \frac{1}{2} \sum_{ij} a_{ij}$ is the number of edges in the graph, $c_i$ is the cluster of vertex $i$ (usually its community), $\delta(c_i, c_j)$ is the Kronecker delta function, its value is 1 if $c_i = c_j$, 0 otherwise (CLAUSET; NEWMAN; MOORE, 2004). This metric compares the number of edges that belong to the given partitions versus a random organization of the edges. This metric can be applied to other types of partitions than communities. For weighted graphs, $k_i$ values can be replaced by $s_i$.

2.3 Community detection

Modularity is used to evaluate the quality of community detection algorithms, and, sometimes also used by algorithms during the process to find out communities. The approaches proposed in Blondel et al. (2008) and Traag, Waltman and Eck (2019) are examples of this kind of algorithms.

Blondel et al. (2008) use modularity as the quality function to link (merge) two clusters in the hierarchical clustering classical algorithm. In the first step, each vertex is addressed as a community of one element. Then, suppose moving each vertex from a community to another, the change that provides the largest increase in modularity value is selected. These moves are repeated until $Q$ stops increasing. Then, in the second step, a new aggregate network is generated; each community found in the first step is transformed into a vertex. All edges connecting vertices of two communities $c_i$ and $c_j$ are joined in one single edge connecting the “new vertices” $c_1$ and $c_2$. The new weight is the sum of all weights of the joined edges. Edges inside the same community become self-loops. These two steps are repeated until the modularity stabilizes. The output is a hierarchical structure (or dendrogram), the number of communities found depends on the point choose to cut it. This method is also known as Louvain, the university where its authors were associated.

Recently, the work proposed in Traag, Waltman and Eck (2019) pointed out some issues in the Louvain algorithm, showing that it can generate communities with disjoint vertices. A new step is added between the two well-known steps from the original algorithm (BLONDEL et al., 2008) to ensure producing only communities composed by one component. For each community found in the first step, each vertex is randomly linked with others if they are originally connected, and if the modularity increases. As a result, it is possible to divide the community found in two or more when there are disjoint vertices.
In Network science, one way to evaluate community structure identification is using Normalized Mutual Information (NMI), which calculates the correlation between two cluster partitions, the one found by the algorithm tested and the known ground truth (DANON et al., 2005). If the partitions are totally compatible, the NMI value is 1; alternatively, if they are entirely incompatible, the value is 0. This metric can be also used to compare how similar the outputs of two different community detection algorithms are.

### 2.4 Edge pruning

Information extraction can be affected by the network size, considering the fast growth in the algorithm complexities associated with the increase in the number of vertices and edges. In this context, edge pruning is an alternative to reduce complexity and clear the data.

One simple approach is defining a weight threshold, so all edges with weight lower than the threshold are removed from the network. In Yan et al. (2018), the authors affirm that the community structure is strong enough to survive after edge pruning processes. However, there are more refined alternatives, such as the proposed in Serrano, Boguná and Vespignani (2009). This algorithm finds the most relevant connections, enabling the removal of all others (not relevant edges). This method ideally better preserves the community structures than the simple threshold pruning.

For each vertex, the weights are normalized by the corresponding strength. In an initial analysis, the higher normalized values indicate the most relevant connections. However, sometimes there are random fluctuations among the real connections, then the authors proposed a filter to identify and remove these noises. The disparity filter generates a random distribution of the weights for a vertex, a null model. The edges that rejected the null model, given a $\alpha$ parameter representing the minimum level of significance desired, are considered relevant heterogeneities that must be preserved. For each weight $w_{ij}$ a corresponding $\alpha_{ij}$ is calculated as

$$\alpha_{ij} = 1 - (k_i - 1) \int_0^{w_{ij}} (1 - x)^{k_i - 2} dx < \alpha.$$  \hspace{1cm} (2.7)

The minimum value between $\alpha_{ij}$ and $\alpha_{ji}$ is used, once it is enough to preserve the edge if it is locally relevant to at least one of its vertices. Figure 2 shows the disparity filter applied to a toy example.

### 2.5 Chapter summary

In this chapter, we presented some of the main definitions and concepts common in Network science. Complex networks were defined as graphs with nontrivial topological patterns, and some classical models were briefly explained: random graphs, Small-world, Scale-free, and networks with community structure. These models are a reference to analyze real-world network
Figure 2 – Disparity filter example. In (a), there is the original weighted network. In (b), there is the $\alpha_{ij}$ values calculated following the Equation 2.7. When $\alpha$ is set to 0.8, the dashed lines are edges removed.

Properties. In Section 2.2, some basic topological features such as degree, strength, entropy, and modularity were detailed. Section 2.3 explained some approaches to identify communities in networks, which is useful to obtain group patterns. Finally, in Section 2.4, edge pruning was showed as a pre-processing method to clear the data, allowing better extraction of information.
In this chapter, there is a brief introduction to the Brazilian political organization, mainly focused on the legislative structure: propositions and Chamber of Deputies (the lower house of representatives). Political terminologies and historical events are explained, providing a time contextualization of the dataset. The dataset’s first records began in 1991. By that time, the deputies were already organized in a multi-party system, which still remains today. Cohesion and coalition (i.e., political alliances) can characterize these groups. Finally, the dataset attributes are detailed with some basic statistics, such as the most common vote options and the distribution of propositions by year.

### 3.1 Brazilian presidential system

The Brazilian political organization is currently a presidential democratic system, divided into three branches. The executive is separated from legislative as well as from the judiciary. The head of the executive is also the president, known as the head of state. In the federal government sphere, the president is elected by direct popular vote for a four years term with the possibility of one reelection. In municipal and state government levels, there are elections for mayors and governors, respectively.

The federal legislative body government is divided into the lower house, also known as the Chamber of Deputies, and the upper house, the Federal Senate. This model is called bicameralism, i.e., a legislature of two chambers. The Chamber of Deputies has 513 seats since 1993\(^1\). Deputies and senators are elected by direct popular vote. The number of deputies by state is proportional to its population size, while the number of senators is always set three by state. Their mandates have different duration: a deputy term is four years, and a senator term is

\(^1\) Source: <https://www2.camara.leg.br/transparencia/acesso-a-informacao/copy_of_perguntas-frequentes/deputados#8>. Accessed on 05/18/2019.
eight years long. Each state has a Legislative Assembly, and each city (or town) has a Municipal Assembly.

The judiciary power is independent, guaranteed by administrative and financial autonomy, its primary function is to guarantee individual, collective, and social rights\(^2\). The municipal government does not have an independent judiciary.

### 3.1.1 Propositions

A legislative proposition (or proposal) is a document submitted for validation, for example, bills (law projects) and constitutional amendments. Deputies, senators, citizens, and the president can present this document that must be analyzed and voted by the Chamber of Deputies and, if approved, voted by the Federal Senate. The final step is the president sanction. Some exceptions can occur, allowing the proposition to be sent to the upper house without being voted in the lower house.

The voting process is usually open, also named as roll call; in this case, everybody knows everybody’s votes. There is four voting options: “yes”, “no”, “abstention”, and “obstruction”. The latter is a way to postpone the vote or request the revision of the preposition document. If the leader of a political party indicates “obstruction”, then all the deputies in the group will absent, and, as a consequence, with fewer deputies than the requisite to a vote happens, they postpone the event until another appropriate time.

### 3.1.2 Political parties

A political party can be defined as a group of people sharing similar ideologies, attempting to get their candidates elected to defend and implement their agenda of a specific community or social movement (VIANA, 2003). This previous definition is more idealized than realistic. In practice, as we show in Chapter 6, political parties do not directly correspond to communities in the deputies co-voting networks. According to Political science, the parties have three main functions (AMARAL, 2013):

- To structure electoral competition;
- To aggregate interests;
- To conduct legislative work.

The dynamics in the Chamber of Deputies are mainly related to the second and third functions above: some alliances emerge to aggregate interests, and some groups disagree about what is being voted and discuss until finding the best way to combine their interests. The

\(^2\) Source: [http://www.brazil.gov.br/government/how-the-government-works/federal-judiciary-branch](http://www.brazil.gov.br/government/how-the-government-works/federal-judiciary-branch). Accessed on 05/20/2020.
3.1. Brazilian presidential system

president, deputies, senators, aldermen, mayors are usually affiliated with political parties. The current Brazilian political system is multi-party, the political parties that appear in our dataset are listed in Table 1, the names were obtained according to Wikipedia, some parties do not have initials or an English name available. The last presidents of Brazil are presented in Table 2.

