
Relacionamento dos políticos brasileiros: um estudo envolvendo Ciência de Redes e Aprendizado de Máquina.

Tiago Bernardes Kerr



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Uberlândia
2025

Tiago Bernardes Kerr

**Relacionamento dos políticos brasileiros: um
estudo envolvendo Ciência de Redes e
Aprendizado de Máquina.**

Dissertação de mestrado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Área de concentração: Ciência da Computação

Orientador: Bruno Augusto Nassif Travençolo
Coorientador: André Ricardo Backes

Uberlândia
2025

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

K41r
2025

Kerr, Tiago Bernardes, 1995-
Relacionamento dos políticos brasileiros [recurso eletrônico]: um estudo envolvendo ciência de redes e aprendizado de máquinas / Tiago Bernardes Kerr. - 2025.

Orientador: Bruno Augusto Nassif Travençolo.

Coorientador: André Ricardo Backes.

Dissertação (Mestrado) - Universidade Federal de Uberlândia, Programa de Pós-graduação em Ciência da Computação.

Modo de acesso: Internet.

Disponível em: <http://doi.org/10.14393/ufu.di.2025.5233>

Inclui bibliografia.

Inclui ilustrações.

1. Computação. I. Travençolo, Bruno Augusto Nassif, 1981-, (Orient.). II. Backes, André Ricardo, 1981-, (Coorient.). III. Universidade Federal de Uberlândia. Faculdade de Computação. Programa de Pós-graduação em Ciência da Computação. IV. Título.

CDU: 681.3

Rejâne Maria da Silva
Bibliotecária-Documentalista - CRB-6/1925



UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Coordenação do Programa de Pós-Graduação em Computação
Av. João Naves de Ávila, 2121, Bloco 1A, Sala 243 - Bairro Santa Mônica, Uberlândia-MG,
CEP 38400-902

Telefone: (34) 3239-4470 - www.ppgco.facom.ufu.br - cpgfacom@ufu.br



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Ciência da Computação			
Defesa de:	Dissertação, 29/2025, PPGCO			
Data:	03 de Setembro de 2025	Hora de início:	13:04	Hora de encerramento:
Matrícula do Discente:	12312CCP033			
Nome do Discente:	Tiago Bernardes Kerr			
Título do Trabalho:	Relacionamento dos políticos brasileiros: um estudo envolvendo Ciência de Redes e Aprendizado de Máquina			
Área de concentração:	Ciência da Computação			
Linha de pesquisa:	Ciência de Dados			
Projeto de Pesquisa de vinculação:	Processo: 402543/2021-1 Universal CNPq - Estratégias visuais de análise de redes dinâmicas com foco nas características e comportamento dos seus nós			

Reuniu-se por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Ciência da Computação, assim composta: Professores Doutores: André Ricardo Backes - DC/UFSCar, Leandro Nogueira Couto - FACOM/UFU, Francisco Aparecido Rodrigues - ICMC/USP e Bruno Augusto Nassif Travençolo- FACOM/UFU, orientador do candidato.

Os examinadores participaram desde as seguintes localidades: André Ricardo Backes - São Carlos/SP, Francisco Aparecido Rodrigues -São Carlos/SP. Os outros membros da banca e o aluno participaram da cidade de Uberlândia.

Iniciando os trabalhos o presidente da mesa, Prof. Dr. Bruno Augusto Nassif Travençolo, a presentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao discente a palavra para a exposição do seu trabalho. A duração do apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir ao candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato:

Aprovado

Esta defesa faz parte dos requisitos necessários à obtenção do título de Mestre.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação

interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Bruno Augusto Nassif Travençolo, Professor(a) do Magistério Superior**, em 04/09/2025, às 13:51, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **André Ricardo Backes, Usuário Externo**, em 04/09/2025, às 14:06, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Francisco Aparecido Rodrigues, Usuário Externo**, em 04/09/2025, às 14:45, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Leandro Nogueira Couto, Professor(a) do Magistério Superior**, em 08/09/2025, às 11:17, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **6590807** e o código CRC **3A6497A9**.

Referência: Processo nº 23117.055874/2025-59

SEI nº 6590807

Agradecimentos

Agradeço à Universidade Federal de Uberlândia (UFU) e à Faculdade de Computação (FACOM) pelo constante incentivo à pesquisa e pelo espaço proporcionado para o desenvolvimento deste trabalho. A ciência no Brasil precisa — e merece — mais ambientes como esse.

Expresso minha profunda gratidão aos professores Bruno Travençolo e André Backes, que aceitaram me orientar, foram sempre solícitos e compreensivos em todas as etapas do processo. Mesmo diante da mudança de tema no meio do projeto, não mediram esforços para me apoiar e oferecer o suporte necessário.

À minha amada esposa, Natália Cunha, e ao meu querido filho, Hélio Kerr, que chegou ao mundo no final de 2024: meu mais sincero agradecimento por todo o amor, paciência e incentivo durante essa jornada. Vocês são minha base e minha maior motivação. Sem vocês, nada disso teria sido possível.

Agradeço também à minha mãe, Rosana, à minha madrasta, Clélia, à minha tia Terezinha, à minha irmã Flávia e ao meu irmão Wick, por sempre acreditarem em mim e apoiarem minhas escolhas e projetos com tanto carinho.

Ao meu sogro Walter e à minha sogra Rute Lea, por todo apoio e incentivo durante essa jornada. O suporte de vocês foi fundamental para que este projeto se tornasse possível.

À memória dos que já não estão entre nós, mas seguem vivos em meu coração: meu pai, Hélio, e meus avós, Lygia e Wick. Carrego comigo os valores, o amor e o impacto que vocês tiveram em minha vida.

Faço um agradecimento especial ao professor José Fausto, meu primeiro orientador e grande amigo, cuja influência e apoio foram fundamentais para que eu trilhasse esse caminho.

Sou grato também ao Wagner, meu chefe e amigo, pelo incentivo constante à pesquisa e pela compreensão e apoio que me permitiram conciliar as demandas do mestrado com as responsabilidades profissionais.

Também agradeço ao Matheus Henrique dos Santos pelas valiosas dicas que me deu ao longo do mestrado. Sua ajuda foi importante em diversos momentos dessa trajetória.

Por fim, agradeço aos meus amigos que, direta ou indiretamente, contribuíram para a realização desta pesquisa. Cada palavra de incentivo, conversa ou gesto de apoio teve grande valor para mim.

“Sem dados, você é apenas mais uma pessoa com uma opinião”
(Prof. W. Edwards Deming)

Resumo

A política brasileira é marcada por uma elevada fragmentação partidária e por relações muitas vezes complexas entre os parlamentares. Com mais de vinte partidos representados na Câmara dos Deputados em diferentes legislaturas, compreender como ocorrem os alinhamentos e as disputas políticas ao longo do tempo torna-se um desafio relevante para pesquisadores e analistas. Nesse cenário, técnicas de ciência de redes oferecem ferramentas eficazes para representar e analisar essas relações de forma estruturada. Este trabalho tem como objetivo estudar o cenário político brasileiro por meio da análise dos padrões de relacionamento entre deputados federais, utilizando técnicas baseadas em redes complexas. Foram analisados dados de votação de 2005 a 2023, considerando o alinhamento partidário e ideológico dos parlamentares. O estudo buscou, primeiramente, propor uma estrutura de rede complexa capaz de representar as relações políticas no Congresso. A partir dessa rede, foram aplicadas métricas clássicas de redes (grau, centralidade de intermediação e coeficiente de *clustering*), além de realizar análises visuais das redes ao longo do tempo. Outro objetivo central foi comparar diferentes metodologias de agrupamento para detecção de grupos de afinidade: o método de detecção de comunidades Louvain, tradicional na ciência de redes, e o método de agrupamento hierárquico Ward, baseado em aprendizado de máquina. A comparação entre essas abordagens permitiu avaliar suas capacidades de identificar agrupamentos políticos relevantes. Os resultados mostram que, independentemente do governo vigente, há uma tendência de formação de blocos majoritários de apoio ao governo, especialmente no início dos mandatos, com redução da coesão ao longo do tempo. Observou-se também maior rigidez ideológica na esquerda e maior flexibilidade de outros grupos em migrar entre blocos de governo e oposição. Apesar da elevada quantidade de partidos representados no Congresso, o número real de agrupamentos identificados nas redes complexas é bem menor, sugerindo que a fragmentação partidária formal não reflete a estrutura efetiva de alinhamento político. A utilização conjunta das diferentes metodologias permitiu uma visão mais completa e detalhada das dinâmicas políticas brasileiras e pode servir de base para estudos futuros.

Palavras-chave: Redes Complexas; Política Brasileira; Detecção de Comunidades.

Abstract

Brazilian politics is characterized by high party fragmentation and often complex relationships among legislators. With more than twenty parties represented in the Chamber of Deputies across different legislatures, understanding how political alignments and disputes evolve over time presents a significant challenge for researchers and analysts. In this context, network science techniques offer effective tools to represent and analyze these relationships in a structured and systematic manner. This study aims to analyze the Brazilian political scenario by examining the relationships among federal deputies through complex network analysis. Voting data from 2005 to 2023 were analyzed, considering both party affiliation and ideological alignment of the legislators. The first objective was to propose a complex network structure capable of representing political interactions in Congress. Based on this network, classical network metrics (degree, betweenness centrality, and clustering coefficient) were applied, alongside visual analyses over time. We aim to compare a central objective was to compare different grouping methodologies to identify affinity groups: the Louvain community detection algorithm, widely used in network science, and the hierarchical Ward clustering method, based on machine learning techniques. The comparison between these approaches allowed for the evaluation of their ability to capture relevant political groupings. The results show that, regardless of the ruling government, there is a tendency for the formation of majority government support blocks, especially at the beginning of presidential terms, with a gradual reduction in cohesion over time. A stronger ideological consistency was observed on the left, while other groups displayed greater flexibility in shifting between government and opposition blocks. Despite the high number of political parties represented in Congress, the actual number of groupings identified in the networks is much smaller, suggesting that formal party fragmentation does not reflect the true structure of political alignment. The integrated use of different methodologies provided a more comprehensive and detailed understanding of Brazil's political dynamics and offers a foundation for future studies.

Keywords: Complex Networks; Brazilian Politics; Community Detection.

Lista de ilustrações

Figura 2 – Exemplo de aplicação da técnica limpeza de aresta.	25
Figura 3 – Exemplo de aplicação da técnica de identificação de comunidade, onde a cor representa os nós centrais	26
Figura 4 – Exemplo de aplicação da técnica de identificação de centralidade (onde a cor representa os nós centrais da rede).	27
Figura 5 – Etapas da metodologia de análise de redes políticas baseada em padrões de votação, conforme apresentada por (BRITO; SILVA; AMANCIO, 2020).	34
Figura 6 – Fluxograma das etapas do método para avaliação.	36
Figura 7 – Distribuição dos Partidos por Ideologia	39
Figura 8 – Quantidade de Deputados por Ideologia ao longo dos anos	43
Figura 9 – Quantidade de partidos com representação no Congresso ao longo dos anos.	44
Figura 10 – Quantidade de Leis votadas ao longo dos anos.	44
Figura 11 – Proporção da resposta de votação por ideologia ao longo dos anos	45
Figura 12 – Média de votação por ideologia ao longo dos anos	46
Figura 13 – Proporção de faixa de valor do coeficiente kappa.	47
Figura 14 – Visualizações das redes complexas dos deputados brasileiros (2005 a 2008).	49
Figura 15 – Visualizações das redes complexas dos deputados brasileiros (2009 a 2012).	50
Figura 16 – Visualizações das redes complexas dos deputados brasileiros (2013 a 2016).	51
Figura 17 – Visualizações das redes complexas dos deputados brasileiros (2017 a 2020).	52
Figura 18 – Visualizações das redes complexas dos deputados brasileiros (2021 a 2023).	53
Figura 19 – Kappa médio por ideologia ao longo do tempo.	55

Figura 20 – Grau médio por ideologia ao longo do tempo.	56
Figura 21 – Centralidade média por ideologia ao longo do tempo.	57
Figura 22 – Clustering médio por ideologia ao longo do tempo	57
Figura 23 – Número de grupos separado por metodologia ao longo dos anos	58
Figura 24 – Composição das comunidades por ideologia ao longo dos anos	60
Figura 25 – Composição dos <i>clusters</i> por ideologia ao longo dos anos	61

Lista de tabelas

Tabela 1 – Métricas anuais das redes dos parlamentares.	55
Tabela 2 – Proporção de igualdade das metodologias por ano.	59

Lista de siglas

PT Partido dos Trabalhadores

PCdoB Partido Comunista do Brasil

PSOL Partido Socialismo e Liberdade

PPL Partido Pátria Livre

PPS Partido Popular Socialista

PDT Partido Democrático Trabalhista

PSB Partido Socialista Brasileiro

PV Partido Verde

SDD Solidariedade

SD Solidariedade (antiga sigla SDD)

REDE Rede Sustentabilidade

SOLIDARIEDADE Solidariedade

CIDADANIA Cidadania

PMDB Partido do Movimento Democrático Brasileiro

PSDB Partido da Social Democracia Brasileira

PMR Partido Municipalista Renovador

PRB Partido Republicano Brasileiro

PRONA Partido de Reedificação da Ordem Nacional

PMN Partido da Mobilização Nacional

PTdoB Partido Trabalhista do Brasil

PAN Partido dos Aposentados da Nação

PROS Partido Republicano da Ordem Social

S.PART. Sem Partidos (ou legenda avulsa)

PTN Partido Trabalhista Nacional

AVANTE Avante

MDB Movimento Democrático Brasileiro

PTC Partido Trabalhista Cristão

DEM Democratas

PHS Partido Humanista da Solidariedade

PSD Partido Social Democrático

PSDC Partido Social Democrata Cristão

PODE Podemos

DC Democracia Cristã

UNIÃO União Brasil

PFL Partido da Frente Liberal

PP Progressistas

PTB Partido Trabalhista Brasileiro

PL Partido Liberal

PSC Partido Social Cristão

PEN Partido Ecológico Nacional

PR Partido da República

PSL Partido Social Liberal

PRP Partido Republicano Progressista

PRTB Partido Renovador Trabalhista Brasileiro

NOVO Partido Novo

PATRIOTAS Patriota

REPUBLICANOS Republicanos

Sumário

1	INTRODUÇÃO	19
1.1	Motivação	20
1.2	Objetivos e Desafios da Pesquisa	20
1.3	Hipótese	21
1.4	Contribuições	21
1.5	Organização da Dissertação	21
2	FUNDAMENTAÇÃO TEÓRICA	22
2.1	Redes Complexas	22
2.1.1	Limpeza de Arestas	24
2.1.2	Comunidade	25
2.1.3	Centralidade	26
2.1.4	Assortatividade	27
2.2	Kappa de Cohen	28
2.3	Aprendizado e Máquina	28
2.3.1	Análise de Agrupamento	29
2.4	Análise Visual em Redes Complexas	30
2.5	Detecção de Outlier	30
3	TRABALHOS RELACIONADOS	32
4	PROPOSTA	36
4.1	Base de dados	37
4.2	Análise Descritiva	39
4.3	Redes Complexas	39
4.3.1	Construção das Redes Complexas	40
4.3.2	Análise visual	40
4.3.3	Extração das principais métricas	40

4.3.4	Detecção de Grupos	41
5	EXPERIMENTOS E ANÁLISE DOS RESULTADOS	42
5.1	Análise Descritiva	42
5.2	Análise Visual	47
5.3	Análise Clássica	54
5.3.1	Análise Geral	54
5.3.2	Análise por ideologia	55
5.4	Análise de Grupos (Comunidade e Cluster)	58
6	CONCLUSÃO	63
6.1	Principais Contribuições	64
6.2	Trabalhos Futuros	65
6.3	Contribuições em Produção Bibliográfica	65
	REFERÊNCIAS	66

CAPÍTULO 1

Introdução

O sistema político brasileiro é estruturado em três poderes distintos – legislativo, judiciário e executivo – cada um com funções específicas: o legislativo elabora leis, o judiciário as aplica e o executivo as administra. O poder legislativo é constituído por senadores, deputados (f federais e estaduais) e vereadores. Em 2023, o Brasil contava com 513 deputados federais distribuídos em 17 partidos políticos. Esses partidos se dividem entre pró-governo, oposição e independentes, sendo o relacionamento entre eles tema de debates e questionamentos frequentes. Apesar do número expressivo de partidos e deputados, muitas vezes é desafiador compreender como se estabelecem as relações entre os deputados no que diz respeito à proposição e votação das leis.

Estudos como os de (MACIEL; ALARCON; GIMENES, 2018), (RODRIGUES, 2002), (REIS, 2010) abordam essa temática, porém não se utilizam de análises quantitativas, mas sim de análises históricas qualitativas. O estudo dessas relações por meio de abordagens matemáticas pode revelar descobertas significativas. Entre as metodologias de destaque para análise de relacionamentos, destacam-se as redes complexas. Essas redes representam estruturas compostas por elementos interligados, cujas interações formam padrões não triviais e frequentemente apresentam propriedades emergentes. Elas constituem representações matemáticas de sistemas reais, desde redes sociais até redes metabólicas em organismos biológicos. Compostas por nós e arestas, representando entidades e suas relações, as redes complexas são fundamentais para a compreensão de fenômenos complexos (BARABÁSI, 2003). Entende-se que os sistemas complexos podem ser adequadamente modelados por meio dessas redes, tornando o estudo dessas estruturas no cenário político brasileiro não apenas viável, mas também de grande interesse. A identificação de padrões nas votações, proposições de emendas e comportamentos dos parlamentares, disponíveis no banco de dados aberto do governo federal (Governo Federal, 2023), é uma possibilidade real por meio da aplicação de análises baseadas em redes complexas.