3.1.3 Party cohesion and coalitions

The political parties are usually characterized by their political orientation, left or right, about different topics such as economy, health, international relations, and education. The group of political parties governing is also known as the government, and those ideologically opposed are the opposition. In this context, the political parties make connections with other parties aiming to be successful in approve propositions of their interest according to their political positions. In this context, alliances emerge. Some parties are more likely to dialogue with different groups, while others do not. In Political science, concepts emerge from these group behaviors, for example, cohesion and coalition.

Party cohesion is the similarity of political positions within the party before any group discussion. A related concept is the party discipline that indicates the party’s ability to control its own members’ votes parliament (TSEBELIS, 1995). Coalition or political alliance is a temporary accordance between members of different parties to achieve a common goal. During the elections, the coalitions are motivated by the common interest in being elected, but after this phase, parties with a common agenda can also make alliances to approve specific propositions in the Chamber of Deputies as well as in the other legislative assemblies.

3.1.4 Impeachment

The impeachment or cassation of the president term can be applied if he or she commits some high crime such as acts against:

- The Nation existence;
- The independence of the legislative, the judiciary and the constitutional power of the states;
- The exercise of political rights, individual and social;
- The internal security of the country;
- Probity in administration;
- Appropriation bill;

3 During the presidential campaign, Bolsonaro was affiliated with PSL. In 2019, he left this group and became independent.
4 Source: <http://www.planalto.gov.br/ccivil_03/leis/l1079.htm>. Accessed on 07/06/2020.
Table 1 – Brazilian political parties in the lower house. In bold, the groups with more representatives elected.

| Initials | Original name | English name |
|----------|---------------|--------------|
| –        | Avante        | Forward      |
| –        | Cidadania     | Citizenship  |
| DEM      | Demócratas    | Democrats    |
| MDB      | Movimento Democrático Brasileiro | Brazilian Democratic Movement |
| NOVO     | Novo          | New Party    |
| PAN      | Partido dos Aposentados da Nação | Party of the Nation’s Retirees |
| –        | Patriota      | Patriot      |
| PCB      | Partido Comunista Brasileiro | Brazilian Communist Party |
| PCdoB    | Partido Comunista do Brasil | Communist Party of Brazil |
| PDC      | Partido Democrata Cristão | Christian Democratic Party |
| PDS      | Partido Democrático Social | Democratic Social Party |
| PDT      | Partido Democrático Trabalhista | Democratic Labour Party |
| PEN      | Partido Ecológico Nacional | National Ecological Party |
| PFL      | Partido da Frente Liberal | Liberal Front Party |
| PHS      | Partido Humanista da Solidariedade | Humanist Party of Solidarity |
| PL       | Partido Liberal | Liberal Party |
| PMB      | Partido da Mulher Brasileira | Party of the Brazilian Woman |
| PMDB     | Partido do Movimento Democrático Brasileiro | Brazilian Democratic Movement Party |
| PMN      | Partido da Mobilização Nacional | Party of National Mobilization |
| PODE     | Podemos       | We can       |
| PP       | Partido Progressista | Progressive Party |
| PBB      | Partido Progressista Brasileiro | Brazilian Progressive Party |
| PLL      | Partido Pátria Livre | Free Fatherland Party |
| PPR      | Partido Progressista Reformador | Reform Progressive Party |
| PPS      | Partido Popular Socialista | Popular Socialist Party |
| PR       | Partido da República | Party of the Republic |
| PRB      | Partido Republicano Brasileiro | Brazilian Republican Party |
| PRN      | Partido da Reconstrução Nacional | National Reconstruction Party |
| PRONA    | Partido de Reedificação da Ordem Nacional | Party of the Reconstruction of the National Order |
| PROS     | Partido Republicano da Ordem Social | Republican Party of the Social Order |
| PRP      | Partido Republicano Progressista | Progressive Republican Party |
| PRS      | Partido das Reformas Sociais | – |
| PRTB     | Partido Renovador Trabalhista Brasileiro | Brazilian Labour Renewal Party |
| PSB      | Partido Socialista Brasileiro | Brazilian Socialist Party |
| PSC      | Partido Social Cristão | Social Christian Party |
| PSD      | Partido Social Democrático | Social Democratic Party |
| PSDB     | Partido Social Democracia Brasileira | Brazilian Social Democracy Party |
| PSDC     | Partido Social Democrata Cristão | Christian Social Democratic Party |
| PSL      | Partido Social Liberal | Social Liberal Party |
| PSOL     | Partido Socialismo e Liberdade | Socialism and Liberty Party |
| PST      | Partido Social Trabalhista | – |
| PSTU     | Partido Socialista dos Trabalhadores Unificado | Unified Workers’ Socialist Party |
| PT       | Partido dos Trabalhadores | Workers’ Party |
| PTB      | Partido Trabalhista Brasileiro | Brazilian Labour Party |
| PTC      | Partido Trabalhista Cristão | Christian Labour Party |
| PTdoB    | Partido Trabalhista do Brasil | Labour Party of Brazil |
| PTN      | Partido Trabalhista Nacional | National Labour Party |
| PTR      | Partido Trabalhista Renovador | Reform Labour Party |
| PV       | Partido Verde | Green Party |
| REDE     | Rede Sustentabilidade | Sustainability Network |
| –        | Republicanos  | Republicans   |
| –        | Solidariedade | Solidarity    |
Table 2 – List of the recent Brazilian presidents.

| President                                      | Took office | Left office | Political party |
|------------------------------------------------|-------------|-------------|-----------------|
| Fernando Collor de Mello (Collor)              | 1990        | December 1992 | PRN             |
| Itamar Franco (Itamar)                         | December 1992 | 1994       | PMDB           |
| Fernando Henrique Cardoso (FHC)                | 1995        | 2002        | PSDB           |
| Luís Inácio Lula da Silva (Lula)               | 2003        | 2010        | PT             |
| Dilma Vana Rousseff (Dilma)                    | 2011        | August 2016 | PT             |
| Michel Miguel Elias Temer Lulia (Temer)        | August 2016 | 2018        | PMDB           |
| Jair Messias Bolsonaro (Bolsonaro)             | 2019        | -           | Independent³   |

- Legal use of public funds;
- Compliance with judicial decisions.

In the recent history of Brazil’s democracy, after 1985 (when the military regime ended), two presidents were impeached. The first was Fernando Collor de Mello, who quit his position during the impeachment process. The second was Dilma Rousseff, she was temporarily removed from office in May 2016, and the consolidation of her impeachment process was in August 2016.

3.2 Our dataset

The dataset was obtained in the web site of the Chamber of Deputies. There are 3753 proposition votes from April 1991 to February 2020. Some of the proposition attributes are:

- The proposition name;
- The proposition voting date;
- The deputies (name and party) and their respective votes;
- The proposition theme.

Figure 3 shows the distribution of propositions over time. In 2020, the lack of data does not allow a proper analysis of the votes, so we will not use this part of the data in future analysis. Table 3 details the percentage of votes of each type by year. The category “obstruction” was not usual in the first records. After 1998 it appears more times but not significantly recurrent when compared to the other types of votes. There is no documentation explaining the meaning of votes registered as “Art. 17” or ‘-’ (the column “Others”), but the most reasonable interpretation for ‘-’ seems to be the deputies’ absence at votations. The “Art. 17” entry seems to be a reference to a specific resolution about the Chamber of Deputies president.

---

³ Source: <https://www2.camara.leg.br/transparencia/dados-abertos/dados-abertos-legislativo>. Accessed on 09/2018.