Em um estudo conduzido por (BRITO; SILVA; AMANCIO, 2020), foi apresentada uma abordagem similar, investigando os relacionamentos entre parlamentares e seus respectivos partidos, analisando os padrões de votação por meio de redes complexas, evi-

denciando sua eficácia como modelo para esse tipo de análise. No entanto, dentro das redes complexas, existem diversas metodologias que podem trazer descobertas distintas ou ampliar a compreensão de uma descoberta ao serem comparadas.

A proposta deste projeto é apresentar e comparar duas abordagens distintas para detecção de comunidades, também analisar o resultado da rede complexa por meio das métricas da rede e por meio da análise visual via ForceAtlas2 (JACOMY et al., 2014), comparando tanto em nível geral quanto em relação a ideologia dos partidos. Isso aplicado ao contexto político brasileiro com o objetivo de compreender e identificar os cenários políticos de forma mais precisa, representa um valor substancial. Essa análise multifacetada possibilitará a identificação da abordagem mais eficaz para compreender as dinâmicas políticas complexas e suas implicações.

1.1 Motivação

A aplicação de redes complexas na análise de sistemas políticos tem se mostrado uma abordagem poderosa para compreender as interações e dinâmicas entre agentes parlamentares. No contexto brasileiro, diferentes metodologias de análise de redes oferecem perspectivas complementares sobre os padrões de relacionamento entre deputados, seus agrupamentos ideológicos e a evolução dessas interações ao longo do tempo. No entanto, ainda há espaço para investigações que adotem uma abordagem comparativa entre diferentes métodos e que explorem séries temporais mais amplas, permitindo uma visão mais integrada da estrutura política nacional. Neste sentido, o presente trabalho propõe uma análise abrangente que articula múltiplas abordagens metodológicas e examina um período estendido de votações parlamentares, contribuindo para uma compreensão mais ampla e multifacetada das redes políticas no Brasil.

1.2 Objetivos e Desafios da Pesquisa

O objetivo principal deste estudo consiste em analisar os padrões de relacionamento entre os políticos brasileiros, considerando variáveis como votação e ideologia partidária. Os objetivos específicos, são elencados a seguir:

- ❑ Propor uma estrutura de rede complexa que descreva o cenário político brasileiro;
- ❑ Realizar análises dessa rede por meio de métricas de redes complexas e de aprendizado de máquina;
- ❑ Realizar uma comparação entre as duas metodologias propostas;

1.3 Hipótese

Considerando a análise dos padrões de relacionamento entre políticos brasileiros por meio de diferentes metodologias de Ciência de Redes, pressupõe-se que a integração de técnicas de aprendizado de máquina em um modelo de rede complexa fornecerá uma representação mais precisa e abrangente das interações políticas, permitindo uma identificação mais precisa de grupos de afinidade, influência e padrões de votação em comparação com os modelos de exclusivos de Ciência de Redes e os modelos de análise visual.

1.4 Contribuições

As principais contribuições deste trabalho são:

- Criação uma rede complexa que mostra o relacionamento dos parlamentares brasileiros;
- Identificação do melhor método para detecção de comunidade dentro dessa rede complexa;
- Identificação um padrão de análise para se basear em relação a redes complexas.

1.5 Organização da Dissertação

O presente trabalho está estruturado em capítulos que abordam progressivamente os elementos teóricos, metodológicos e analíticos da pesquisa. Inicialmente, no Capítulo 2, **Fundamentação Teórica**, são apresentados os principais conceitos relacionados à ciência de redes complexas, aprendizado de máquina e estatística, considerados essenciais para o entendimento da proposta. Em seguida, o Capítulo 3, de **Trabalhos Relacionados**, reúne estudos que analisam a política brasileira sob diferentes enfoques, oferecendo um panorama da literatura existente. No Capítulo 4 é apresentada a **Proposta**, em que descrevem-se as etapas adotadas na pesquisa, incluindo a coleta de dados, o tratamento inicial e a análise descritiva. Os principais achados são discutidos no Capítulo 5, **Experimentos e Análise dos Resultados**, que contempla o diagnóstico da base de dados, a criação, caracterização e análise das redes complexas, bem como a aplicação de técnicas de agrupamento para identificação de padrões estruturais. Por fim, a **Conclusão**, no Capítulo 6, sintetiza os resultados obtidos, destacando as contribuições do estudo e sugerindo caminhos para pesquisas futuras.

CAPÍTULO 2

Fundamentação Teórica

Este capítulo apresenta os fundamentos teóricos e metodológicos utilizados no desenvolvimento da pesquisa, com ênfase no conceito de redes complexas, abordando desde sua construção — por meio da definição de nós e arestas — até a aplicação de métricas estruturais para análise topológica. Destaca-se o uso do coeficiente kappa de Cohen como medida de similaridade entre pares de elementos, sendo empregado para atribuir pesos às arestas da rede, de modo a refletir o grau de concordância nas interações analisadas. Além disso, são introduzidos conceitos de aprendizado de máquina, com foco em técnicas de análise de agrupamento (*clustering*) para identificação de padrões estruturais. Por fim, são discutidas estratégias de análise visual e critérios estatísticos para a detecção de outliers, com destaque para o método baseado no intervalo interquartil (IQR), utilizado na construção do box plot.

2.1 Redes Complexas

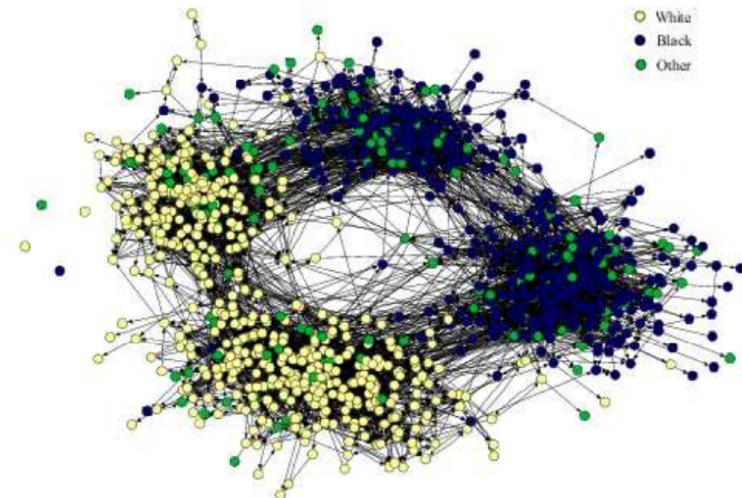
As redes complexas são estruturas matemáticas comumente modeladas por grafos $G = (V, E)$, nos quais V representa o conjunto de nós (ou vértices) e $E \subseteq V \times V$ representa o conjunto de arestas (ou conexões) entre pares de nós. Essas redes são utilizadas para representar sistemas reais compostos por múltiplos elementos interconectados de forma não trivial, frequentemente com padrões heterogêneos, dinâmicos e escaláveis (??).

Essas estruturas são aplicáveis a diversos domínios. Por exemplo, redes sociais como o Facebook ou o Twitter podem ser representadas por grafos em que cada nó $v \in V$ representa um usuário, e cada aresta $(u, v) \in E$ representa uma relação social ou uma interação entre usuários. A Internet pode ser vista como uma rede complexa em que os nós são roteadores ou servidores e as arestas representam conexões físicas ou lógicas entre esses dispositivos. Em biologia, redes de interação proteica representam proteínas como nós e interações bioquímicas como arestas (BARABÁSI, 2016).

A análise de redes complexas permite investigar propriedades estruturais, como a distribuição de grau dos nós, coeficiente de agrupamento (*clustering*), centralidade, mo-

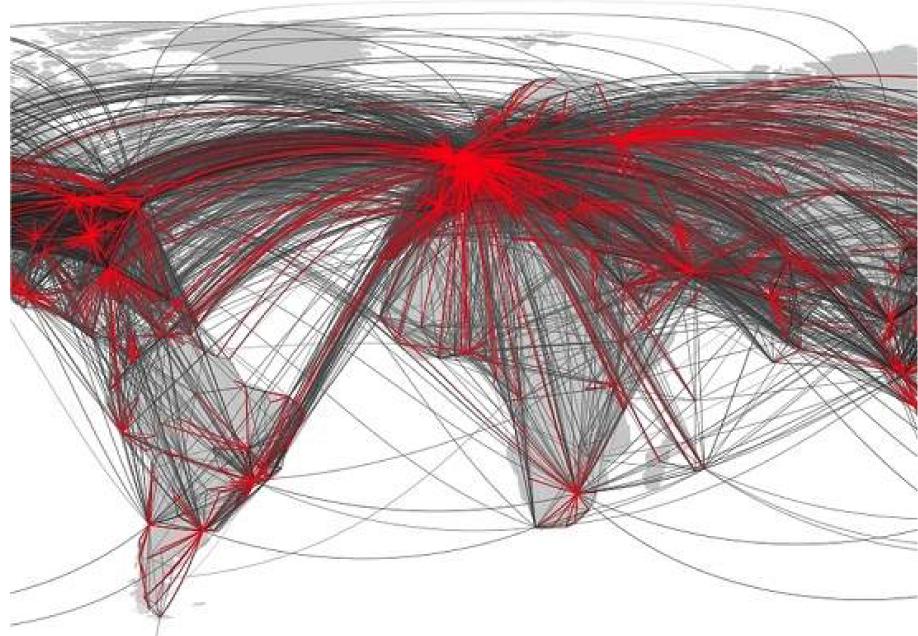
dularidade e caminhos mínimos, bem como estudar dinâmicas sobre a rede, como difusão de informações, robustez a falhas e propagação de epidemias. Assim, essas redes oferecem uma estrutura poderosa para a compreensão de padrões emergentes e da organização de sistemas complexos em diferentes contextos científicos e tecnológicos (COHEN; HAVLIN, 2010; ??; ??).

Existem diferentes tipos de redes complexas, como os grafos aleatórios de Erdős–Rényi, que conectam nós de forma aleatória (ERDÖS; RÉNYI, 1959); o modelo de Barabási–Albert, que gera redes com crescimento preferencial, em que novos nós tendem a se ligar a nós já bem conectados (BARABÁSI; ALBERT, 1999); e redes que seguem uma lei de potência, típicas de sistemas reais como a internet ou redes sociais, em que poucos nós têm muitas conexões e muitos nós têm poucas conexões (NEWMAN, 2005). Os grafos aleatórios são modelos teóricos de redes complexas nos quais as conexões entre os nós são estabelecidas de maneira probabilística, sem preferência ou tendência por determinados nós (ERDÖS; RÉNYI, 1959; NEWMAN, 2010). Esses grafos são caracterizados por uma distribuição de grau dos nós que segue uma distribuição de Poisson, sendo que a maioria dos nós tem um número de conexões semelhante, resultando em uma estrutura homogênea. Por outro lado, o Modelo de Barabási-Albert descreve uma classe de redes complexas baseadas no crescimento preferencial, em que novos nós são adicionados à rede e conectados preferencialmente a nós existentes com alta conectividade, resultando em uma distribuição de grau seguindo a Lei de Potência, no qual poucos nós têm muitas conexões, enquanto a maioria tem poucas. A Lei de Potência é uma característica fundamental em redes complexas, indicando que a distribuição de conectividade dos nós segue uma relação de potência, em que a probabilidade de um nó ter k conexões é proporcional a k elevado a um expoente negativo, sugerindo a presença de hubs ou nós altamente conectados na rede. Além desses, outros tipos de redes complexas incluem redes hierárquicas, redes modulares, redes livre de escala, entre outros, cada uma com características estruturais específicas que as distinguem e que podem ser aplicadas em diversos contextos científicos, sociais, tecnológicos e biológicos para compreender a organização e dinâmica dos sistemas complexos. Exemplos desse relacionamento podem ser vistos nas Figuras 1(a) e 1(b), que retratam o relacionamento de estudantes e aeroportos, respectivamente.



(a) Relacionamento de estudantes estadunidenses pela raça.

(Fonte: (NEWMAN, 2003b))



(b) Relacionamento entre aeroportos.

(Fonte: (GRADY; THIEMANN; BROCKMANN, 2012))

2.1.1 Limpeza de Areostas

A limpeza de areostas fracas em uma rede complexa é um procedimento que envolve a remoção ou filtragem de conexões de baixa intensidade ou peso na estrutura da rede. Essa abordagem tem como objetivo reduzir o ruído e focar nas relações mais significativas ou fortes entre os nós, visando simplificar a representação da rede. A eliminação das areostas mais fracas pode ajudar a destacar padrões mais relevantes e identificar conexões mais

importantes dentro da rede, resultando em uma representação mais clara e significativa das relações entre os nós. Essa prática de limpeza de arestas fracas pode ser aplicada em diversos contextos, como análise de redes sociais, biológicas ou de infraestrutura, em que a identificação de interações mais robustas e relevantes é crucial para compreender a estrutura e dinâmica da rede complexa. Exemplos de algoritmos são: algoritmos de poda (QUINLAN, 1987), reduções de dimensionalidade (PCA) (PEARSON, 1901), entre outros. Um exemplo de aplicação com dados aleatórios pode ser visto na Figura 2.

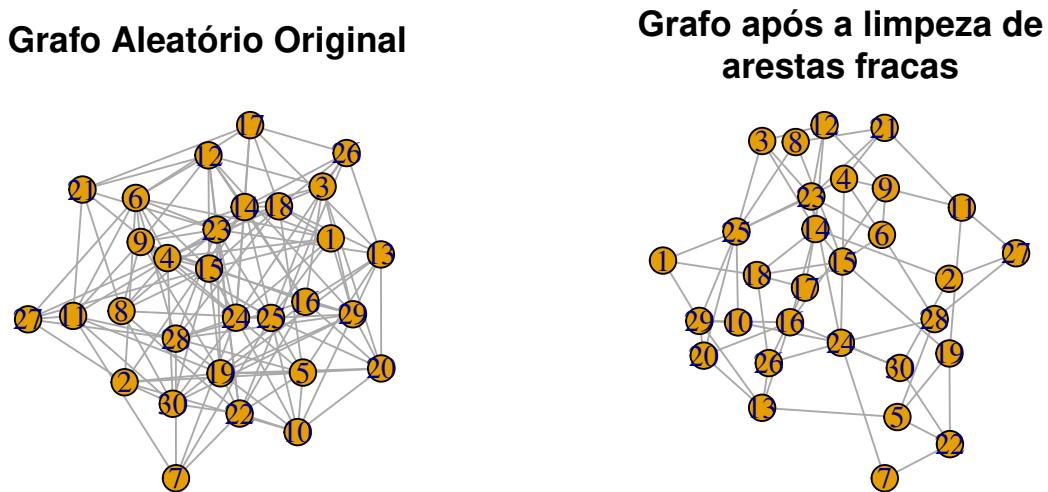


Figura 2 – Exemplo de aplicação da técnica limpeza de aresta.

Fonte: Autoria Própria.

2.1.2 Comunidade

A detecção de comunidade em redes complexas refere-se ao processo de identificar estruturas organizacionais significativas dentro de um sistema de interconexão de elementos, como uma rede social, biológica ou tecnológica. Essas estruturas, conhecidas como comunidades, são conjuntos densamente conectados de nós (ou elementos) dentro da rede, em que os membros compartilham características ou têm interações mais frequentes entre si do que com indivíduos fora desses conjuntos. Métodos de detecção em comunidade utilizam algoritmos e técnicas analíticas para revelar essas estruturas, possibilitando uma compreensão mais profunda das dinâmicas e funcionamento da rede (LANCICHINETTI; FORTUNATO; RADICCHI, 2008).

Exemplo de algoritmos são comumente usados em Ciência de Redes: Algoritmo de Louvain (BLONDEL et al., 2008), Algoritmo de Girvan-Newman (GIRVAN; NEWMAN, 2002). Um exemplo de aplicação com dados aleatórios pode ser visto na Figura 3.

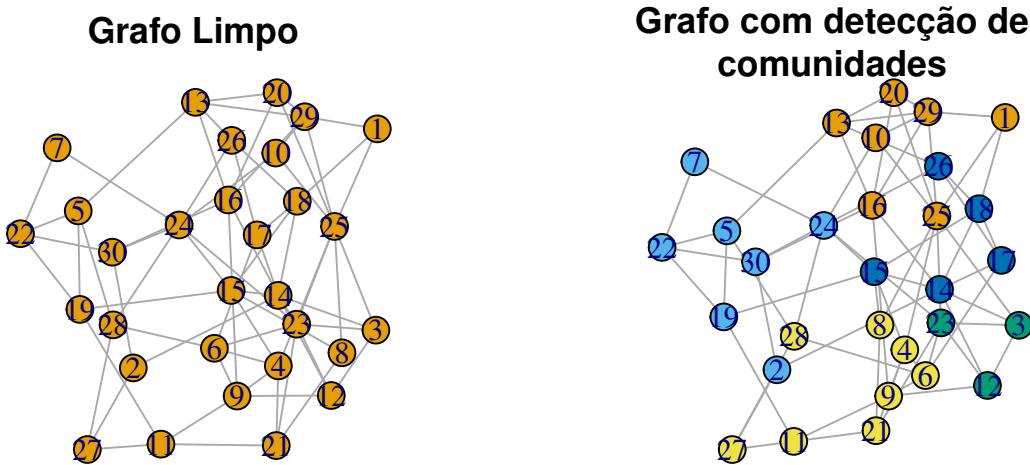


Figura 3 – Exemplo de aplicação da técnica de identificação de comunidade, onde a cor representa os nós centrais

Fonte: Autoria Própria.

2.1.2.1 Louvain

O algoritmo de Louvain é uma técnica heurística amplamente utilizada para detecção de comunidades em redes complexas. Ele tem como objetivo maximizar a modularidade da partição da rede, uma medida que quantifica a densidade de conexões dentro de comunidades em comparação com as conexões entre diferentes comunidades. A modularidade Q é definida por:

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

em que A_{ij} representa o peso da aresta entre os nós i e j , k_i e k_j são os graus dos nós i e j , m é o número total de arestas da rede, c_i e c_j são as comunidades às quais os nós pertencem, e δ é a função delta de Kronecker, que vale 1 quando $c_i = c_j$ e 0 caso contrário (BLONDEL et al., 2008). O algoritmo opera em duas fases iterativas: na primeira, cada nó é atribuído à comunidade do vizinho que proporciona o maior ganho local de modularidade; na segunda, as comunidades identificadas são agregadas para formar uma nova rede, cujos nós representam as comunidades. Esse processo é repetido até que não haja mais aumento significativo na modularidade, permitindo identificar estruturas comunitárias de forma eficiente mesmo em redes de grande escala.

2.1.3 Centralidade

A centralidade em redes complexas refere-se à medida ou métrica que avalia a importância ou influência dos nós dentro de uma rede com base em sua posição topológica. Essas medidas são fundamentais para identificar os nós mais significativos ou centrais em uma rede, desempenhando um papel crucial na compreensão da estrutura, dinâmica e

funcionamento dos sistemas complexos. Existem várias formas de centralidade, cada uma capturando diferentes aspectos da importância dos nós na rede. A centralidade de grau avalia a importância de um nó com base no número de conexões diretas que ele possui na rede. A centralidade de intermediação (*betweenness*) mede a frequência com que um nó atua como ponte ao longo do caminho mais curto entre outros pares de nós na rede. Já a centralidade de proximidade considera a distância média de um nó para todos os outros nós na rede, essas métricas são cruciais para identificar nós-chave em redes complexas, sendo aplicadas em diversas áreas, como ciência de redes, sociologia, biologia e computação (FREEMAN et al., 2002). Exemplos de algoritmos comumente usados em Ciência de Redes: Detecção de Nós de Maior Grau (FREEMAN et al., 2002), Detecção de Nós com Alta Intermediação (FREEMAN, 1977). Um exemplo de aplicação com dados aleatórios pode ser visto na Figura 4.

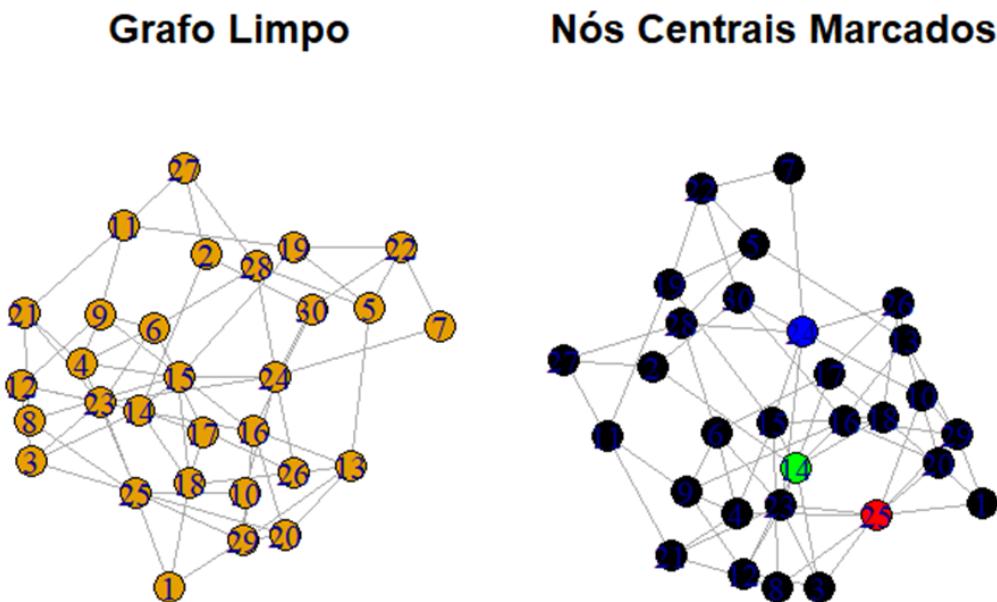


Figura 4 – Exemplo de aplicação da técnica de identificação de centralidade (onde a cor representa os nós centrais da rede).

Fonte: Autoria Própria.

2.1.4 Assortatividade

A assortatividade é uma propriedade fundamental de redes complexas que quantifica a tendência de nós com características similares se conectarem entre si. No contexto de redes, a assortatividade de grau mede especificamente se nós com graus semelhantes tendem a formar conexões preferenciais (NEWMAN, 2002). Esta medida é expressa pelo coeficiente de assortatividade r , que varia entre -1 e $+1$, onde valores positivos indicam que nós de alto grau tendem a se conectar com outros nós de alto grau (rede assortativa),

valores negativos indicam que nós de alto grau preferem se conectar com nós de baixo grau (rede dissolutiva), e valores próximos a zero sugerem conexões aleatórias sem preferência por grau (NEWMAN, 2003a).

A análise da assortatividade fornece *insights* importantes sobre a estrutura e dinâmica de redes reais. Redes sociais tipicamente exibem assortatividade positiva, refletindo a tendência de indivíduos bem conectados interagirem entre si, fenômeno conhecido como homofilia (MCPPERSON; SMITH-LOVIN; COOK, 2001). Por outro lado, redes tecnológicas e biológicas costumam apresentar assortatividade negativa, na qual hubs conectam-se preferencialmente a nós periféricos. Essa estrutura aumenta a robustez a falhas aleatórias, embora reduza o limiar epidêmico e amplifique a vulnerabilidade à disseminação de processos contagiosos (PASTOR-SATORRAS; VESPIGNANI, 2001).

2.2 Kappa de Cohen

O coeficiente kappa de Cohen (coeficiente kappa) é uma medida estatística de concordância usada para avaliar a concordância entre dois ou mais observadores em relação a categorias mutuamente exclusivas. É calculado como a proporção da concordância observada menos a concordância esperada, dividida por 1 menos a concordância esperada (COHEN, 1960):

$$\kappa = \frac{P_o - P_e}{1 - P_e} \quad (1)$$

κ é o coeficiente kappa de Cohen;

Onde: P_o é a proporção de concordância observada;

P_e é a proporção de concordância esperada.

O coeficiente kappa, conforme definido por (COHEN, 1960), assume valores que variam de menores que zero até 1. Sua interpretação segue os seguintes critérios: valores menores que zero indicam ausência de concordância; entre 0,01 e 0,20, concordância desprezível a leve; de 0,21 a 0,40, concordância razoável; de 0,41 a 0,60, concordância moderada; de 0,61 a 0,80, concordância substancial; e entre 0,81 e 1, concordância quase perfeita.

2.3 Aprendizado e Máquina

O aprendizado de máquina (*machine learning*) é um subcampo da inteligência artificial que se concentra no desenvolvimento de algoritmos capazes de aprender padrões a partir de dados, com o objetivo de realizar previsões ou tomar decisões sem serem explicitamente programados para isso (MITCHELL, 1997). Os métodos de aprendizado de máquina são geralmente classificados em duas grandes categorias: aprendizado supervisionado e aprendizado não supervisionado. No aprendizado supervisionado, os algoritmos

são treinados com base em um conjunto de dados rotulado, isto é, contendo pares de entrada e saída, com o objetivo de prever rótulos para novas entradas (BISHOP, 2006). Em contrapartida, no aprendizado não supervisionado, os dados não possuem rótulos e o foco está na descoberta de estruturas subjacentes, como padrões, associações ou agrupamentos, dentro dos dados (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). Esta última abordagem é particularmente útil quando se deseja explorar a organização interna dos dados ou reduzir sua dimensionalidade para facilitar a interpretação.

2.3.1 Análise de Agrupamento

Dentro do escopo do aprendizado não supervisionado, a análise de agrupamento (*clustering*) é uma das abordagens mais utilizadas para explorar a estrutura oculta dos dados. Ao contrário de métodos supervisionados, em que há uma variável-alvo previamente definida, o agrupamento busca identificar automaticamente agrupamentos naturais presentes nos dados com base em suas características (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). As análises de agrupamento consistem em um conjunto de técnicas exploratórias que, com base na distância entre os objetos estudados, identificam padrões de similaridade, agrupando objetos semelhantes em um mesmo grupo. O objetivo é que os elementos dentro de um grupo sejam mais semelhantes entre si do que em relação aos elementos de outros grupos, garantindo a heterogeneidade entre os grupos formados (FÁVERO; BELFIORE, 2017). Os algoritmos de agrupamento podem ser divididos em hierárquicos e não hierárquicos. Nos métodos hierárquicos, a quantidade de grupos pode ser definida a posteriori por meio da análise do dendrograma, embora demandem maior custo computacional. Já os métodos não hierárquicos exigem a definição prévia do número de grupos. No presente trabalho, optou-se pela utilização do algoritmo hierárquico de Ward. O principal fator que motivou a escolha do método Ward foi o fato de se tratar de um método hierárquico, que dispensa a definição prévia do número de grupos, permitindo uma exploração mais flexível da estrutura dos dados. Entre os algoritmos de agrupamento frequentemente empregados em tarefas de aprendizado de máquina destacam-se também o K-Means (MACQUEEN, 1967) e o DBSCAN (ESTER et al., 1996), cada um com suas especificidades e aplicações conforme a natureza dos dados analisados.

2.3.1.1 Ward

O método de Ward é uma técnica de agrupamento hierárquico aglomerativo que visa minimizar a variância dentro dos grupos a cada iteração de fusão. A lógica central do método consiste em escolher, a cada passo, a fusão entre pares de grupos que resulte no menor aumento da soma total dos quadrados das diferenças dentro dos grupos (*Within-Cluster Sum of Squares – WCSS*). Formalmente, seja E a soma total dos quadrados dentro

dos grupos, o critério de Ward minimiza o incremento ΔE na função objetivo, definido como:

$$\Delta E = \frac{n_A n_B}{n_A + n_B} \|\bar{x}_A - \bar{x}_B\|^2$$

em que n_A e n_B são os tamanhos dos grupos A e B , respectivamente, e \bar{x}_A e \bar{x}_B são seus respectivos centróides. A distância entre dois grupos é, portanto, calculada com base no aumento da variância total que ocorreria caso fossem fundidos. Esse critério torna o método de Ward particularmente eficaz na formação de grupos homogêneos e bem definidos, com aplicação relevante em contextos onde se busca a maximização da homogeneidade interna dos agrupamentos (ANDERBERG, 1973).

2.4 Análise Visual em Redes Complexas

(LINHARES et al., 2020) definem visualização de redes complexas como o processo de representar graficamente estruturas e processos em redes que evoluem ao longo do tempo. Essas redes contêm nós (elementos) e arestas (conexões) que mudam dinamicamente. O objetivo é utilizar técnicas de visualização para facilitar a compreensão de padrões e anomalias na estrutura e na atividade das redes ao longo do tempo, destacando conexões recorrentes e atividades dos nós.

A análise visual de redes pode ser realizada por meio de diversos softwares que incorporam algoritmos de *layout* por padrão. Atualmente, um dos algoritmos mais vantajosos para redes de grande porte é o ForceAtlas2. Trata-se de uma técnica de visualização baseada em simulações físicas, cujo objetivo é posicionar os nós de maneira a evidenciar estruturas latentes e agrupamentos dentro da rede. O algoritmo simula forças de repulsão entre todos os nós (semelhante a cargas elétricas) e forças de atração entre nós conectados (semelhantes a molas), até atingir um ponto de equilíbrio que proporciona uma representação espacial mais interpretável. Dentre suas vantagens, destacam-se a capacidade de revelar comunidades, centralidade e padrões de conectividade, sendo especialmente útil em redes sociais e políticas. Além disso, o ForceAtlas2 incorpora otimizações de desempenho, como o uso da aproximação de repulsão via método de Barnes-Hut, o que o torna aplicável a redes extensas, mantendo a qualidade visual e informativa da estrutura representada. Mais detalhes sobre o algoritmo podem ser encontrados em (JACOMY et al., 2014).

2.5 Detecção de Outlier

A detecção de outliers univariados é uma etapa fundamental na análise exploratória de dados. Uma das técnicas mais utilizadas para essa detecção é o método gráfico do

box plot, proposto por (TUKEY et al., 1977), que baseia-se na distribuição interquartil dos dados. Nesse método, define-se o intervalo interquartil (IQR) como a diferença entre o terceiro quartil (Q3) e o primeiro quartil (Q1). Consideram-se como outliers os valores que se situam abaixo de $Q1 - 1,5 * IQR$ ou acima de $Q3 + 1,5 * IQR$. Essa abordagem é amplamente adotada por sua simplicidade e eficácia na identificação de observações discrepantes, especialmente em distribuições assimétricas ou sem pressuposto de normalidade.

CAPÍTULO 3

Trabalhos Relacionados

Este capítulo apresenta uma revisão de estudos que abordam a organização e o comportamento político no Brasil, com ênfase em abordagens metodológicas distintas aplicadas à análise do sistema partidário e da atuação parlamentar. São discutidos trabalhos que exploram desde perspectivas históricas e ideológicas até análises empíricas baseadas em dados de votações legislativas. Em especial, destaca-se a aplicação da ciência de redes como ferramenta analítica para compreender padrões de interação entre deputados, identificar coalizões e avaliar a coerência ideológica das práticas legislativas. Essa diversidade metodológica contribui para uma compreensão mais ampla das estruturas e dinâmicas do sistema político brasileiro.

A forma de organização da política brasileira é tema de estudos de várias áreas, diferentes metodologias tentam compreender e explicar como os políticos se organizam e atuam, se a configuração partidária realmente tem um impacto na tomada de atitudes, ou se resumem a poucos grupos independente do partido. Essas configurações podem ser respondidas por estudos históricos, filosóficos, metanálise ou mesmo por redes complexas. (BRAGA; Pimentel Jr, 2011) aborda o comportamento eleitoral no Brasil (2002 a 2010) num estudo descritivo, identificando o papel dos partidos políticos nesse comportamento, eles identificaram que a simpatia partidária, especialmente pelo PT e PSDB, desempenha um papel crucial na decisão de voto, superando outras variáveis tradicionalmente consideradas relevantes, como perfil socioeconômico e ideologia. Essa simpatia tem se fortalecido ao longo do tempo, polarizando o eleitorado e moldando as escolhas eleitorais. Mesmo entre os eleitores menos escolarizados, a influência dos partidos na decisão do voto é evidente, indicando que, embora não plenamente internalizados, os partidos têm impacto considerável no comportamento dos eleitores brasileiros, estruturando, assim, a dinâmica do sistema político-partidário contemporâneo.

(TAROUCO; MADEIRA, 2015) buscaram avaliar o posicionamento ideológico dos partidos políticos no Brasil contemporâneo, desafiando a percepção comum de que a maioria dos partidos carece de uma ideologia clara. Por meio de uma pesquisa com especialistas, o estudo discute a aplicação dessa metodologia e analisa seus resultados. Essa aborda-

gem, apesar de suas limitações, destaca-se como uma maneira importante de estimar as posições políticas dos partidos. Os especialistas, ao classificarem os partidos em uma escala de esquerda a direita, revelaram uma polarização percebida no sistema partidário brasileiro. Embora alguns partidos tenham sido menos conhecidos ou difíceis de classificar, a análise demonstrou uma notável concordância entre as avaliações dos especialistas, conferindo precisão e confiabilidade à mensuração da ideologia partidária. Comparando esses resultados com outros estudos, conclui-se que há consistência na identificação das posições ideológicas dos partidos brasileiros, apesar das diferentes metodologias. Esta coerência permite afirmar que, apesar das complexidades do sistema político brasileiro, há uma clara ordenação ideológica dos partidos, o que é crucial para compreender seu comportamento e suas interações na democracia recente do país.

(AMARAL, 2013) destacam a tendência de cartelização dos sistemas políticos, o que pode ter consequências perigosas para a democracia, abrindo espaço para lideranças populistas e movimentos antipartidos. Por outro lado, outros estudiosos argumentam que os partidos políticos ainda desempenham funções essenciais para a conexão entre a sociedade e o governo, mesmo diante de novos desafios como a ascensão de lideranças carismáticas antipartidárias e práticas de democracia direta. No contexto brasileiro, embora haja algumas monografias de qualidade sobre alguns partidos políticos, como o PT e o PSDB, e alguns estudos comparativos, ainda existem lacunas significativas no estudo da organização partidária. Faltam informações detalhadas sobre como essas agremiações selecionam líderes e candidatos, administram suas finanças, tomam decisões e se relacionam com movimentos sociais e outras organizações da sociedade civil. Para preencher essas lacunas, os autores sugeriram a realização de pesquisas empíricas sólidas, combinando análises detalhadas dos principais partidos políticos brasileiros com estudos sobre os filiados aos partidos, suas opiniões, níveis de participação e redes de associação.

Um dos trabalhos de referência na aplicação de redes complexas à análise política é o estudo de (BRITO; SILVA; AMANCIO, 2020), que propõe uma metodologia baseada em redes para investigar a dinâmica política da Câmara dos Deputados brasileira. Nesse estudo, os autores constroem redes em que os nós representam deputados e as arestas indicam semelhança nos padrões de votação. A partir de dados abertos compreendendo o período de 1991 a 2019, as redes foram analisadas ao longo do tempo utilizando métricas como modularidade, diversidade efetiva, distância média entre nós e medidas de isolamento e fragmentação partidária.