⁶ Source: <https://www2.camara.leg.br/legin/fed/rescad/1989/resolucaodacamaradosdeputados-17-21-setembro-1989.html>. Accessed on 10/2020.
3.3 Chapter summary

In this chapter, the concepts necessary to contextualize the data used in this research were presented, beginning with an overview of the powers in Brazilian political organization as executive, legislative, and judiciary. The latter is the most distant and independent. The legislative power discusses laws for the entire country, states, or cities. Propositions are documents that must be voted by the lower house, if approved the document proceeds to the upper house, and eventually become a law. The Brazilian political system is multi-party, and some of the most popular (with more deputies elected) political parties are MDB, PT, PP, PSDB, and DEM. The Political science concepts of cohesion and coalition can help understand the political dynamics once they are related to the inter- and intra-organizations of these groups. Finally, the dataset is detailed as a list of propositions with the deputies votes from 1991 until 2020.
Table 3 – Percentage of each type of vote by year (“yes”, “no”, “abstention”, “obstruction” e “others”). The last column indicates the total of votes.

| Year | Propositions | Yes     | No      | Abstention | Obstruction | Others | Entries |
|------|--------------|---------|---------|------------|-------------|--------|---------|
| 1991 | 40           | 60.67   | 37.28   | 2.05       | 0.00        | 0.00   | 12029   |
| 1992 | 42           | 73.89   | 23.99   | 2.12       | 0.00        | 0.00   | 11317   |
| 1993 | 51           | 53.47   | 44.73   | 1.79       | 0.00        | 0.00   | 16410   |
| 1994 | 16           | 66.36   | 31.08   | 2.56       | 0.00        | 0.00   | 3716    |
| 1995 | 115          | 62.43   | 36.03   | 1.54       | 0.00        | 0.00   | 46254   |
| 1996 | 115          | 60.99   | 37.21   | 1.80       | 0.00        | 0.00   | 44670   |
| 1997 | 141          | 52.59   | 45.77   | 1.64       | 0.00        | 0.00   | 53651   |
| 1998 | 93           | 48.92   | 44.05   | 1.35       | 0.21        | 5.47   | 36865   |
| 1999 | 125          | 48.28   | 26.06   | 0.34       | 0.30        | 25.02  | 64081   |
| 2000 | 167          | 45.33   | 29.51   | 0.43       | 0.68        | 24.05  | 85641   |
| 2001 | 98           | 43.72   | 25.52   | 0.42       | 1.39        | 28.95  | 50233   |
| 2002 | 42           | 48.42   | 18.48   | 0.45       | 1.01        | 31.64  | 21536   |
| 2003 | 150          | 35.39   | 29.49   | 0.50       | 3.07        | 31.36  | 76925   |
| 2004 | 118          | 35.46   | 29.54   | 0.50       | 3.08        | 31.42  | 60520   |
| 2005 | 84           | 27.28   | 26.87   | 0.55       | 5.11        | 40.19  | 43079   |
| 2006 | 99           | 34.68   | 22.47   | 0.39       | 2.53        | 39.93  | 50761   |
| 2007 | 217          | 29.57   | 33.97   | 0.45       | 4.75        | 31.26  | 111266  |
| 2008 | 158          | 30.50   | 26.92   | 0.37       | 5.34        | 36.87  | 80999   |
| 2009 | 154          | 34.87   | 26.82   | 0.36       | 2.77        | 35.18  | 78966   |
| 2010 | 83           | 38.43   | 19.78   | 0.42       | 3.04        | 38.33  | 42573   |
| 2011 | 99           | 31.62   | 31.73   | 0.44       | 3.07        | 33.14  | 50762   |
| 2012 | 79           | 34.13   | 27.00   | 0.45       | 2.68        | 35.75  | 40491   |
| 2013 | 160          | 30.59   | 26.54   | 0.54       | 3.25        | 39.07  | 81611   |
| 2014 | 94           | 42.90   | 17.23   | 0.32       | 1.75        | 37.80  | 46901   |
| 2015 | 275          | 46.55   | 31.43   | 0.65       | 1.37        | 20.00  | 140592  |
| 2016 | 204          | 35.00   | 29.40   | 0.36       | 5.82        | 29.42  | 104518  |
| 2017 | 232          | 33.10   | 29.46   | 0.48       | 6.01        | 30.95  | 117941  |
| 2018 | 152          | 5.62    | 29.86   | 0.37       | 9.13        | 35.03  | 77928   |
| 2019 | 332          | 34.54   | 33.94   | 0.24       | 5.11        | 26.17  | 168999  |
This chapter presents an overview of studies focused on legislators’ behavior and political parties’ dynamics, defined as cohesion, coalitions, and government strength. Section 4.1 reviews some Political science methods such as Poole and Rosenthal (1985) used to find legislators’ positions in a Euclidean space according to their voting patterns. Section 4.2 presents Network science studies that use co-voting networks and topological metrics to characterize political party groups (MASO et al., 2014; CHEREPNALKOSKI et al., 2016; LEVORATO; FROTA, 2016; FERREIRA; MATOS; ALMEIRA, 2018). In particular, Brazilian legislators are objects of study in Zucco (2009), Leite and Trento (2018), Levorato and Frota (2016), Ferreira, Matos and Almeira (2018).

### 4.1 Political science

This section reviews some Political science works used to visualize legislators or political parties’ dynamics.

#### 4.1.1 Poole and Rosenthal (1985)

The Nominal Three-step Estimation method, also known as the NOMINATE method, was a landmark in Political science. Since it presented a good solution to represent politicians in a Euclidean space. In Poole and Rosenthal (1985), each legislator has an ideal point which indicates her/his position in the political spectrum represented by a unidimensional space. Each roll call is represented by the points corresponding to the ‘yes’ or ‘no’ voting alternatives (also called the roll call parameters). The utility function estimates the probability of legislator vote in a specific alternative given its affinity with her/his ideal point. When the legislator does not follow her/his expected political position, it occurs an error. The model considers both spatial and non-spatial errors, i.e., errors in the spacial estimation and non expected political positions motivated by unknown external events.
In a perfect roll call, all ideal points of the legislators are known. Also there is a midpoint where all legislators on the left voted on the same option and those on the right voted on the other option. The NOMINATE method tries to estimate these ideal points by first estimating the utility function parameters and then the roll call parameters. While the parameters of a specific step are being calculated, the other values are held fixed. This three-step loop continues until the correlation of the parameters in the previous and the current iterations reach up a value specified by the user.

This method allows analyzing the legislators by their position in a one-dimensional space varying from liberal to conservative. After this first version, some improvements were proposed. For instance, the D-NOMINATE variation allows estimating the ideal point of a legislator over time (POOLE; ROSENTHAL, 1991), and the DW-NOMINATE designed for personal computers. Besides proposing the NOMINATE method, Poole and Rosenthal (1985) also applies it to the United States Congress roll calls (1957 and 1958). Moreover, some authors used the NOMINATE method for Brazilian roll calls. In particular, Zucco (2009) uses this method to interpret the Brazilian political context that is different from the North-American. The first one is a multi-party system, while the second is a two-party system.

Zucco (2009) argues that there is a particular pattern usual in the Brazilian legislative: the president influences the deputies and senators’ behavior on votes by offering nominations in strategical political positions. The author analyzed the Chamber of Deputies’ roll calls grouped by legislatures (from 1987 to 2007). The output positions calculated by the W-NOMINATE method indicate, at first sight, a one-dimensional political spectrum varying from left to right.

Despite this, Zucco (2009) shows that this organization reflects the political parties alignment with the government. An additional analysis using the probability of a legislator takes a position aligned to the government based on her/his position in this political spectrum supports the idea that this space correlates with the legislator’s degree of alignment with the government. The author also discusses the political party’s role as the bridge connecting the president to the group that supports the government.

### 4.1.2 Leite and Trento (2018)

In order to avoid using the NOMINATE methods family, Leite and Trento (2018) use PCA (Principal Component Analysis) to visualize the political party dynamics. The authors propose a matrix of votes where each row represents a roll call, and each column is a legislator’s voting history. The alternatives are 1 for “yes”, -1 for “no” and 0 in the case of “abstention”, “obstruction” or “absence”. Before applying PCA, each element is subtracted by its average row value.

The visualization consists of a radar map with a colored circle for each political party. The members’ positions calculated by PCA are used to find the parties’ centroids. When radar
maps are generated for consecutive legislatures, it is possible to identify the position changes as the legislature changes. However, once the PCA generates different spaces, the authors proposed a rotation function to minimize the different positions of the same parties in consecutive radar maps. Figure 4 shows an example of the visualization proposed. The rotation allows comparing the two radar maps. In the first period (2001-2002), the radar map shows PSDB in the left and PT in the right. Then, in the second legislature (2003-2004), the MDB (at that time, named PMDB) relative position changes from left to right.

The roll calls were obtained from the Chamber of Deputies and Federal Senate (from 2007 to 2014). There are also voting records of the Municipal Assembly of São Paulo (from 2013 to 2016). As a visualization method, this approach enables an analysis of well known political party dynamics in the Brazilian legislative such as the sizeable ideological distance between PT and PSDB, as indicated in Figure 4.

### 4.2 Network science

As an alternative to Political science works, there is co-voting networks modeling. Within this area of research, Maso et al. (2014), Cherepnalkoski et al. (2016) investigate political party cohesion and coalitions. In particular, (Levorato; Frota, 2016; Ferreira; Matos; Almeira, 2018) are important related works since they analyze the Brazilian roll calls in a similar fashion we do in this research.
4.2.1 **Maso et al. (2014)**

Maso et al. (2014) use co-voting networks and topological measures to analyze political polarization, cohesion, government strength. The data used is the Italian Parliament roll calls (between April 2013 and December 2013). The co-voting network connects the deputies by the similarity of their alternatives chosen in votings, the weight connecting two legislators is the total of times they voted in the same option (“yes”, “no” or “abstention”) divided by the total of roll calls.

To detect party cohesion, the authors used two metrics. The first is defined as the internal cluster density

$$d_{int}(C) = \frac{\sum_{ij \in C} w_{ij}}{n_c(n_c - 1) / 2},$$

(4.1)

while the second is the external cluster density

$$d_{ext}(C) = \frac{\sum_{i \in C, j \notin C} w_{ij}}{n_c(n - n_c)},$$

(4.2)

where $n_c$ is the total of vertices in the cluster, $n$ is the total of vertices in the network, and $w_{ij}$ is the weight of the edge between vertices $i$ and $j$. When the internal density is higher than the network average link density, it indicates that the cluster is more internally connected than usually occurs for the entire network. The external density lower than the average is desired since it indicates that there are just a few connections between the cluster and the other ones, less than the usual for the network analyzed. To analyze community structure and legislators, the authors used the $dQ$ metric, which is defined as the difference between the modularity of the original network and the modularity of the network when shifting a legislator from one community to another. For the polarization analysis by swapping two legislators of opposite political parties, the negative $dQ$ was used as a reference value.

The internal and external cluster densities revealed some cohesive political parties, i.e., the members of a cluster following the same ideologies when voting. Conversely, these metrics also revealed some legislators voting aligned to parties that they do not belong to. The authors found out that political parties are not the best way to maximize the network modularity. Alternatively, when using community detection, two large groups were reveled: the government and the opposition. The $dQ$ variations for the opposition indicated non-cohesive groups. They probably changed their coalitions according to the themes being voted. Finally, a temporal analysis showed details in the political parties’ change of alignment over the months.

4.2.2 **Cherepnalkoski et al. (2016)**

Cherepnalkoski et al. (2016) investigate the internal cohesion and coalitions of political parties. For this purpose, the authors compare two approaches, one based on an agreement metric, known as Krippendorff’s Alpha, and the other based on co-voting networks. In the latter method,
4.2. Network science

Exponential Random Graph Models are used to identify relevant factions affecting the edge’s presence in the networks. Additionally, they use tweets to analyze political endorsing on Twitter.

Their first dataset consists of the European Parliament roll calls between October 2014 and February 2016. For each roll call, a co-voting network was generated where the legislators are vertices and the edges weighted by the number of times the legislators chose the same alternative. For the same time interval, the authors collected tweets and retweets of the Parliament members, i.e., their original publications and contents shared from other profiles on Twitter. The retweets were interpreted as indicators of endorsement employed by some legislators to show their support to political allies. From this second dataset, a retweet network was generated by connecting two legislators when one of them retweets the other legislator’s tweet.

The first method uses the co-voting measure of agreement known as Krippendorff’s Alpha. This value is calculated based on the real disagreement between the legislators divided by the disagreement expected by chance and can be used to evaluate the group’s internal cohesion. The authors also extended the Alpha measure by adapting it to analyze the coalition between groups. For the second approach, Exponential Random Graph Models (ERGM) were used to investigate the factors behind the network formation. The analysis developed took into account the number of edges connecting members of the same nationality, edges connecting members of the same party of the same country, the degree of coalition for each pair of political parties, and the number of mutual retweets between the legislators on Twitter.

The main results indicated the maintenance of high levels of cohesion, as already demonstrated by other authors (HIX; NOURY; ROLAND, 2005). Both methods, the agreement measure, and the ERGM, revealed similar results for the groups classified as cohesive and non-cohesive. The roll calls were also analyzed according to their policy areas, in this case, the already known cohesive groups persists. The coalitions revealed the cooperation between groups of different ideological positions. In particular, larger groups tends to collaborate with small groups of different alignments. Finally, the endorsement analysis on Twitter showed a correlation between aligned voting behavior on Parliament and retweet patterns, except for the Economic area.

4.2.3 Levorato and Frota (2016)

Levorato and Frota (2016) use co-voting networks and Correlation Clustering problems (BECKER, 2005) to analyze the dynamics of Brazilian political party groups, such as coalitions, cohesion, and polarization. Correlation Clustering problems are defined as grouping techniques based on the similarity between elements, given a predefined number of clusters and a cost function. In addition to the original version, there is the Symmetric Relaxed Correlation Clustering (SRCC) problem that allows identifying polarization, mediation, and differential popularity in social networks. The authors chose the second approach since it identified more coherent groups than those found using the first definition of the problem.
Annual co-voting networks were generated using roll calls of the Brazilian Chamber of Deputies between 2011 and 2016. Each vertex is a legislator, and there are two ways to calculate the weight of the edge connecting them. For each roll call, when both legislators chose the same voting option, it counts as +1, if they disagreed, it counts as -1. The alternative way ignores when one of the votes is an abstention, i.e., it counts as 0, while the other considers this situation as a half agreement and counts +0.5. Finally, this sum is divided by the total of roll calls.

An initial analysis of the partitions revealed that the political parties usually do not follow their first coalition formed during the period of elections. The temporal investigation suggests significant changes in Dilma Rousseff’s second term when compared to the first one. The government coalition lost the support of many deputies by the end of the first term. Despite her reelection, the newly elected deputies were predominantly composed of other coalitions not aligned with the president. Besides, center parties and opposition came closer when the government lost the coalition.

**4.2.4 Ferreira, Matos and Almeira (2018)**

The main focus of this study is to investigate legislators’ ideological communities changing over the years. For this purpose, Ferreira, Matos and Almeira (2018) use co-voting networks and topological metrics. The dataset consists of the Brazilian and North-American lower houses roll calls between 2003 and 2007.

The co-voting networks generation follows the typical approach of legislators as vertices and edges weighted by the number of coincident votes divided by the total of roll calls. Since Brazil has more alternatives to vote than the US, the authors opted to consider “obstruction” as they do with “yes” and “no” options. They interpreted that this option shows the legislator’s position against the roll call proposal. The authors also proposed the topological measures Partisan discipline and Party discipline. The first is defined as

\[
pdm = \frac{\sum_{i=1}^{n} I(m, pm, i)}{n},
\]

where \( n \) is the total of roll calls, \( I(m, pm, i) \) is 1 when the legislator \( m \) voted on the most popular option choose by her/his political party \( pm \), 0 otherwise. The second measure is the average of the Partisan discipline of all members of \( pm \).

Despite the existence of a large number of political parties in Brazil, the Party discipline found for communities are greater or equal to the Party discipline calculated for the political parties. In North-America, the recognized communities and political parties partitions are similar. In addition, the authors performed an analysis of the party polarization using the legislators’ neighborhood. To this end, given the two legislators connected by an edge, the authors calculated the Jaccard index between the two sets of neighbors. The neighborhood’s overlap values below a threshold defined with tests were removed to increase the network modularity but without producing sparse networks. For the Brazilian co-voting networks, this analysis supported the idea
that there are not polarized communities since the threshold selected was higher than the one selected for the US networks. The legislators’ majority were connected within the communities but are not strongly connected. In contrast, the US legislators removed from the main components after the edge punning were not a significant fraction. In conclusion, the political parties in the US are more cohesive than the ones in Brazil.