Apesar do sistema multipartidário brasileiro, os resultados indicam que o número efetivo de comunidades topológicas detectadas nas redes é significativamente menor que o número formal de partidos, girando em torno de três comunidades. Esse achado sugere que a pluralidade ideológica esperada não se reflete de forma clara nas práticas legislativas. Além disso, o estudo identificou padrões de coalizões entre partidos ao longo do tempo e evidências de isolamento partidário precedendo eventos políticos críticos, como

o processo de impeachment presidencial. Os autores também propõem uma interessante representação visual do processo metodológico (Figura 5), que sintetiza desde a coleta e comparação dos votos até a construção, filtragem e análise da rede — imagem que será utilizada nesta dissertação para ilustrar a abordagem de redes complexas aplicada à política.

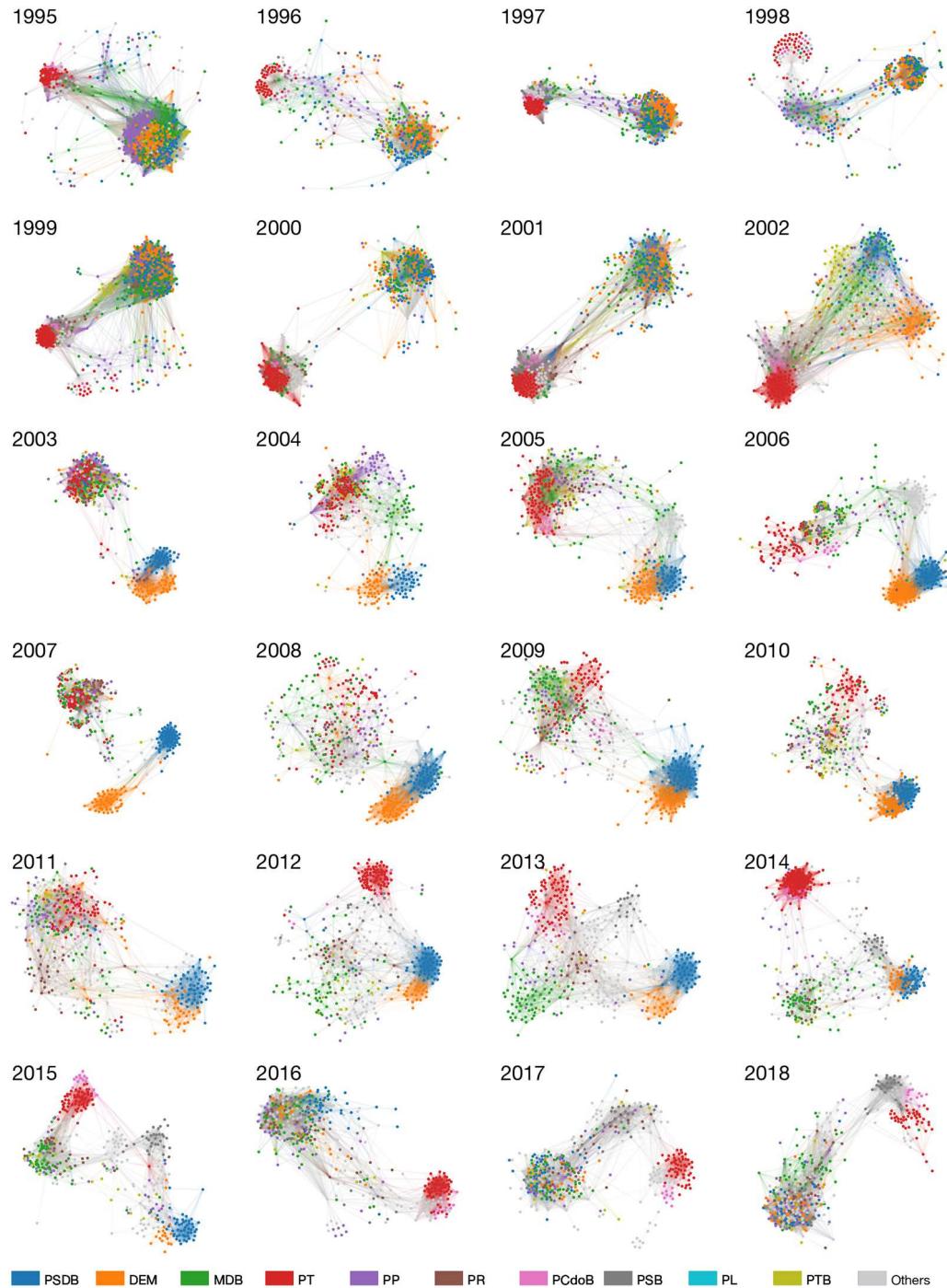


Figura 5 – Etapas da metodologia de análise de redes políticas baseada em padrões de votação, conforme apresentada por (BRITO; SILVA; AMANCIO, 2020).

A pesquisa de (BRITO; SILVA; AMANCIO, 2020) destaca-se por demonstrar como conceitos da ciência de redes podem ser utilizados para mensurar e visualizar aspectos

políticos relevantes como coesão, polarização e alinhamentos partidários, oferecendo um instrumental robusto que pode ser replicado em outros contextos parlamentares.

CAPÍTULO 4

Proposta

A fim de estudar o cenário político brasileiro por meio da análise dos padrões de relacionamento entre deputados federais, foi desenvolvido um método para avaliação que subdividiu-se nas etapas descritas na Figura 6.

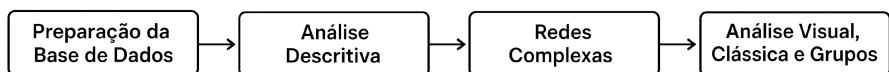


Figura 6 – Fluxograma das etapas do método para avaliação.

Fonte: Autoria Própria

Inicialmente, com a **base de dados** coletada, foram adicionadas duas novas informações: ideologia e coeficiente kappa. A primeira permitiu a segmentação da base de acordo com a filiação ideológica partidária dos deputados, possibilitando a análise comportamental por grupo ideológico, os partidos e suas ideologias. A segunda, o coeficiente kappa, foi adotado como medida de concordância entre os parlamentares, sendo a métrica utilizada para estabelecer as conexões na construção das redes complexas.

A **análise descritiva** teve como finalidade compreender melhor as características gerais da base, bem como os padrões específicos de cada ideologia ao longo do tempo. Essa etapa inicial serviu como subsídio para a posterior interpretação dos resultados oriundos da análise de redes complexas.

O objetivo central deste estudo consistiu em investigar o comportamento dos deputados brasileiros por meio de **redes complexas**, comparando diferentes metodologias. Para isso, foram criadas redes estáticas anuais, seguidas por análises visuais, cálculo de métricas estruturais e extração de grupos a partir de duas metodologias distintas. Os resultados foram analisados e comparados entre si, sendo também relacionados às evidências encontradas na etapa descritiva.

A presente proposta de pesquisa realizou uma **análise** abrangente da estrutura de redes complexas, empregando métodos de análise visual, de detecção de comunidades e métricas de centralidade, tanto os tradicionalmente utilizados na Ciência de Redes quanto

abordagens baseadas em Aprendizado de Máquina. Inicialmente, os dados de votação dos deputados foram obtidos a partir de fontes abertas do governo federal (Governo Federal, 2023). Com esses dados, foi construída uma rede complexa para cada ano analisado, na qual os nós representam os deputados e as arestas indicam interações entre eles, baseadas na similaridade de seus votos. Foram extraídas diversas métricas, incluindo centralidade, coeficiente de *clustering* e grau, além da aplicação de algoritmos de detecção de comunidades e análise de agrupamento. Os resultados obtidos foram analisados e comparados de forma integrada. A pesquisa visa contribuir para uma melhor compreensão da dinâmica estrutural da atuação parlamentar brasileira sob a perspectiva de redes complexas, oferecendo subsídios analíticos relevantes e um modelo metodológico que pode ser aplicado em estudos futuros nesta área.

4.1 Base de dados

Os dados de votação e as informações partidárias foram coletados de fontes públicas disponíveis no site Dados Abertos da Câmara Federal (Governo Federal, 2023). Nesse site estão registradas diversas informações da atividade parlamentar, como votação de projetos, pareceres, gastos dos deputados entre outros. Para este trabalho, foi realizada a coleta abrangendo todas as votações realizadas entre os anos de 2005 a 2023. Para a variável partidária, foi considerado o partido ao qual o deputado estava filiado no respectivo ano. Nos casos em que houve troca de partido dentro do mesmo ano, adotou-se como referência o partido em que o deputado permaneceu por mais tempo durante o período. Os dados foram baixados diretamente de (Governo Federal, 2023) e vieram em formato csv, além das informações trabalhadas neste trabalho, constavam informações as seguintes informações: ID Votacao, URI Votacao, data Hora Voto, voto, deputado ID, deputado URI, deputado nome, deputado siglaPartido, deputado URI Partido, deputado sigla Uf, deputado ID Legislatura, deputado URL Foto. A descrição dos dados das bases de dados geradas é vista na próxima sessão deste trabalho.

Com a base consolidada, algumas modificações foram implementadas. Foi criada a variável ideologia, na qual os partidos políticos foram classificados segundo suas respectivas orientações ideológicas. Para essa categorização, inicialmente foram consultados os sites oficiais de cada partido; nos casos em que a posição ideológica não estava explicitamente definida, realizou-se uma deliberação entre os autores para a definição da ideologia atribuída. Para essa deliberação, os autores fizeram um levantamento no site Wikipedia e consultas a modelos de LLM (ChatGPT, Meta AI e Claude), classificando a ideologia conforme o partido fosse classificado pela maioria desses 4 votos. O resultado final dessa classificação encontra-se apresentado na Figura 7.

A lista de partidos considerada foi: Partido dos Trabalhadores (PT), Partido Comunista do Brasil (PCdoB), Partido Socialismo e Liberdade (PSOL), Partido Pátria

Livre (PPL), Partido Popular Socialista (PPS), Partido Democrático Trabalhista (PDT), Partido Socialista Brasileiro (PSB), Partido Verde (PV), Solidariedade (SDD), Solidariedade (antiga sigla SDD) (SD), Rede Sustentabilidade (REDE), Solidariedade (SOLIDARIEDADE), Cidadania (CIDADANIA), Partido do Movimento Democrático Brasileiro (PMDB), Partido da Social Democracia Brasileira (PSDB), Partido Municipalista Renovador (PMR), Partido Republicano Brasileiro (PRB), Partido de Reedificação da Ordem Nacional (PRONA), Partido da Mobilização Nacional (PMN), Partido Trabalhista do Brasil (PTdoB), Partido dos Aposentados da Nação (PAN), Partido Republicano da Ordem Social (PROS), Sem Partidos (ou legenda avulsa) (S.PART.), Partido Trabalhista Nacional (PTN), Avante (AVANTE), Movimento Democrático Brasileiro (MDB), Partido Trabalhista Cristão (PTC), Democratas (DEM), Partido Humanista da Solidariedade (PHS), Partido Social Democrático (PSD), Partido Social Democrata Cristão (PSDC), Podemos (PODE), Democracia Cristã (DC), União Brasil (UNIÃO), Partido da Frente Liberal (PFL), Progressistas (PP), Partido Trabalhista Brasileiro (PTB), Partido Liberal (PL), Partido Social Cristão (PSC), Partido Ecológico Nacional (PEN), Partido da República (PR), Partido Social Liberal (PSL), Partido Republicano Progressista (PRP), Partido Renovador Trabalhista Brasileiro (PRTB), Partido Novo (NOVO), Patriota (PATRIOTAS) e Republicanos (REPUBLICANOS).

Também foi calculado o coeficiente kappa entre os pares de deputados, além do número de votações em que ambos participaram no mesmo ano. Para garantir a robustez das conexões na construção da rede, foi aplicado um filtro, considerando apenas pares com coeficiente kappa $\geq 0,61$ e com, no mínimo, 30 votações em comum. Essa restrição visa evitar distorções causadas por casos em que dois deputados participaram de apenas uma votação conjunta e concordaram, resultando em um coeficiente kappa igual a 1, mas sem representar uma relação substancial entre eles.

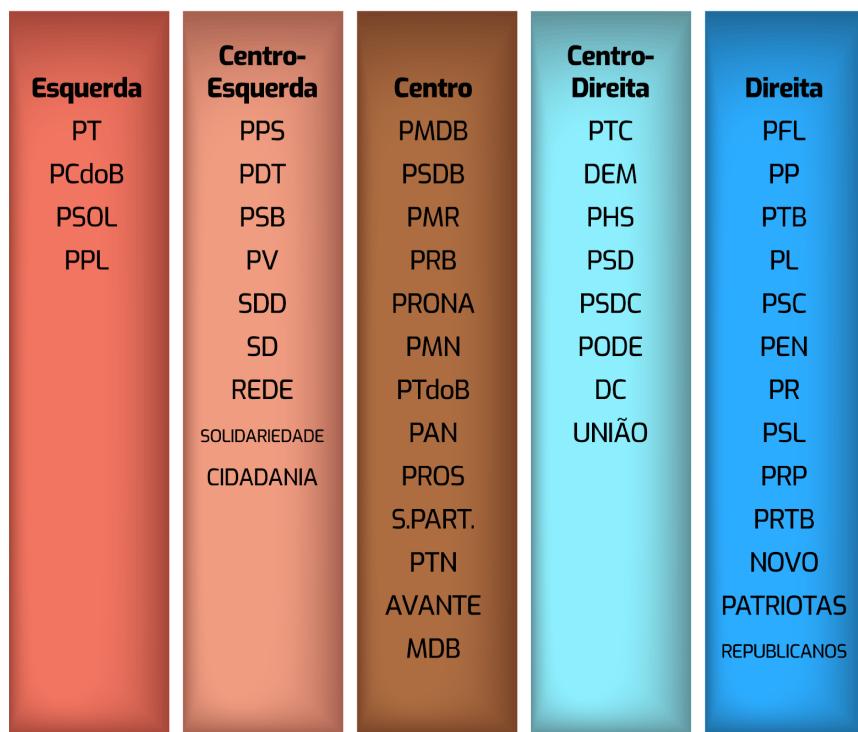


Figura 7 – Distribuição dos Partidos por Ideologia. Clique na sigla para encontrar nome por extenso do partido: PT, PCdoB, PSOL, PPL, PPS, PDT, PSB, PV, SDD, SD, REDE, SOLIDARIEDADE, CIDADANIA, PMDB, PSDB, PMR, PRB, PRONA, PMN, PTdoB, PAN, PROS, S.PART., PTN, AVANTE, MDB, PTC, DEM, PHS, PSD, PSDC, PODE, DC, UNIÃO, PFL, PP, PTB, PL, PSC, PEN, PR, PSL, PRP, PRTB, NOVO, PATRIOTAS e REPUBLICANOS

Fonte: Autoria Própria.

4.2 Análise Descritiva

Por meio de gráficos temporais de contagem e medidas de tendência central, segmentados por ideologia ou de forma agregada, é possível identificar tanto o comportamento geral dos dados quanto os padrões específicos de cada ideologia. A utilização de medidas proporcionais visa evitar vieses decorrentes da volumetria desigual entre as ideologias. O principal objetivo dessa análise é fornecer um diagnóstico geral, utilizando métricas e representações gráficas que sejam representativas do conjunto de dados.

4.3 Redes Complexas

Após a coleta dos dados, procederemos à construção e análise de uma rede complexa utilizando essas informações. A investigação detalhada dessa rede torna-se fundamental para a identificação e compreensão dos comportamentos-chave presentes nos dados,

assim como para o aprimoramento das estratégias de extração de métricas relevantes. A natureza específica da rede gerada desempenhará um papel determinante na definição e refinamento das metodologias utilizadas para extrair e interpretar essas métricas, permitindo uma análise mais precisa e contextualizada.

4.3.1 Construção das Redes Complexas

Para a construção da rede foi adotada a seguinte estrutura: os nós representam os deputados, e as arestas indicam a coincidência ou similaridade das votações entre deputados, calculada via coeficiente kappa de Cohen (Equação 1). As redes são não direcionadas e são ponderadas, sendo o peso da aresta igual ao valor resultante do coeficiente kappa.

4.3.2 Análise visual

Por meio do software Gephi 0.10.1, foi aplicado o algoritmo ForceAtlas2 para a geração do *layout* das redes complexas. Com todas as redes estruturadas para os diferentes anos, foi conduzida uma análise descritiva das configurações geradas, com o objetivo de identificar padrões estruturais ao longo do tempo. Essa análise permitiu compreender como os deputados se organizavam e interagiam em cada período, revelando mudanças significativas nos agrupamentos e nas conexões conforme os governos e as composições ideológicas variavam ao longo dos anos.

4.3.3 Extração das principais métricas

Após a definição das redes complexas, foram extraídas as seguintes métricas gerais: Grau, Centralidade de *Betweenness*, Coeficiente de *Clustering* e Assortividade (apenas geral), calculadas tanto em nível geral quanto segmentadas por ideologia, para cada ano do período analisado.

Na análise geral, foi conduzida uma avaliação de outliers aplicando-se a regra do intervalo interquartil (IQR) multiplicado por 1,5 para cada métrica. Essa abordagem possibilitou identificar a quantidade de deputados que apresentam valores acima do padrão esperado para cada métrica, permitindo assim uma melhor compreensão do comportamento geral da rede em cada ano.