4.3 Chapter summary

This chapter presented an overview of the most recent techniques and results for investigations focused on political party groups behaviors using Complex networks. The first section briefly reviews some works from Political science areas focused on calculate politician positions in a Euclidean space (POOLE; ROSENTHAL, 1985) and its adaptations to other objectives such as temporal and group analysis (ZUCCO, 2009). In the Brazilian political context, these methods were useful to revel political parties’ alignments with the president instead of with their ideological ideas. The analysis performed by Leite and Trento (2018) showed some alignment changes in the largest Brazilian political parties. The second part of this chapter focused on complex network methods. Co-voting networks and topological metrics are usually used to investigate party cohesion, government strength, polarization, coalitions, and party discipline. Levorato and Frota (2016), Ferreira, Matos and Almeira (2018) studied the Brazilian Federal Legislative. Their results revealed that Dilma Rousseff’s government had lose coalition in the period preceding her impeachment. In addition, the Brazilian political parties are not a good way to identify cohesive clusters of legislators as community detection seems to be.
Here we introduce the framework used to analyze group dynamics in co-voting networks extracted from voting data. We focus on characterizing the political parties according to their legislators’ behavior voting. Some topological metrics are used to evaluate cohesion and coalitions. Although our application is specific for politics, the methodology proposed in this chapter can be also useful to other social networks that present group dynamics evolving over time – for instance, co-authorship networks and groups of research.

5.1 Framework

The framework described consists of a sequence of procedures that can be followed to conduct a group analysis in co-voting networks. The main steps are:

- Network construction;
- Edge pruning;
- Community detection;
- Characterization metrics for clusters: fragmentation and isolation.

5.1.1 Network construction

The network construction step consists of two parts: defining how the edge weight must be calculated and, since our goal also incorporates temporal aspects, defining the time interval slices. Our co-voting networks are constructed similarly to the instances mentioned in Chapter 4. The deputies are nodes, that are connected by edges when they agreed more than disagreed during the votes. That is, when they chose the same option in a roll call, it counts +1. In the case of different options, it counts -1. Then this sum is divided by the total of roll calls. If the final
value is positive, then an undirected and weighted edge connects the two deputies. The valid options of vote are “yes”, “no”, “abstain” and “obstruction”, i.e., when the deputy is absent, it does not count.

Since the dataset covers an extended period, we slice the roll calls in time intervals to generate temporal networks allowing detailed analysis of the group dynamics changes. We employed two different approaches to generate the networks: yearly and, alternatively, overlapping windows of three months. We generated yearly networks between 1991 and 2019. This approach allows a general overview of the deputies’ behavior patterns. However, when there were some interest events in specific periods, we used overlapped slices of three months to generate the networks. For example, if the first network was generated based on propositions voted in February-March-April, the next one would be March-April-May. This second approach, in particular, can be more appropriate to reveal details.

5.1.2 Edge pruning

We propose using edge pruning techniques to remove undesired weak connections between the deputies, not statistically significant. Instead of using a single threshold, we apply the backbone extraction presented by Serrano, Boguná and Vespignani (2009). As earlier explained in Chapter 2, this method has a parameter $\alpha$ defined as the minimum significance necessary to preserve an edge on the network. Since a unique value is not appropriate to filter edges in all networks, each network has a specific $\alpha$ adjusted to preserve a predefined percentage of the original network size. As a result, the consistency between consecutive networks is maintained. This percentage must be adjusted according to the application context. In particular, our co-voting networks were filtered, preserving 80% of the network’s original size, i.e., 80% of the nodes in the principal component were kept connected.

5.1.3 Community detection

The community structures reveal how the deputies are effectively divided into groups by analyzing their voting behaviors. In this step, we use the Leiden Algorithm (TRAAG; WALTMAN; ECK, 2019) to find these structures, as it is an improved version of the Louvain Algorithm (BLONDEL et al., 2008). In particular, we compare the number of communities with the number of political parties. Also, we compare the absolute number of partitions (communities or political parties) with its corresponding effective number obtained by applying the diversity measure (Equation 2.5) (CHAO; CHIU; JOST, 2016). For example, the true number of political parties is calculated based on the distribution of deputies among these groups.
5.1.4 Characterization metrics for clusters

The political parties are characterized according to their members’ behavior. The distance between two clusters $A$ and $B$ is defined as

$$d(A, B) = \frac{1}{|A \times B|} \sum_{a,b \in A \times B} l(a, b), \quad (5.1)$$

where $l$ is the shortest path length between $a$ and $b$. Since the edge weights are based on the similarity of voting patterns, we transformed these values into distance using

$$\Delta(w_{ij}) = \sqrt{2 \times (1 - w_{ij})}. \quad (5.2)$$

Given the metrics defined above, we are able to define isolation and fragmentation for clusters. Given a cluster $A$, the first metric is defined as

$$I(A) = \sum_{B \neq A} |B| \times d(A, B) / \sum_{B \neq A} |B| \quad (5.3)$$

and the second is

$$F(A) = d(A, A). \quad (5.4)$$

The isolation measures the average distance between a political party and all other parties. In politics, it is a result of the coalitions established between these groups as the fewer connections a party makes, the more isolated it becomes. Finally, the fragmentation is the distance among the members who belong to the same cluster. It indicates how close or united the group members are. In this sense, this concept may go alongside that of cohesion, derived from Political Science.

Here we use metrics derived from the shortest path lengths between the legislators. This aspect differentiates our work from the others mentioned in Chapter 4. Short distances are used as an indicator of similarity between deputies behaviors. Moreover, these metrics allow the generation of temporal series to describe the groups changing over time. When taking into account historical events, they have the potential to improve our understanding of Brazilian political history.

5.2 Implementation details

The dataset and code are available at <https://github.com/carolmb/political>. Python3 was the programming language used during the research, graph representation and algorithms were generated with igraph library. The machine used to run the experiments had the following specs: Ubuntu 18.04.3 LTS, 32GB RAM, Intel(R) Core(TM) i7-3770 CPU @ 3.40GHz.
5.3 Chapter summary

In this chapter, we presented the methodology used to model and analyze the dataset of propositions. We defined a set of procedures summarized as follows. The first step specifies how to construct the networks based on the legislators’ behavior similarity. The second consists of applying an edge pruning algorithm to reduce data noise. Subsequently, we propose the use of diversity metric to compare the political party partition and the actual groups revealed by applying a community detection algorithm. Finally, two metrics are defined to characterize the political parties according to their members’ behaviors: isolation and fragmentation.
Here we present the results obtained applying the framework specified in Section 5.1 to the Brazilian propositions dataset presented in Section 3.2. We discuss the characterization of the co-voting networks, political parties, and communities, taking into account temporal changes and historical events.

### 6.1 Networks characterization

Figures 5, 7, and 6 were generated using only the principal components of the yearly networks. The cumulative number of deputies distributed among the political parties over time is presented in Figure 5. As we can observe, the distribution among the parties has changed. For instance, traditional parties such as MDB and PT decreased their size, while the category “others”, a reference to non traditional groups, increased substantially after 2013. The peaks in 1995 and 2000 are exceptions explained by the join of deputies of different legislatures: the deputies that were leaving office at the end of January, and the newly elected ones that took office in February first. This specific period did not use to have the Congress recess during January.

Before exploring the partition structures such as political parties or communities, we briefly examine the complete structure of the networks. In general, the average distance among deputies (Equation 2.3) increased recently (from 2003 to 2017) when compared to the years between 1994 and 2002, as shown in Figure 6. As stated earlier in Subsection 5.1.4, this metric is inversely proportional to the similarity of deputies behaviors. Taking this into consideration, it seems the deputies are in more disagreement (or less cohesive) than before.

While Figure 6 provides a general view, Figure 7 shows an initial description of the group structures present in these networks (Equation 2.6 applied to the network communities and political parties). Since the communities modularity is higher than political parties modularity, this fact supports the idea that communities are more accurate to group deputies than the parties,
Figure 5 – Number of seats distributed among political parties in the Chamber of Deputies changing over the years.

![Figure 5](image_url)

Source – Prepared by the author.

Figure 6 – Average shortest path lengths over the years. The values ranges approximately between 2 and 5.

![Figure 6](image_url)

Source – Prepared by the author.

as already shown in literature (MASO et al., 2014). Also, the deputies are more connected within the communities. In this case, their voting behavior possibly does not follow their parties’ ideological ideas. Besides this, the curves seem to be usually correlated: the parties’ modularity increases (or decreases) following the communities’ modularity. In particular, between 2016 and 2018, the political parties’ modularity values were quite different from their neighbor’s values. This period coincides with the post-impeachment of Dilma Rousseff.

Figure 8 presents the NMI between communities and political parties evolving over the years (DANON et al., 2005). These values suggest a moderate correlation between the two partitions. This result is in line with the previous one. In particular, higher values occurred between 2012 and 2015, just before the Dilma Rousseff’s impeachment.

As shown in Figures 7 and 8, it seems political party and community partitions are distinct. Figure 9 is in line with the previous results. The comparison between the total of communities ($P_{total}$) and political parties ($c_{total}$) shows that they are, in fact, quite different. The green line indicates the total number of political parties evolving over the years, these
values are considerably higher \((13 \leq p_{total} \leq 29)\) than the total of communities \((3 \leq c_{total} \leq 7)\). We also compare these values with the effective number of political parties and communities (Equation 2.5). The diversity values calculated for the deputies distribution over political parties are always smaller than any \(p_{total}\) (blue and green lines in Figure 9). Finally, the comparison between the effective number of communities and the total of political parties reveals a strong difference between them. As of 2005, the number of political parties began to increase, while the community diversity values remained stable. The highest difference between these values occurs in 2019 with \(D(community) = 3.75\) and \(p_{total} = 29\). The lowest diversity occurred in 2000, \(D(community) = 1.90\) and \(p_{total} = 16\). In other words, despite this large number of political parties, the deputies are organized into just a few groups.

Figure 10 shows an overview of the co-voting networks evolving over the years, each row corresponds to a different president term and the network vertices are colored by political party. In the initial years (1995 to 2001), there is a configuration consisting of two large groups that seem to be the government (PP, DEM, MDB, and PSDB) and the opposition (PT, PP, and MDB). After this period, in 2002, there is a transition to a new configuration. Between 2003 and
Chapter 6. Results and discussion

Figure 9 – Comparison between diversity and size of the communities and political parties partitions.

Source – Prepared by the author.

2011, which coincides with the PT’s government, DEM and PSDB behaved as one unique group distant from the others. In 2012 and the following years, we see progressive transformations where PT seems to lose coalition and becomes more isolated. In fact, MDB and PP always act as a connection between the dominant groups. Since 2012, the “others” parties tend to appear more in the networks, in accordance with Figure 5. The low effective number of communities (Figure 9) is in line with the groups revealed in these networks visualizations.

6.2 Political parties characterization

The following results take into account a small set of political parties composed of those that have more deputies in the parliament. According to our selection criteria, at least in one year, a political party to be shown in the results should have to be part of the top 3 largest parties in terms of the representatives filling chairs in the Chamber of Deputies. Although PSL does not meet the criteria described, the current president Bolsonaro was affiliated to this group when elected, therefore we considered important to include PSL in the results too. In some figures, it is possible to notice that PSL line is not appearing. It happens because it is a small size party and sometimes, after the edge pruning, none of the deputies attached to the principal component of the networks were from this political party.

It is also important to mention that, over the years, some political parties changed their names, and, for simplicity, we verified and unified these names when it was necessary. For instance, recently, PMDB was renamed as MDB. We processed both as a unique group. Another example is PP that also encompasses in past other groups such as PDS, PDC, PPB, PPR, and PDS.

Figure 11 shows the distance between the political parties evolving over the years. In general, it seems to be different behaviors associated with the political parties in the presidency. For example, between 1993 and 2002, during FHC’s term, PT seems to be the only party distant from the other major groups. During this period, PSDB and DEM were more distant from PT than MDB and PP by a moderate difference. These values are possible related to the MDB and
Figure 10 – Visualization of the co-voting networks evolving over time.

Source – (BRITO; SILVA; AMANCIO, 2020)
PP role as “mediators” between groups of different political alignment, as also shown in Figure 10.

In PT’s government, these dynamics changed considerably. MDB and PP became close to PT, this strong coalition seems to last until 2011. Then it gradually began to weaken and culminated with Rousseff’s impeachment. In 2016, PT was again isolated as it used to be by the 1990s and the beginning of the 2000s. It is also noticeable that PSDB and DEM are always close to each other and apart from PT as shown in Figure 10.

In general, it seems PT is a cohesive group, at least when compared to the other political parties. MDB and PP are the flexible ones, that is, they tend to make coalition with the current party in the presidency. But, when they changed their alignment during Rousseff’s term, her party in the presidency was not able to keep its power governance. In conclusion, despite not being the political party in the presidency, they have a strong and influential base. Finally, PSDB and DEM behave almost as a single group. During FHC’s presidency, they were able to interact with other major groups, except for PT. The distance between these two groups remains unchanged in the following years.

In order to detail the dynamics months before, during, and after Rousseff’s impeachment, we used the isolation (Equation 5.3) and fragmentation (Equation 5.4) metrics to characterize the political parties. In this analysis, the co-voting networks used were generated using propositions voted in three-month-windows with overlapping from 2015 to 2019. Figure 12a shows that PT was already isolated in comparison with the other political parties, before the Rousseff’s impeachment process. Moreover, just after Temer took office, the PT isolation increased into the highest value among the major parties. In general, all the other groups presented lower similar values.

The fragmentation values show again PT as the less fragmented party in the sense that this group is the most cohesive among the major groups, as already mentioned before. Despite this, in the gray window (just after Rousseff left office because of her process of impeachment), it presented a peak.

It would be interesting to analyze PSL behavior since it was the president’s political party until 2019. However, PSL used to be a small group, so there are not much data available to make possible this investigation. Despite this, it is possible to examine the beginning of Bolsonaro’s term. It seems PT’s isolation is decreasing while the other parties remain stable. Also, there is a local peak of PT’s fragmentation, but except for this, there is no indication of changes in comparison with the years before.
6.2. Political parties characterization

Figure 11 – Average shortest path length between a political party and all the others. The background color indicates the political party in the presidency in the given period. In particular, the gray area corresponds to the Dilma Rousseff’s process of impeachment.

(a) Average shortest path length between PT and the other political parties.

(b) Average shortest path length between PSDB and the other political parties.

(c) Average shortest path length between PP and the other political parties.

(d) Average shortest path length between MDB and the other political parties.

(e) Average shortest path length between DEM and the other political parties.

Source – Prepared by the author.
Figure 12 – *Isolation* and *fragmentation* of the political parties. The background color indicates the political party in presidency in the given period. In particular, the gray window corresponds to the Dilma Rousseff’s process of impeachment.

(a) Isolation evolving over the years.

(b) Fragmentation evolving over the years.

Source – Prepared by the author.

6.3 Inside the communities

As shown in Figure 9, there is a significant difference between the effective number of political parties and the total of communities found in our co-voting networks. Here we detail how the political parties are distributed among the communities, in particular, focusing in three moments: the beginning of Lula’s first term (Figure 13); the beginning of Rousseff’s first term and her last two years in the presidency; and the ending of Temer’s term (Figures 14 and 15). The parties are listed from left to right according to their political stance (DUCROQUET, 2018). Each community is shown in a different bar chart named with a capital letter with its associated total of deputies informed in parenthesis.