Para a análise das métricas segmentadas por ideologia, foi realizada uma avaliação gráfica temporal, analisando a variação das métricas ao longo do tempo. Dessa forma, foi possível comparar o comportamento de cada ideologia tanto em relação ao seu próprio histórico quanto em relação às demais ideologias.

4.3.4 Detecção de Grupos

A detecção de grupos foi conduzida por meio de duas metodologias distintas, cujos resultados foram posteriormente analisados. Para a detecção de comunidades, foi aplicado o algoritmo de Louvain à rede complexa construída para cada ano. Já para a análise de agrupamento via método Ward, utilizaram-se as principais métricas da rede — Grau, Centralidade de *Betweenness* e Coeficiente de *Clustering* — combinadas com o valor médio de Kappa de cada deputado por ano. Os valores dessas variáveis foram previamente normalizados e, em seguida, aplicou-se o algoritmo hierárquico, adotando-se um ponto de corte fixo no dendrograma (três grupos) para todos os anos.

Com os grupos extraídos por ambas as metodologias, realizou-se uma análise de concordância, verificando-se a sobreposição entre os deputados classificados simultaneamente no grupo 1 da comunidade e no grupo 1 do *cluster*. A proporção de concordância foi então calculada, e, em seguida, foi realizada uma avaliação da composição de cada grupo, com o objetivo de identificar suas diferenças e características específicas.

Devido à aleatoriedade inerente aos algoritmos de agrupamento, cada execução dos modelos (tanto de detecção de comunidades quanto de análise de *cluster*) pode atribuir diferentes rótulos aos grupos gerados. Como foram aplicados dois métodos distintos para cada ano, tornou-se necessário adotar uma estratégia que possibilitasse a comparação direta entre os grupos ao longo do tempo e entre metodologias. Para isso, foi realizada uma reordenação padronizada dos grupos, com base em seu tamanho relativo. Assim, o Grupo 1 passou a representar o grupo com o maior número de deputados, o Grupo 2 o segundo maior, e assim sucessivamente. Essa ordenação visa garantir coerência interpretativa nas análises comparativas entre anos e entre as diferentes abordagens de agrupamento.

CAPÍTULO 5

Experimentos e Análise dos Resultados

Este capítulo apresenta uma análise descritiva da base de dados construída a partir das votações nominais na Câmara dos Deputados, com o objetivo de compreender o comportamento legislativo por meio de redes complexas. Inicialmente, são descritas as principais características da base, incluindo distribuição temporal, volume de interações e padrões gerais de votação. Em seguida, são construídas redes anuais não dirigidas, em que os nós representam deputados e as arestas indicam o grau de concordância entre eles, medido pelo coeficiente kappa. Para cada rede, são calculadas métricas estruturais como número de nós e arestas, grau médio, kappa médio, *betweenness centrality* e *clustering coefficient*. Adicionalmente, são apresentados visualmente os grafos gerados, permitindo observar a formação de blocos e padrões de agrupamento político ao longo do tempo. Por fim, realiza-se uma análise de comunidades por meio de dois métodos: o algoritmo de Louvain e o método de agrupamento hierárquico de Ward, cujos resultados são analisados e comparados a fim de identificar agrupamentos coesos de deputados com comportamentos legislativos semelhantes, contribuindo para revelar alinhamentos ideológicos e estruturas de coalizão dentro do legislativo brasileiro. As análises realizadas neste trabalho foram executadas com o uso da linguagem Python versão 3.10.6 (bibliotecas utilizadas: pandas, numpy, matplotlib, seaborn, networkx e scipy), e as redes complexas foram visualizadas com o software Gephi versão 0.10 (BASTIAN; HEYMANN; JACOMY, 2009).

5.1 Análise Descritiva

A análise descritiva teve como objetivo compreender os padrões gerais dos dados, identificar possíveis inconsistências e verificar se os resultados obtidos nas etapas posteriores estão alinhados com o comportamento observado na base. Para essa etapa, foram realizadas diversas explorações iniciais, incluindo: a contagem total de deputados por ideologia ao longo dos anos (Figura 8), a evolução do número de partidos representados no Congresso (Figura 9), o número total de leis analisadas por ano (Figura 10), a proporção de respostas de votos por ideologia ao longo do tempo (Figura 11), a média de

votações por ideologia (Figura 12), e, por fim, a distribuição dos valores do coeficiente kappa (Figura 13).

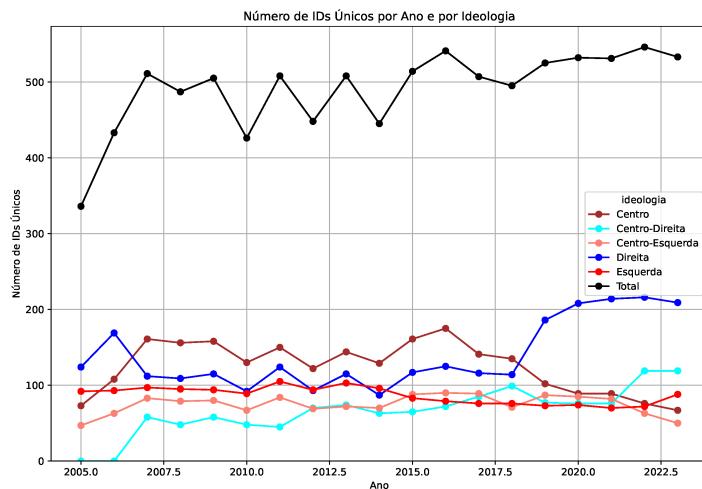


Figura 8 – Quantidade de Deputados por Ideologia ao longo dos anos

Fonte: Autoria Própria

Na Figura 8 temos a quantidade de deputados ao longo dos anos, distribuídos por ideologia. Observa-se que, até 2014, os partidos de Esquerda, Centro-Esquerda, Centro e Direita apresentaram padrões semelhantes de variação. A exceção foi a Centro-Direita, que, embora menos representativa em termos absolutos, apresentou crescimento contínuo durante esse período. Após 2014, verifica-se uma queda na representação da Esquerda, possivelmente associada a uma migração de parlamentares para a Centro-Esquerda. A Centro-Direita também apresentou um crescimento expressivo entre 2014 e 2018. O Centro, por sua vez, cresceu significativamente entre 2014 e 2016, mas sofreu uma redução posterior. Já a Direita manteve relativa estabilidade entre 2014 e 2018, mas passou a apresentar um aumento expressivo após esse período, tornando-se a ideologia com maior número de representantes — o que pode indicar uma migração de parlamentares anteriormente alinhados ao Centro e à Centro-Direita. De modo geral, independentemente da ideologia, observa-se uma tendência de crescimento no número total de deputados ao longo do tempo.

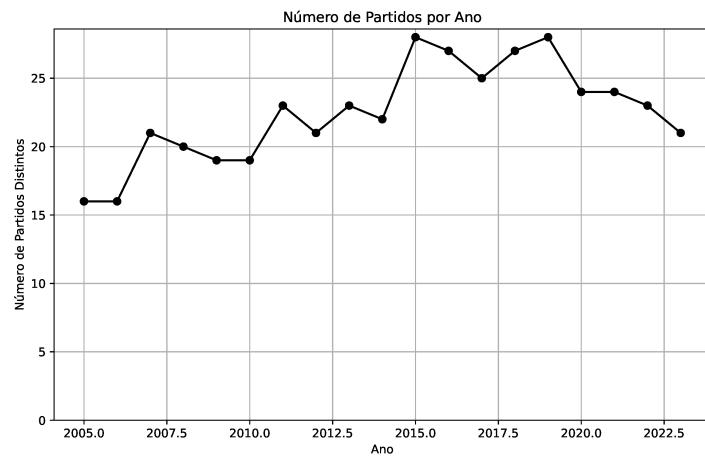


Figura 9 – Quantidade de partidos com representação no Congresso ao longo dos anos.

Fonte: Autoria Própria.

A quantidade de partidos representados é vista na Figura 9, é possível notar um aumento entre 2015 e 2019. A partir desse período, observa-se uma tendência de redução, embora o número de partidos tenha permanecido elevado. Entre 2011 e 2023, o total de legendas com representação parlamentar manteve-se consistentemente acima de 20, atingindo um pico de 29 partidos.

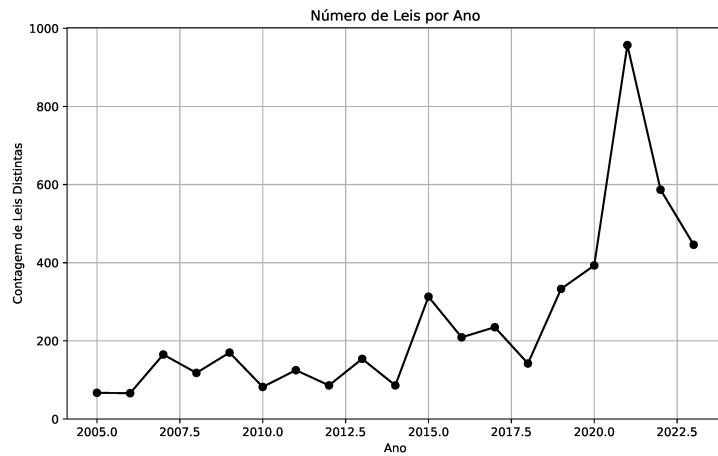


Figura 10 – Quantidade de Leis votadas ao longo dos anos.

Fonte: Autoria Própria.

Na Figura 10 é possível notar que o número de leis votadas manteve um padrão relativamente estável, com variações reduzidas, entre 2005 e 2014. Em 2015, observa-se um pico, seguido por novo crescimento a partir de 2018, alcançando o valor máximo em 2021, ano marcado pela pandemia.

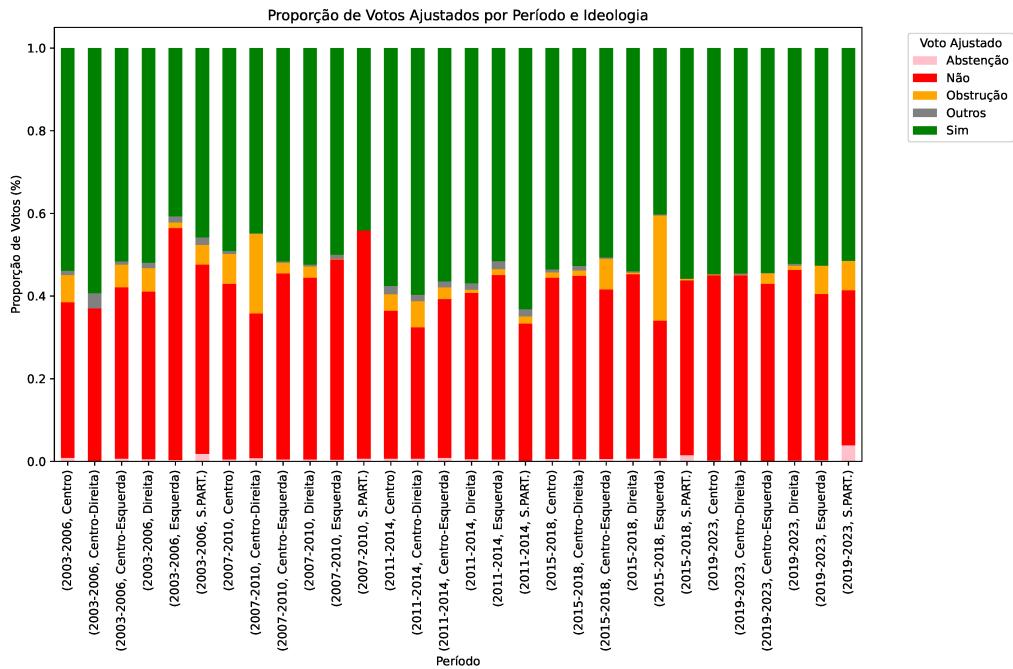


Figura 11 – Proporção da resposta de votação por ideologia ao longo dos anos

Fonte: Autoria Própria

Ao avaliar, por período, a proporção dos tipos de voto por ideologia, visto na Figura 11, observa-se que apenas entre 2003 e 2006 (para a Esquerda) e entre 2007 e 2010 (para parlamentares sem partido) o voto “Não” foi majoritário. No período de 2011 a 2014, a opção “Sim” predominou em todas as ideologias. Destaca-se, ainda, que durante o período do impeachment, a Esquerda apresentou a maior proporção de votos do tipo “Obstrução” de todo o intervalo analisado.

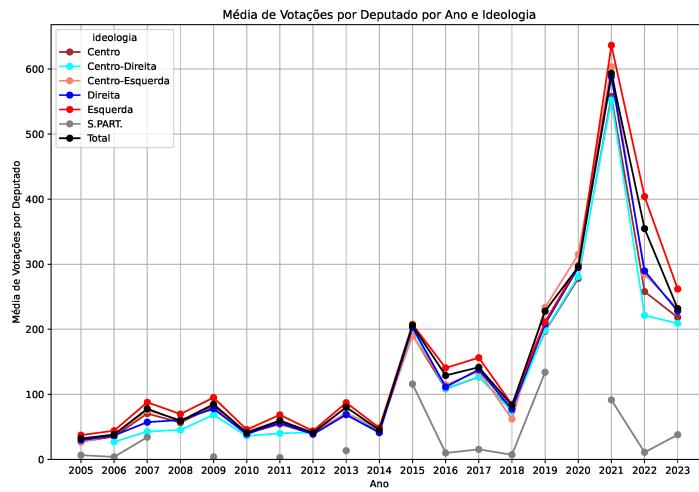


Figura 12 – Média de votação por ideologia ao longo dos anos

Fonte: Autoria Própria

Na Figura 12, nota-se que, de forma geral, o padrão de variação das médias de votação por ideologia segue uma mesma tendência de crescimento e decréscimo ao longo do tempo. No entanto, observa-se que, em apenas dois períodos, a Esquerda não apresentou a maior média de votações; nesses intervalos, a liderança ficou com a Centro-Esquerda. Ao comparar a curva da média de votações com a quantidade de deputados por ideologia é possível inferir que a Esquerda, por contar com um número menor de parlamentares, tende a participar de mais votações por deputado. Em contrapartida, a Direita, por dispor de um contingente maior, apresenta uma média de votações por deputado inferior.

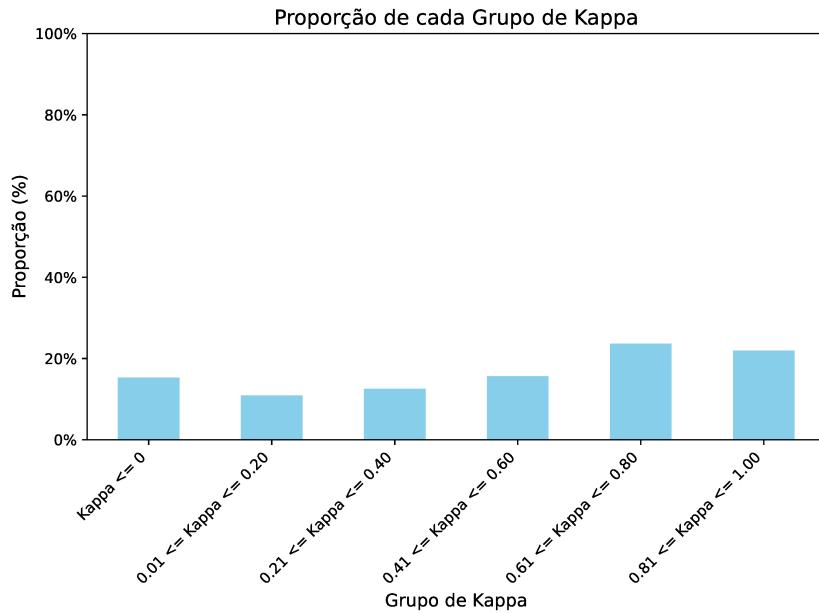


Figura 13 – Proporção de faixa de valor do coeficiente kappa.

Fonte: Autoria Própria.

A Figura 13 mostra o coeficiente kappa agrupado em blocos. Quando analisado em blocos, observa-se que os maiores agrupamentos concentram-se em níveis elevados de concordância (valores superiores a 0,61). Considerando que a rede complexa só é significativa entre parlamentares que compartilham alta concordância, serão utilizadas, para a construção da rede, apenas as observações em que o valor do coeficiente kappa é igual ou superior a 0,61, o que corresponde a 45,6% da base analisada.

5.2 Análise Visual

A análise visual das redes complexas permite a extração de diversos insights sobre o comportamento parlamentar ao longo do tempo. Essa análise foi realizada a partir das Figuras 14 a 18. A Figura 14 contempla os anos de 2005 a 2008, correspondendo ao final do primeiro mandato e início do segundo mandato do governo Lula. A Figura 15 abrange os anos de 2009 a 2012, que compreendem o término do segundo mandato de Lula e o início do primeiro mandato do governo Dilma. A Figura 16 apresenta o período de 2013 a 2016, incluindo o fim do primeiro mandato de Dilma, todo o segundo mandato (interrompido pelo impeachment) e o início do governo Temer. A Figura 17 reúne os anos de 2017 a 2020, cobrindo o final do governo Temer e o início do governo Bolsonaro. Por fim, a Figura 18 representa os anos de 2021 a 2023, abrangendo o término do governo Bolsonaro e o início do novo mandato do presidente Lula.