Figure 13 shows two years preceding Lula’s election and the first once he took office. In particular, the effective number of communities displayed its lowest value in 2000 (see Figure 9), that it is not a coincidence that there are only two major communities, one composed of the left-ring parties and other of the center and right-ring as detailed in Figure 13a. This configuration remains almost the same in 2002 (Figure 13b), but, after Lula took office, it changed completely once the center political parties joined the left-ring forming a new coalition. This pattern of behavior, also in accordance with Figure 11, details a tendency of some political parties to seek out political alliances with the current party in the presidency even if it is not in accordance with their ideologies. When the motivations behind these emerging coalitions are specifically to obtain favors such as special important positions or even illegal funds, in
Political science, is the definition for cronyism (KHATRI; TSANG; BEGLEY, 2006). Whereas, in 2003, it was happening the Mensalão scandal, a famous Brazilian case of a vote-buying scheme involving these groups (NEWS, 2013), so it can be a possible explanation for these new coalitions. Although, in the following years after the scandal outbreak, these groups remained the same. Only in 2011, some signs of change had appeared as new communities emerged composed only by center-to-right parties.

Figure 14a shows the first year of Rousseff in the presidency, elected with the support of the former president Lula. By that year, the PT’s coalition was about to dissolve. In the next year, there was only one community predominantly composed of PT’s deputies and the emergence of two center-to-right communities. At the beginning of her second term (2015 and 2016), the left parties were distributed among communities C, D, and E, while the right-to-center parties among A and B. As a conclusion, despite have being reelected, Rousseff did not have a coalition supporting her government. These results are in line with the temporal series showed in Figure 11.

Figure 13 – Political parties distribution among communities (2000, 2002, and 2003).

Figure 15a shows the last year of Temer in the presidency, Rousseff’s vice president and her successor after the impeachment process conclusion. This new distribution seems similar to the one displayed in 2000, however, the right-to-center ring was divided into two communities A and B. On the other hand, the left ring is isolated in the community C. The following year is the beginning of Bolsonaro’s term, the current president of Brazil (Figure 15b). The left ring was spread among three different communities A, B, and C, while the center-to-right ring was concentrated in community D. It is also noticeable the rapid rise of PSL that used to have just a small number of deputies elected until 2018. But this was not a coincidence, PSL was Bolsonaro’s party when he was elected.
Figure 14 – Political parties distribution among communities (2011, 2012, 2015, and 2016).

(a) 2011
(b) 2012
(c) 2015
(d) 2016

Source – Prepared by the author.

6.4 Chapter summary

In this chapter, we presented the main results of analyzing the Chamber of Deputies using the framework proposed in Chapter 5. The analysis is divided into three sections: network characterization, political parties characterization, and community characterization according to the political parties’ distribution among them. In the first section, we examined the networks globally. We showed that the communities’ partition seems more accurate to group the deputies according to their voting patterns instead of the political parties’ partition. When comparing the total of communities and political parties, there is a significant difference that might indicate less plurality of ideas than our large number of parties suggest. In Section 6.2, we used the average distance between groups to analyze the coalitions evolving over the years. In particular, the
beginning of Lula’s first term and the early ending of Rousseff’s second term draw attention since it is noticeable some interesting processes of emergence and rupture of coalitions. The weakness of Rousseff’s second term was visible months before her impeachment. Finally, in Section 6.3, we investigated inside the communities, detailing the political parties’ distribution among them. In line with the second section, these results also show the coalitions between the major parties in the Chamber of Deputies during PT’s government. In addition, we also analyzed the last year of Temer’s term and the first year of Bolsonaro in the presidency.
Political science plays an essential role in analyzing systems of governance and their associated political practices. In particular, for countries governed by democratic regimes (i.e., by elected representatives of the people), it is important to understand how their governmental agents act defending public policies to ensure social and economic development. This critical disposition allows ordinary citizens to claim for their interests to be fully considered and collaborates with the maintenance of democratic institutions.

In this work, we investigated the Brazilian legislators’ behavior based on their voting patterns. Due to its recent democracy, multi-party system, and easy access to voting data, Brazil is an interesting case study. In order to perform this research, the data was modeled using complex networks. It is well known that Political science has been using statistical methods to investigate hidden patterns on voting data of legislators around the world, as illustrated in the related works (Chapter 4). Nevertheless, Network science also studies social networks, including the co-voting networks used in this research. Extensive possibilities are available when applying topological analysis, mathematical models, and visualization techniques.

Our data consists of propositions voted by the Brazilian Chamber of Deputies (the lower house of the Brazilian National Congress) between 1991 and 2019. Our overall goal was to investigate the representatives’ voting patterns taking into account the temporal aspects and focusing on the group dynamics, be them the communities’ structures or political parties. In addition, we were also interested in characterizing these dynamics combining them with political concepts such as party cohesion and coalitions. For this purpose, we proposed a framework, that is, a set of procedures that was followed to deploy these analyses.

Using the framework, we analyzed the co-voting networks composed by deputies linked according to the similarity of their voting patterns. The first part of our analysis is a global characterization of the networks and a comparison of the communities and political parties partitions. At this point, some tendencies, such as less cohesive networks in recent years,
were observed compared with the initial years analyzed. Moreover, the partitions’ comparison indicated communities as more appropriate to cluster the deputies instead of the political parties. Somewhat the idea of a multiparty system with such many different political parties and different ideologies is not real in practice, the small number of effective communities compared with the large total of political parties supports this conclusion.

We calculated the average distance between parties as well as fragmentation and isolation. These metrics, displayed as temporal series, showed some interesting patterns associated with the change of terms. Some parties seem to be more “flexible” than others when choosing their coalitions, such as MDB and PP, while others formed durable coalitions, for instance, DEM and PSDB. Different from the other groups, PT used to be isolated, but during their period on the presidency, this group showed its ability to form a coalition. The analysis of before, during, and after the Rousseff’s impeachment process revealed that though the PT was in the presidency, it did not have a strong coalition anymore as it used to be in Lula’s first and second terms. Finally, the political parties’ distribution among the communities showed the emergence and rupture of coalitions.

All these results are in line with the ones shown by the related works of Political and Network sciences. We believe that metrics based on topological distance and the visualization, plotting temporal series, offer a natural alternative to analyze voting data, allowing global and mesoscopic analysis according to the interests of the researchers.

7.1 Limitations and future works

The current methodology used in this research does not consider the propositions’ attributes of theme or the party authorship. The analysis deployed until this moment did not use this information since it was not available. It would be necessary to aggregate data obtained from other sources, such as the one available in Deputados (2020). After this, it would be interesting to analyze if there are specific coalitions formed according to the propositions’ themes (policy areas).

Another unexploited aspect of the group dynamics is the motivations behind the emergence, union, division, or dissolution of political parties. Up to this point, we analyzed the coalitions emerging and dissolving by analyzing the distribution of the political parties among the networks’ communities. Despite this, several political parties continue going through changes in their internal organization. It could be interesting to analyze if there are signals of rupture or union before these events occur.

Finally, once the proposed methodology was employed and tested only in this Brazilian dataset, it would be interesting to analyze voting data from other countries as a way to validate generalization of the framework for other political circumstances. Alternatively, the analyzes of group positions can also be useful in different contexts unrelated to politics. For example,
research groups in co-authoring networks. In this case, we could explore the distribution of research topics among authors’ communities as well as understand how these groups appear and give rise to new groups.

7.2 Publication

The main contributions of this research were reported in the following publication:

- **Brito, A. C. M., Silva, F. N., & Amancio, D. R. (2020).** A complex network approach to political analysis: Application to the Brazilian Chamber of Deputies. Plos one, 15(3), e0229928.
ALBERT, R.; JEONG, H.; BARABÁSI, A.-L. Diameter of the world-wide web. *nature*, Nature Publishing Group, v. 401, n. 6749, p. 130–131, 1999. Citation on page 24.

AMARAL, O. E. d. O que sabemos sobre a organização dos partidos políticos: uma avaliação de 100 anos de literatura. Universidade Federal do Rio Grande do Sul. Instituto de Filosofia e Ciências . . . , 2013. Citation on page 32.

BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. *science*, American Association for the Advancement of Science, v. 286, n. 5439, p. 509–512, 1999. Citation on page 25.

BECKER, H. A survey of correlation clustering. *Advanced Topics in Computational Learning Theory*, p. 1–10, 2005. Citation on page 43.