No primeiro mandato do governo Lula (PT), observa-se que o grupo governista era majoritário e contava com representantes de diversos partidos. A principal oposição,

naquele momento, era composta por setores da Direita e parte do Centro, porém com uma distância baixa em relação ao grupo governista. Com o passar dos anos, a concordância do grupo governista aumentou, evidenciada pela maior proximidade dos nós na rede. A oposição passou a se concentrar no Centro e na Centro-Direita, sendo os anos de 2007 e 2008 marcados por maior isolamento. A partir de 2009, ainda que o grupo governista permanecesse majoritário, os nós se apresentavam mais dispersos, indicando uma redução nesta concordância em relação aos anos anteriores.

No início do governo Dilma, em 2011, a rede indicava elevada concordância entre os parlamentares governistas. No entanto, essa coesão começou a se dissipar já em 2012, quando visualmente se formava um terceiro grande grupo. O isolamento da Esquerda tornou-se evidente, embora até 2014 os demais grupos ainda se apresentassem dispersos, refletindo baixa concordância interna. Em 2015, surgem três grupos claramente definidos, sendo o grupo da Esquerda o menor e formado quase exclusivamente por seus representantes. Apesar de certa aproximação com o grupo majoritário, a concordância entre eles já era limitada. O fim do mandato do governo Dilma ocorreu via impeachment e foi marcado pelo completo isolamento da Esquerda e pela união das demais ideologias em um único grupo opositor, que incluía até mesmo deputados de esquerda.

Durante o governo Temer (2016–2018) a Esquerda manteve-se isolada. O grupo pró-governo era majoritário e apresentava alta concordância, embora no último ano tenha ocorrido um aumento na dispersão, sinalizando maior divergência interna.

Nos dois primeiros anos do governo Bolsonaro (2019–2020), observou-se a maior coesão entre os deputados alinhados ao governo em todo o período analisado. Com o avanço do mandato, o grupo governista tornou-se mais disperso, embora visualmente a concordância entre seus membros ainda permanecesse elevada.

No retorno do PT ao poder (2023), verificou-se a retomada da Esquerda como grupo majoritário. No entanto, a concordância foi a menor entre todos os inícios de mandato analisados. O grupo governista era composto por parlamentares de diferentes ideologias, enquanto a oposição passou a ser majoritariamente da Direita.

De forma geral, os dados indicam que a maioria dos deputados tende a se alinhar com o governo, independentemente da ideologia presidencial. Apesar da alta fragmentação partidária, os agrupamentos de votação são relativamente reduzidos, sugerindo que a quantidade de partidos não reflete, necessariamente, a dinâmica de votação. Observa-se, ainda, que a maior coesão entre os parlamentares ocorre, via de regra, no início dos mandatos presidenciais, diminuindo à medida que o ciclo avança.

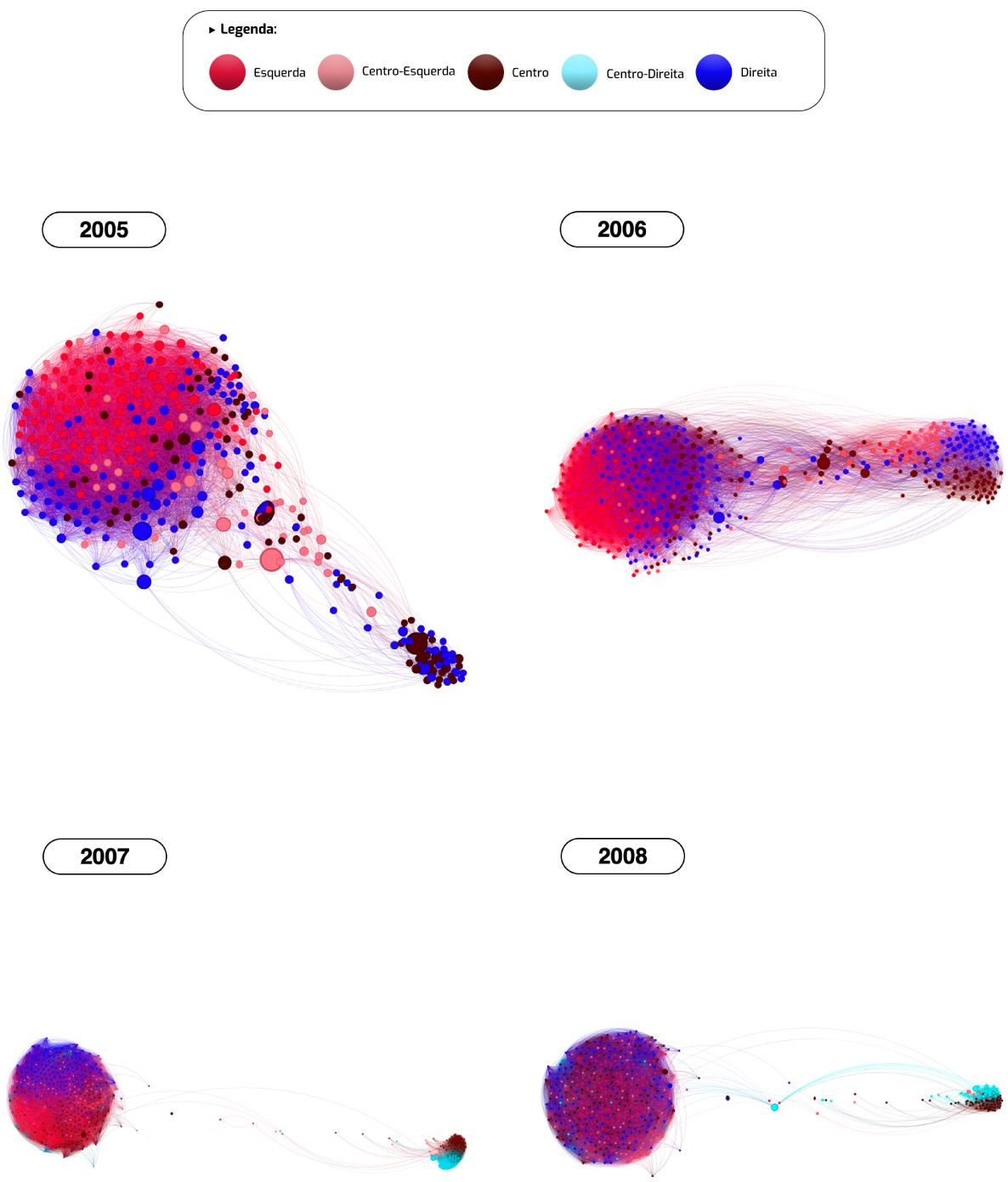


Figura 14 – Visualizações das redes complexas dos deputados brasileiros (2005 a 2008).

Fonte: Autoria Própria.

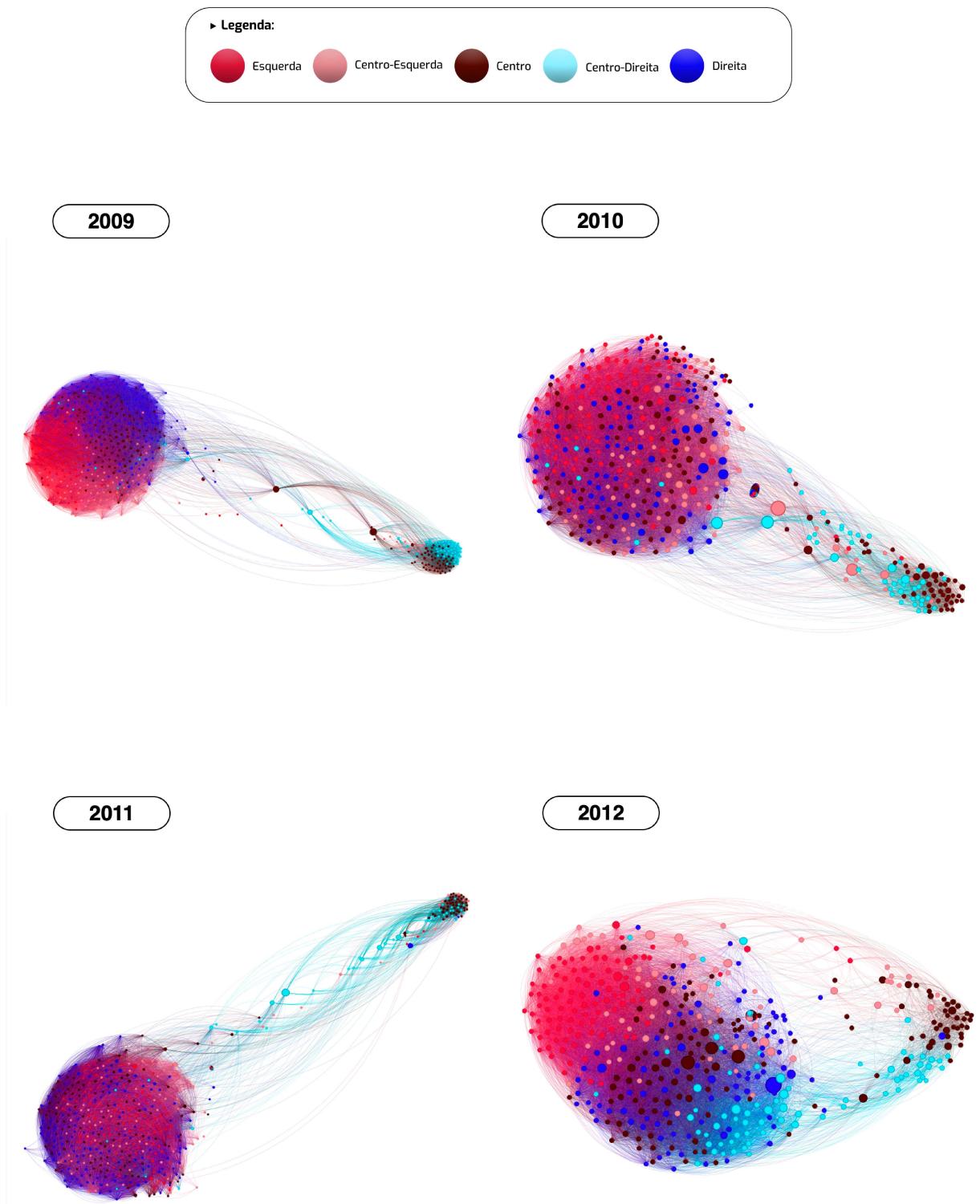


Figura 15 – Visualizações das redes complexas dos deputados brasileiros (2009 a 2012).

Fonte: Autoria Própria.

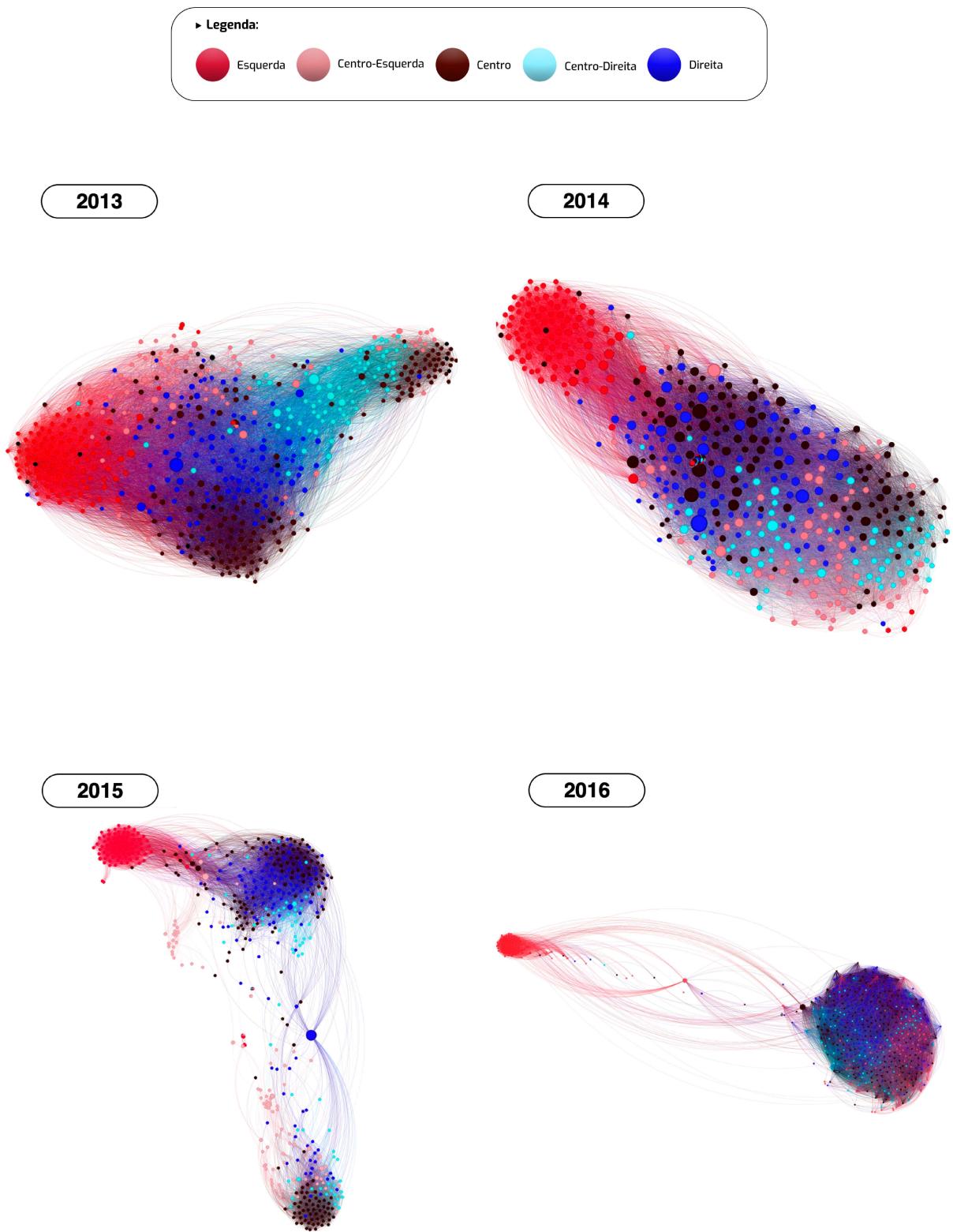


Figura 16 – Visualizações das redes complexas dos deputados brasileiros (2013 a 2016).

Fonte: Autoria Própria.

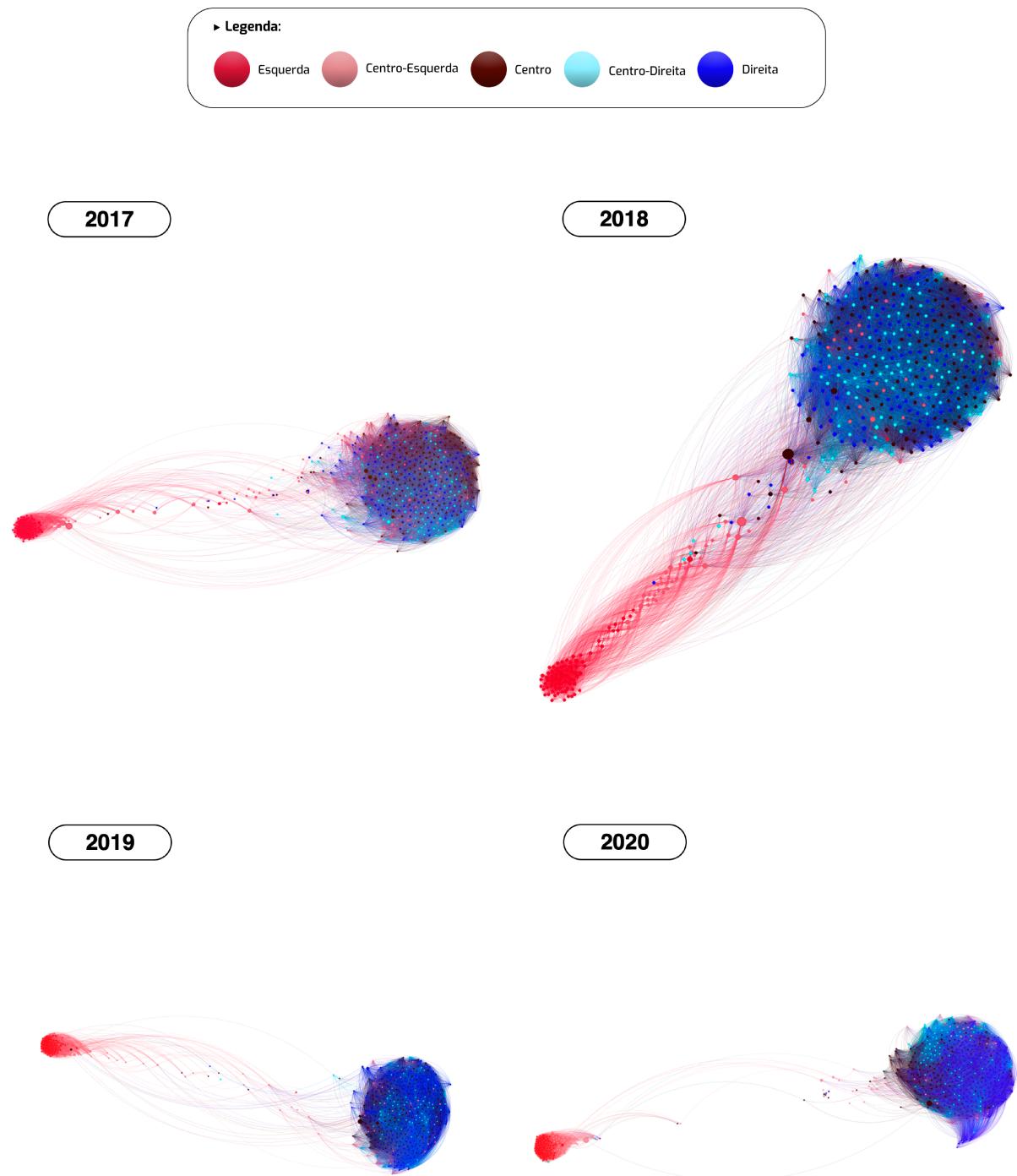


Figura 17 – Visualizações das redes complexas dos deputados brasileiros (2017 a 2020).