BLONDEL, V. D.; GUILLAUME, J.-L.; LAMBIOTTE, R.; LEFEBVRE, E. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, IOP Publishing, v. 2008, n. 10, p. P10008, 2008. Citations on pages 27 and 48.

BONDY, J. A.; MURTY, U. S. R. *et al.* *Graph theory with applications*. [S.l.]: Citeseer, 1976. Citation on page 26.

BRITO, A. C. M.; SILVA, F. N.; AMANCIO, D. R. A complex network approach to political analysis: Application to the brazilian chamber of deputies. *Plos one*, Public Library of Science San Francisco, CA USA, v. 15, n. 3, p. e0229928, 2020. Citations on pages 55 and 59.

CANCHO, R. F. I.; SOLÉ, R. V. The small world of human language. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, The Royal Society, v. 268, n. 1482, p. 2261–2265, 2001. Citation on page 24.

CHAO, A.; CHIU, C.-H.; JOST, L. Phylogenetic diversity measures and their decomposition: a framework based on hill numbers. *Biodiversity Conservation and Phylogenetic Systematics*, Springer, Cham, p. 141, 2016. Citations on pages 26 and 48.

CHEREPNALKOSKI, D.; KARPF, A.; MOZETIČ, I.; GRČAR, M. Cohesion and coalition formation in the european parliament: roll-call votes and twitter activities. *PloS one*, Public Library of Science, v. 11, n. 11, p. e0166586, 2016. Citations on pages 17, 19, 39, 41, and 42.

CLAUSET, A.; NEWMAN, M. E.; MOORE, C. Finding community structure in very large networks. *Physical review E*, APS, v. 70, n. 6, p. 066111, 2004. Citation on page 27.

COSTA, L. d. F.; RODRIGUES, F. A.; TRAVIESO, G.; Villas Boas, P. R. Characterization of complex networks: A survey of measurements. *Advances in physics*, Taylor & Francis, v. 56, n. 1, p. 167–242, 2007. Citations on pages 23 and 24.

DANON, L.; DIAZ-GUILERA, A.; DUCH, J.; ARENAS, A. Comparing community structure identification. *Journal of Statistical Mechanics: Theory and Experiment*, IOP Publishing, v. 2005, n. 09, p. P09008, 2005. Citations on pages 28 and 52.
DEPUTADOS, C. dos. *Pesquisa simplificada*. 2020. Available: <https://www.camara.leg.br/buscaProposicoesWeb/pesquisaSimplificada>. Citation on page 64.

DUCROQUET, S. *Genealogia dos partidos*. 2018. Available: <https://arte.folha.uol.com.br/poder/eleicoes-2018/genealogia-dos-partidos/>. Citation on page 58.

ERD, P. *et al.* On random graphs, i. *Publicationes mathematicae*, v. 6, p. 290–297, 1959. Citation on page 25.

FERREIRA, C. H. G.; MATOS, B. de S.; ALMEIRA, J. M. Analyzing dynamic ideological communities in congressional voting networks. In: SPRINGER. *International Conference on Social Informatics*. [S.I.], 2018. p. 257–273. Citations on pages 17, 39, 41, 44, and 45.

GIRVAN, M.; NEWMAN, M. E. Community structure in social and biological networks. *Proceedings of the national academy of sciences*, National Acad Sciences, v. 99, n. 12, p. 7821–7826, 2002. Citations on pages 24 and 25.

HIX, S.; NOURY, A.; ROLAND, G. Power to the parties: cohesion and competition in the european parliament, 1979–2001. *British Journal of Political Science*, Cambridge University Press, v. 35, n. 2, p. 209–234, 2005. Citation on page 43.

JALILI, M. Social power and opinion formation in complex networks. *Physica A: Statistical mechanics and its applications*, Elsevier, v. 392, n. 4, p. 959–966, 2013. Citation on page 19.

KHATRI, N.; TSANG, E. W.; BEGLEY, T. M. Cronyism: A cross-cultural analysis. *Journal of International Business Studies*, Springer, v. 37, n. 1, p. 61–75, 2006. Citation on page 59.

KUMAR, R.; NOVAK, J.; TOMKINS, A. Structure and evolution of online social networks. In: *Link mining: models, algorithms, and applications*. [S.I.]: Springer, 2010. p. 337–357. Citation on page 19.

LEITE, L.; TRENTO, S. Análise de votações nominais do legislativo brasileiro utilizando componentes principais. *Leviathan (São Paulo)*, n. 12, p. 120–163, fev. 2018. Available: <https://www.revistas.usp.br/leviathan/article/view/143408>. Citations on pages 17, 39, 40, 41, and 45.

LEVORATO, M.; FROTA, Y. Brazilian congress structural balance analysis. *arXiv preprint arXiv:1609.00767*, 2016. Citations on pages 17, 39, 40, 41, and 45.

MASO, C. D.; POMPA, G.; PULIGA, M.; RIOTTA, G.; CHESSA, A. Voting behavior, coalitions and government strength through a complex network analysis. *PloS one*, Public Library of Science, v. 9, n. 12, p. e116046, 2014. Citations on pages 17, 19, 39, 41, 42, and 52.

NEWS, B. *Q&A: Brazil’s ‘big monthly’ corruption trial*. 2013. Available: <https://www.bbc.com/news/world-latin-america-19081519>. Citation on page 59.

POOLE, K. T.; ROSENTHAL, H. A spatial model for legislative roll call analysis. *American Journal of Political Science*, JSTOR, p. 357–384, 1985. Citations on pages 17, 39, 40, and 45.

_____. Patterns of congressional voting. *American journal of political science*, JSTOR, p. 228–278, 1991. Citation on page 40.

SEKARA, V.; STOPCZYNSKI, A.; LEHMANN, S. Fundamental structures of dynamic social networks. *Proceedings of the national academy of sciences*, National Acad Sciences, v. 113, n. 36, p. 9977–9982, 2016. Citation on page 19.
SERRANO, M. Á.; BOGUNÁ, M.; VESPIGNANI, A. Extracting the multiscale backbone of complex weighted networks. *Proceedings of the national academy of sciences*, National Acad Sciences, v. 106, n. 16, p. 6483–6488, 2009. Citations on pages 28 and 48.

TANG, J.; SUN, J.; WANG, C.; YANG, Z. Social influence analysis in large-scale networks. In: ACM. *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, [S.l.], 2009. p. 807–816. Citation on page 19.

TRAAG, V.; WALTMAN, L.; ECK, N. J. van. From louvain to leiden: guaranteeing well-connected communities. *Scientific reports*, Nature Publishing Group, v. 9, 2019. Citations on pages 27 and 48.

TSEBELIS, G. Decision making in political systems: Veto players in presidentialism, parliamentarism, multicameralism and multipartyism. *British journal of political science*, Cambridge University Press, v. 25, n. 3, p. 289–325, 1995. Citation on page 33.

VIANA, M. P.; AMANCIO, D. R.; COSTA, L. d. F. On time-varying collaboration networks. *Journal of Informetrics*, Elsevier, v. 7, n. 2, p. 371–378, 2013. Citation on page 19.

VIANA, N. O que são partidos políticos. *Goiânia: Germinal*, 2003. Citation on page 32.

WAGNER, A.; FELL, D. A. The small world inside large metabolic networks. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, The Royal Society, v. 268, n. 1478, p. 1803–1810, 2001. Citation on page 24.

WANG, B.; TANG, H.; GUO, C.; XIU, Z. Entropy optimization of scale-free networks’ robustness to random failures. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 363, n. 2, p. 591–596, 2006. Citation on page 26.

WATTS, D. J.; STROGATZ, S. H. Collective dynamics of ‘small-world’networks. *nature*, Nature Publishing Group, v. 393, n. 6684, p. 440, 1998. Citation on page 25.

YAN, X.; JEUB, L. G.; FLAMMINI, A.; RADICCHI, F.; FORTUNATO, S. Weight thresholding on complex networks. *Physical Review E*, APS, v. 98, n. 4, p. 042304, 2018. Citation on page 28.

ZUCCO, C. Ideology or what? legislative behavior in multiparty presidential settings. *Journal of Politics*, v. 71, n. 3, p. 1076–1092, 2009. Citations on pages 39, 40, and 45.