Fonte: Autoria Própria.

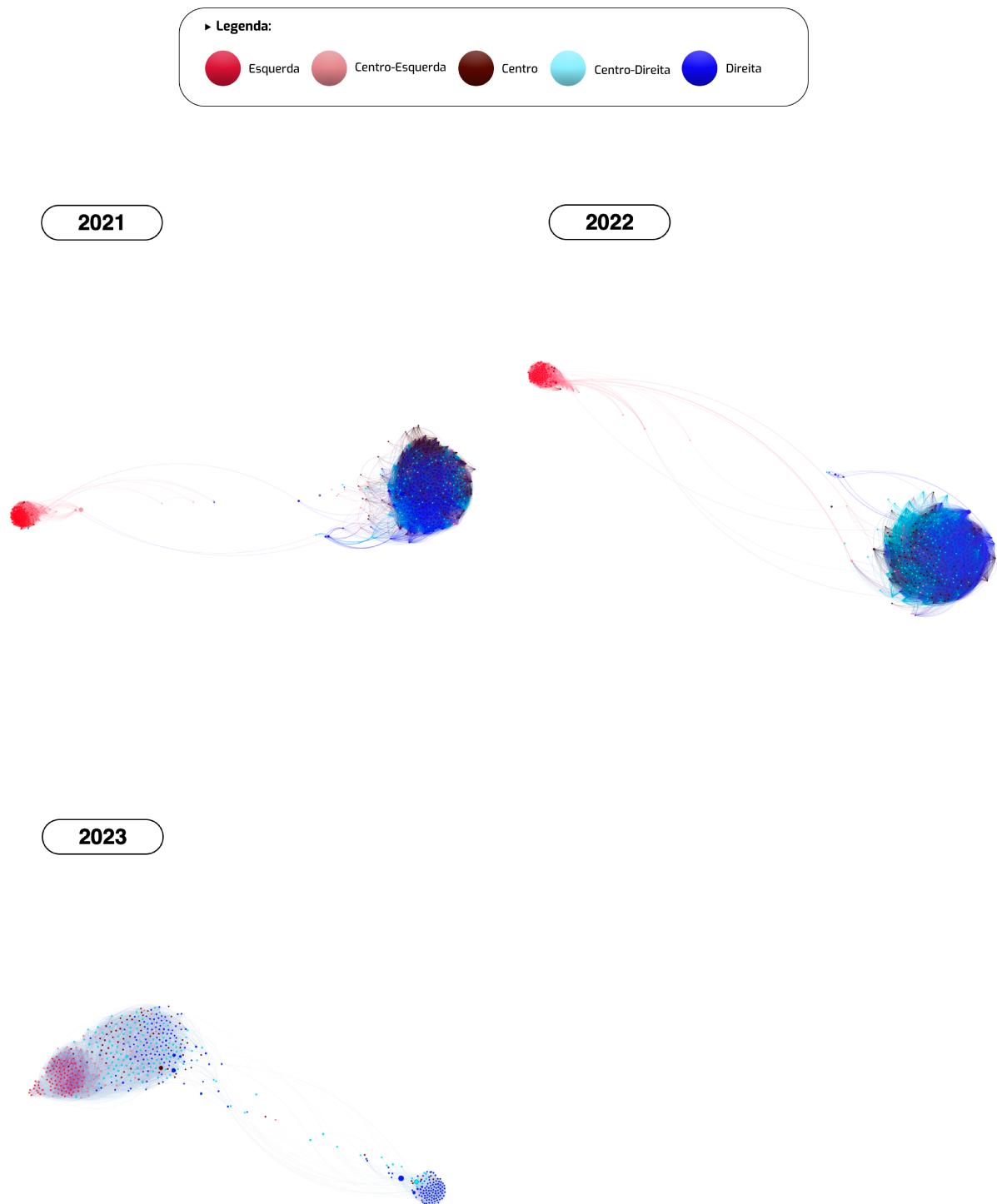


Figura 18 – Visualizações das redes complexas dos deputados brasileiros (2021 a 2023).

Fonte: Autoria Própria.

5.3 Análise Clássica

5.3.1 Análise Geral

Ao longo dos anos observa-se na Tabela 1 que as maiores médias de concordância entre os deputados ocorreram durante o governo Lula (principalmente no início do segundo mandato). Embora os valores de concordância sejam geralmente elevados ao longo de todo o período analisado, os primeiros anos do governo Lula destacam-se como aqueles em que os parlamentares apresentaram maior alinhamento nas votações. Em contraste, os menores valores de grau e número de arestas foram registrados durante o governo Dilma, especialmente em 2015, o que sugere uma baixa coesão entre os deputados e uma maior fragmentação da rede em pequenos grupos. A quantidade de deputados com destaque central em suas respectivas comunidades atingiu o menor valor em 2008, indicando uma distribuição mais equitativa entre os nós da rede naquele ano, e o maior valor em 2015, evidenciando maior concentração de influência em poucos parlamentares. No que se refere à centralidade de intermediação (*betweenness*), o governo Dilma apresenta valores abaixo da média observada, o que reforça a ideia de uma maior segmentação e menor articulação entre os blocos políticos. O coeficiente de *clustering* também foi mais baixo nesse período, indicando uma menor densidade de conexões entre os vizinhos dos deputados, ou seja, uma menor coesão interna. A assortatividade apresenta valores positivos e elevados durante os períodos de maior coesão parlamentar, com picos em 2020 (0,7721) e 2007 (0,7527), indicando que deputados com padrões de votação semelhantes tendem a se conectar entre si, formando comunidades mais homogêneas. Por outro lado, valores próximos a zero ou negativos, como observado em 2005 (-0,1692) e 2010 (0,0042), sugerem uma estrutura de rede mais heterogênea, com maior interação entre deputados com perfis de votação distintos. Esse padrão de baixa assortatividade, combinado com outros indicadores, reforça a interpretação de menor polarização e maior fragmentação da rede em determinados períodos.

O ano de 2015 se destaca como o mais atípico da série, coincidindo com o período que antecede o processo de impeachment da presidente Dilma Rousseff. A análise sugere que anos com indicadores estruturais significativamente baixos, como 2015, podem sinalizar instabilidades políticas relevantes, sendo, portanto, recomendável o monitoramento contínuo desses indicadores para avaliar futuros cenários de fragmentação parlamentar. Embora o ano de 2005 também apresente valores médios relativamente baixos, esse resultado está associado ao menor número de deputados na base de dados para aquele ano. Quando os indicadores são relativizados pela quantidade de nós, 2005 deixa de ser um caso de destaque negativo.

Tabela 1 – Métricas anuais das redes dos parlamentares.

Ano	Kappa Médio (valor)	Qtd de Nós	Qtd de Arestas	Média do Grau	Qtd Betweenness Outlier	Qtd Clustering Outlier	Média do Clustering	Média Assortatividade
2005	0,79	336	14303	85	32	54	0,58	-0,1692
2006	0,76	433	24285	112	24	53	0,54	0,3411
2007	0,8	511	67913	266	22	85	0,72	0,7527
2008	0,85	487	61473	252	22	94	0,74	0,4481
2009	0,75	505	58780	233	16	69	0,63	0,5380
2010	0,77	426	30295	142	2	57	0,59	0,0042
2011	0,83	508	65020	256	16	83	0,72	0,4681
2012	0,75	448	23190	104	22	49	0,47	0,0660
2013	0,72	508	27492	108	18	23	0,42	0,2052
2014	0,75	445	28091	126	18	26	0,47	0,0735
2015	0,71	514	16898	66	32	66	0,52	0,1063
2016	0,78	541	69370	256	15	71	0,68	0,5364
2017	0,76	507	54665	216	19	81	0,66	0,5658
2018	0,8	495	56535	228	15	66	0,68	0,5775
2019	0,77	525	67235	256	18	75	0,69	0,6941
2020	0,78	532	70481	265	21	86	0,72	0,7721
2021	0,77	531	63517	239	25	82	0,71	0,7153
2022	0,78	546	72072	264	19	69	0,73	0,6872
2023	0,73	533	45949	172	22	60	0,59	0,3807

5.3.2 Análise por ideologia

Para esta etapa da análise, foi elaborado um gráfico para cada uma das métricas estudadas — Kappa, Grau, Centralidade de *Betweenness* e Coeficiente de *Clustering* — segmentando os resultados por ideologia e observando sua variação ao longo do tempo. Os resultados obtidos podem ser visualizados nas Figuras 19, 20, 21 e 22.

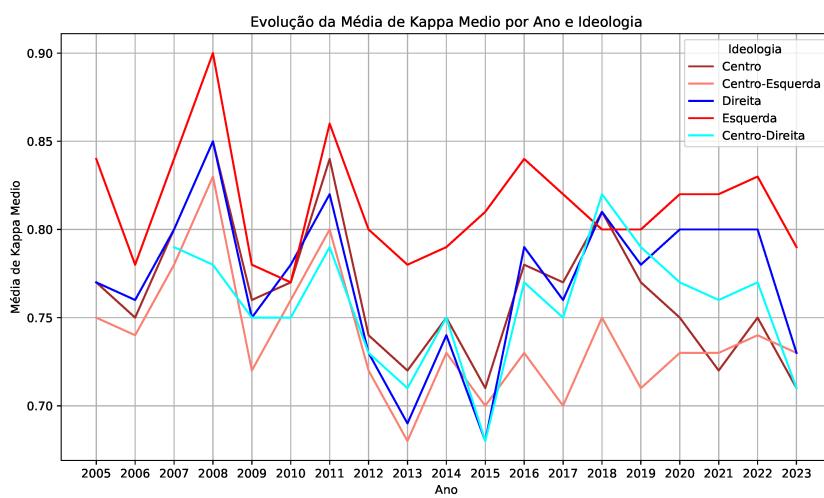


Figura 19 – Kappa médio por ideologia ao longo do tempo.

Fonte: Autoria Própria.

A análise do kappa médio ao longo do tempo, visto na Figura 19, indica que a Es-

querda manteve, de forma geral, os maiores níveis de concordância entre seus membros, com exceção do ano de 2018, em que foi superada por outras ideologias, embora com valores muito próximos. Os partidos de Direita apresentaram uma trajetória de queda na concordância média entre 2010 e 2015, seguida por um crescimento gradual nos anos subsequentes. É relevante destacar que, mesmo durante o período do impeachment, a Esquerda manteve a maior média de concordância, sugerindo maior coesão interna em relação às demais ideologias.

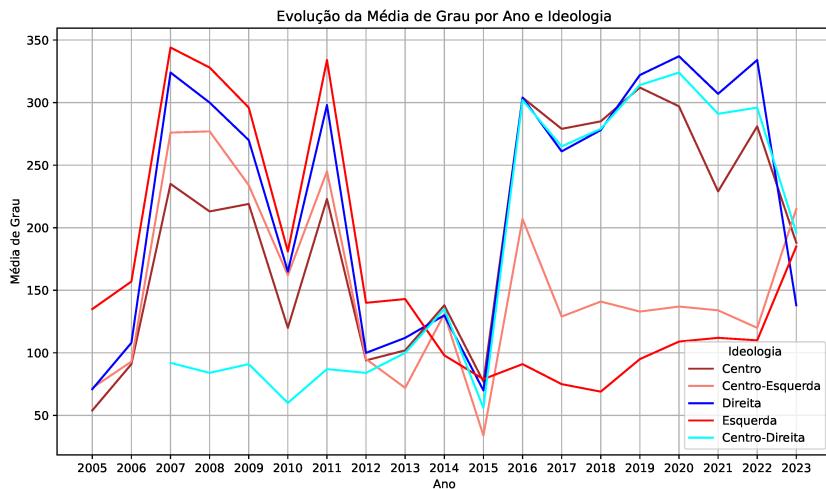


Figura 20 – Grau médio por ideologia ao longo do tempo.

Fonte: Autoria Própria.

Na Figura 20, a análise do grau médio revela que, até 2015, todas as ideologias apresentaram um padrão semelhante de variação, ainda que com valores distintos. Nesse período, a Esquerda registrou os maiores valores, o que é coerente com sua posição no governo. A partir de 2015, observa-se uma significativa divergência entre os grupos ideológicos: Esquerda e Centro-Esquerda passaram a apresentar valores de grau médio consideravelmente inferiores aos das demais ideologias, sugerindo um processo de isolamento. Esse cenário persistiu até 2021, quando se iniciou uma recuperação do grau médio da Esquerda, paralelamente à redução dos valores observados na Direita.

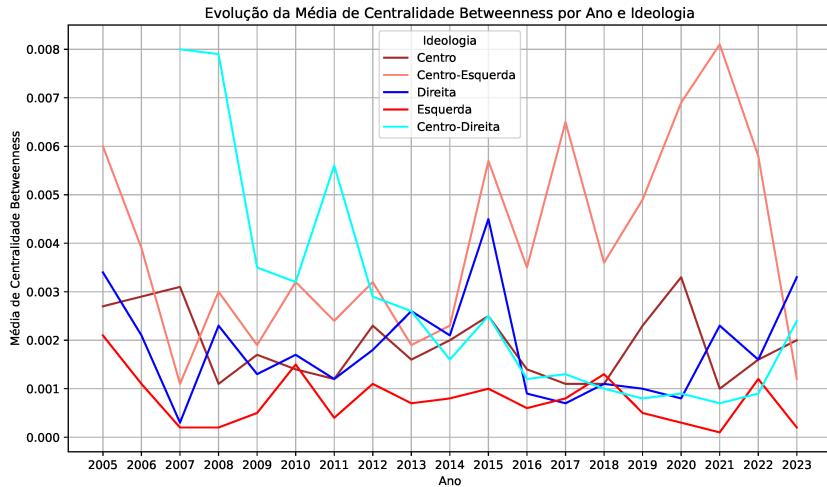


Figura 21 – Centralidade média por ideologia ao longo do tempo.

Fonte: Autoria Própria.

A análise da centralidade ao longo do período visto na Figura 21, revela que no início da série temporal, os partidos de Centro-Direita ocupavam posições mais centrais na rede. Já ao final do período analisado, os maiores valores de centralidade foram observados nos partidos de Centro-Esquerda. A Esquerda, de modo geral, apresentou os menores níveis de centralidade ao longo do tempo, o que sugere uma estrutura mais fechada e menos conectada com outras ideologias. Por outro lado, os partidos de Centro-Esquerda mantiveram, na maior parte do período, os maiores valores de centralidade, indicando uma posição intermediária e a capacidade de articulação com diferentes espectros ideológicos.

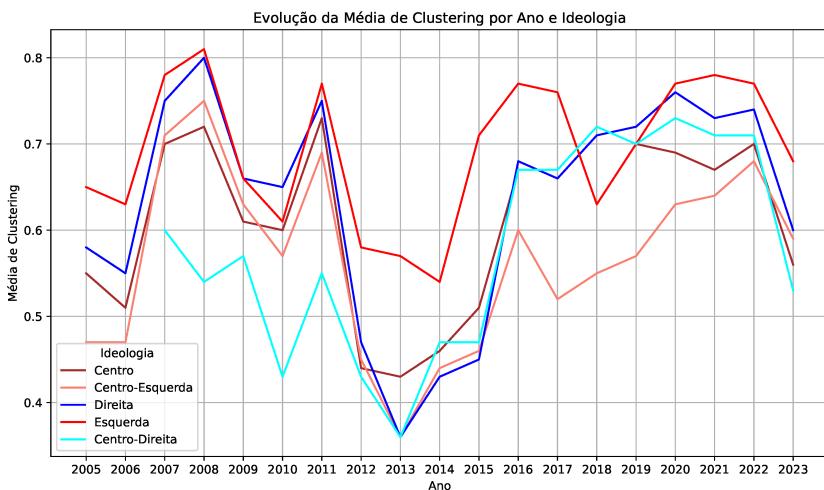


Figura 22 – Clustering médio por ideologia ao londo do tempo

Fonte: Autoria Própria

A análise do coeficiente de agrupamento (*clustering*), visto na Figura 22, mostra que, até 2017, todas as ideologias apresentavam padrões semelhantes. No entanto, nesse ano,

observa-se uma queda no valor médio da Esquerda, enquanto todas as demais ideologias apresentaram aumento. A partir de 2018, o comportamento voltou a se uniformizar entre os grupos ideológicos. Durante a maior parte do período analisado, a Esquerda apresentou os maiores valores médios de *clustering*, sugerindo maior coesão interna e interações mais frequentes dentro do próprio grupo. Em contrapartida, os menores valores médios foram predominantemente observados nos partidos de Centro-Esquerda, indicando menor densidade nas conexões internas.

5.4 Análise de Grupos (Comunidade e Cluster)

Duas metodologias para a extração de grupos foram propostas e comparadas, tanto em um nível geral — considerando a quantidade de grupos formados (Figura 23) — quanto em um nível mais específico, avaliando a composição interna desses grupos (Figuras 24 e 25, e Tabela 2).

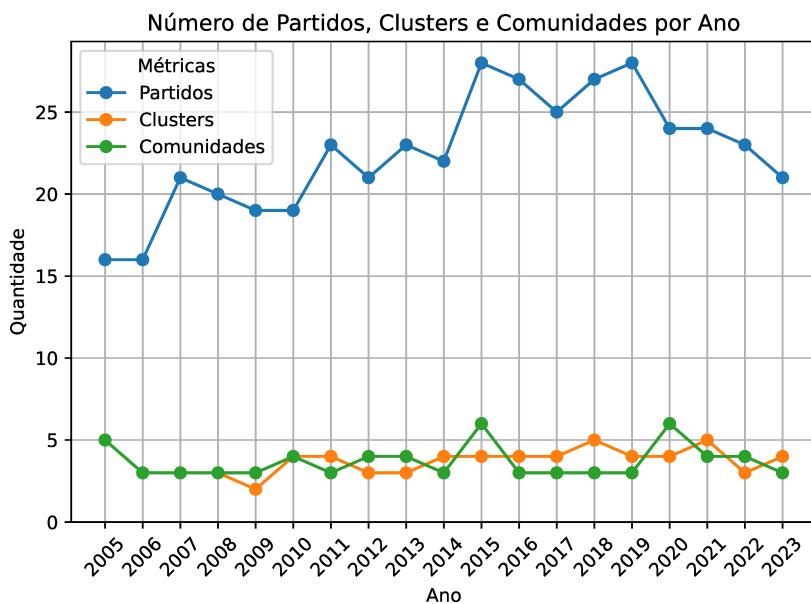


Figura 23 – Número de grupos separado por metodologia ao longo dos anos

Fonte: Autoria Própria

O elevado número de partidos não se traduz em uma grande quantidade de comunidades detectadas por nenhuma das metodologias utilizadas. Os anos de 2015 e 2020 apresentaram o maior número de comunidades, sendo 2015 particularmente marcado por uma alta fragmentação entre os deputados, o que é coerente com o contexto político conturbado vivenciado pelo país naquele período. O menor número de comunidades identificado ao longo da série foi três, valor que também apresentou a maior frequência. Esse padrão é consistente com a análise visual, que aponta para a existência de dois grandes

grupos e um terceiro composto por deputados mais dispersos. Ressalta-se que, considerando a existência de cinco ideologias e um mínimo de 16 partidos ao longo do período analisado — chegando a 29 em determinados anos —, seria esperado um número mais elevado de comunidades. A constância na distribuição das comunidades ao longo dos anos sugere que o número atual de partidos é excessivo frente à estrutura real de agrupamento político observada.

Tabela 2 – Proporção de igualdade das metodologias por ano.

Ano	Proporção de Igualdade
2005	0,375
2006	0,058
2007	0,423
2008	0,474
2009	0,404
2010	0,223
2011	0,596
2012	0,533
2013	0,510
2014	0,400
2015	0,358
2016	0,425
2017	0,572
2018	0,370
2019	0,537
2020	0,600
2021	0,665
2022	0,755
2023	0,452

A Tabela 2 evidencia o grau de divergência entre os métodos de agrupamento no que se refere à classificação dos deputados nos mesmos grupos. Observa-se que, na maior parte dos anos analisados, menos da metade dos parlamentares foram alocados em grupos coincidentes entre os diferentes métodos. O ano de 2022 apresentou o maior índice de concordância, com aproximadamente 75% dos deputados classificados de forma semelhante, enquanto 2006 apresentou o menor índice, com menos de 1% de coincidência entre os agrupamentos. Esses resultados indicam que, mesmo quando o número de grupos identificados é próximo ou idêntico entre os métodos, a distribuição interna dos deputados entre esses grupos pode variar significativamente. Tal variação reforça a necessidade de uma análise qualitativa mais aprofundada para compreender as características e a composição política de cada grupo identificado.

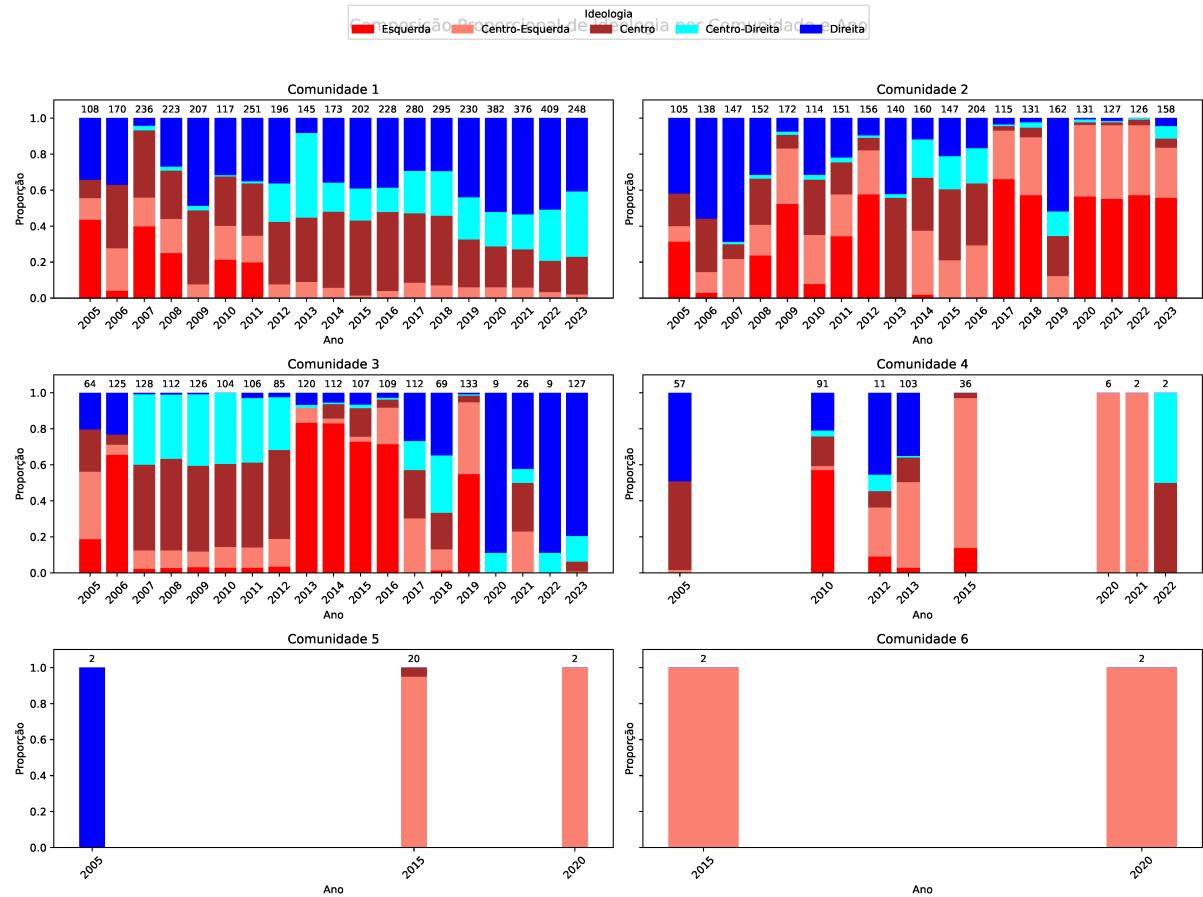
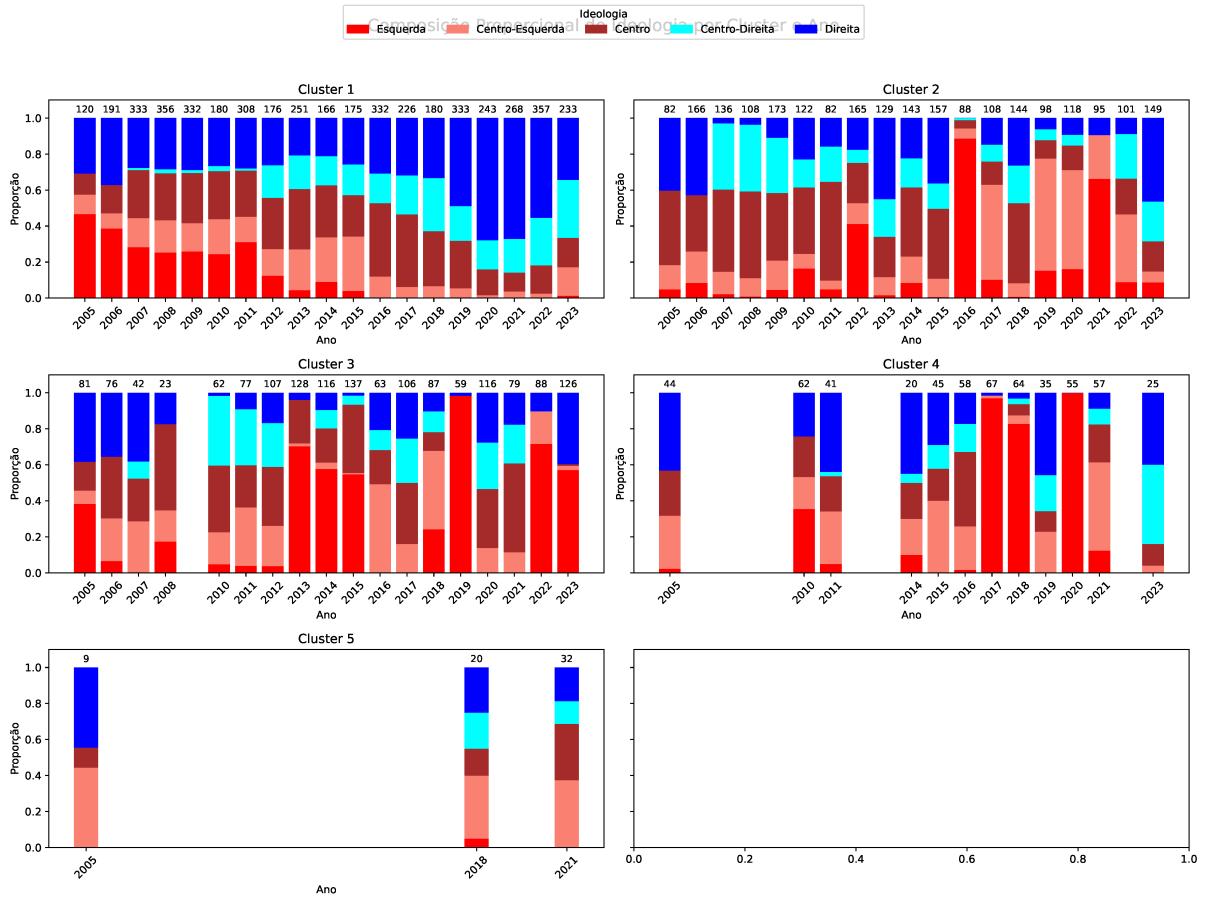


Figura 24 – Composição das comunidades por ideologia ao longo dos anos

Fonte: Autoria Própria

É visto na Figura 24 a segmentação dos grupos identificados por meio da detecção de comunidades, o qual revela uma estrutura mais detalhada da rede complexa. Durante o governo Lula (até 2010), a oposição esteve majoritariamente concentrada na Comunidade 3 (menos nos anos de 2005 (que ficou na comunidade 4) e 2006 (comunidade 1)). No período do governo Dilma, a oposição ficou mais claramente concentrada na Comunidade 1, enquanto os deputados governistas se distribuíram entre as Comunidades 2, 3 e 4 — com destaque para o ano de 2016, no qual a Esquerda se isolou na Comunidade 3.

Com a queda de Dilma Rousseff, as Comunidades 1 e 3 assumiu a posição de governo, e a Comunidade 2 passou a representar a oposição, papel que manteve até retornar ao poder em 2023. As demais comunidades, por apresentarem menor volume, refletem subdivisões internas nos grandes componentes da rede. A ausência de uma configuração claramente dominante de algumas comunidades indica uma segmentação mais refinada dentro dos principais agrupamentos da rede complexa. Um exemplo dessa dinâmica ocorre em 2006, quando, apesar de os Componentes 1 e 3 estarem vinculados ao governo, a comunidade ideologicamente mais alinhada estava concentrada na Comunidade 3, que apresentava o maior volume, em consonância com a análise visual.

Figura 25 – Composição dos *clusters* por ideologia ao longo dos anos

Fonte: Autoria Própria

A análise dos *clusters*, vista na Figura 25, revelou padrões consistentes na estrutura das redes ao longo do tempo. O *Cluster 1* esteve majoritariamente vinculado ao governo em exercício, com uma composição que variou conforme os mandatos: observou-se uma diminuição da presença da Esquerda e um crescimento da Direita ao longo dos anos, seguido posteriormente por nova queda da Direita. Durante o governo Dilma, embora o *Cluster 1* estivesse relacionado ao governo, ele praticamente não incluía deputados de Esquerda, o que é evidenciado na análise visual da rede. Durante este governo, o *Cluster 3* passou a representar de forma mais fiel os deputados mais alinhados ao governo de esquerda (porém, no ano do impeachment (2016) o *Cluster* que melhor representava o governo de esquerda é o 2). O *Cluster 2* por sua vez, aparenta ser composto pela oposição ao governo vigente. Já o *Cluster 3* aparenta estar ligado ao governo vigente, porém não tão associado quanto ao *Cluster 1* (salvo no governo Dilma, que este *Cluster* representava melhor o governo vigente). Deputados com baixa conectividade, que apresentam não se integrar claramente a nenhum grupo maior, estiveram majoritariamente relacionados ao *Cluster 4* — embora, em alguns anos, como 2018, este tenha se consolidado como um grupo estruturado. Por fim, o *Cluster 5* agregou parlamentares que não apresentavam vínculos diretos com os principais blocos da rede. De forma geral, a clusterização foi

eficaz em caracterizar os principais agrupamentos políticos, delimitando de forma clara os blocos governistas, oposicionistas e independentes. Um ponto de destaque é que, no atual governo do PT, a estrutura da rede assumiu uma configuração tão distinta que o grupo governista se formou de maneira isolada, algo que, até então, só havia sido observado durante o governo Dilma.

CAPÍTULO 6

Conclusão

O presente trabalho teve como objetivo principal **investigar o comportamento dos deputados federais brasileiros ao longo do tempo por meio da construção e análise de redes complexas ponderadas pela concordância ideológica em votações parlamentares**. Para isso, foram utilizadas abordagens complementares — análise visual, métricas de redes e algoritmos de agrupamento — que permitiram identificar padrões de alinhamento, coesão e transformação ideológica nas interações legislativas.

De forma geral, a análise descritiva permitiu identificar dinâmicas de mudança ideológica entre os deputados ao longo do tempo. Observa-se, por exemplo, que quando a representação da esquerda diminui, há um crescimento proporcional da centro-esquerda; o declínio do centro é acompanhado por uma ascensão da centro-direita; e, por fim, essa também perde espaço para o aumento da direita. Tais movimentos sugerem que, independentemente da ideologia, possíveis crises ou desgastes associados a determinados espectros políticos provocam o afastamento de parlamentares dessas ideologias. Já na análise preliminar, antes mesmo da construção da rede, verificou-se que, de forma geral, os deputados mais discordam entre si do que concordam.

Visualmente, os grafos permitem identificar padrões consistentes de alinhamento e coesão parlamentar ao longo dos diferentes governos. Independentemente da ideologia presidencial, há uma tendência de formação de blocos majoritários de apoio ao governo, especialmente no início dos mandatos, caracterizados por maior coesão e concordância entre os parlamentares. Essa coesão, no entanto, tende a se dissipar com o tempo. É possível observar um processo de “migração” ideológica de deputados, especialmente nos momentos de transição de governo, com muitos se aproximando da nova ideologia no poder. Quando as métricas são avaliadas por ideologia ao longo do tempo, nota-se que a maior parte dos deputados tende a apoiar o governo vigente: durante mandatos de esquerda, por exemplo, a esquerda apresentou os maiores valores de grau médio, padrão que se inverte quando a direita assume o poder.

Ainda assim, a esquerda apresenta valores consistentemente mais elevados de concordância (κ) e coeficiente de *clustering*, o que pode ser explicado pelo fato de possuir,

em geral, menor número de deputados. Isso implica na necessidade de maior coesão interna para exercer influência no legislativo. Ao avaliar os governos individualmente, destaca-se negativamente o segundo mandato de Dilma Rousseff, que apresentou os piores indicadores de coesão e centralidade nas redes analisadas. Por outro lado, o primeiro mandato de Lula figura como o mais positivo, evidenciando forte adesão da maioria dos deputados ao governo, independentemente da ideologia.

No que tange à quantidade de partidos representados na Câmara, que chegou a 29 em determinados anos, observa-se que esse número elevado não se reflete em uma fragmentação proporcional nos agrupamentos detectados. Em todas as abordagens — análise visual, detecção de comunidades e *clustering* — identificaram-se pequenos blocos de atuação política que transcendem a lógica partidária formal. Em geral, há um grupo majoritário de apoio ao governo e outro menor, de oposição. No entanto, há diferenças relevantes dependendo da ideologia presidencial: quando a direita está no poder, os grupos governistas tendem a excluir parlamentares de esquerda; já quando a esquerda governa, observa-se a presença de deputados de diferentes espectros ideológicos em seu bloco de apoio. Tal padrão pode indicar uma maior coerência da esquerda, uma vez que, mesmo fora do poder, não se observa sua adesão a blocos formados com a direita para integrar a base governista.

Em relação às metodologias empregadas, destaca-se a possibilidade de uso complementar ou comparativo entre elas. A análise visual contribui para a interpretação inicial e intuição da rede; as métricas oferecem fundamentos quantitativos robustos; e os algoritmos de agrupamento — como *clustering* e detecção de comunidades — permitem segmentações que variam do mais geral ao mais específico. O modelo de *cluster* se mostrou mais abrangente, ideal para interpretações iniciais, enquanto a detecção de comunidades forneceu uma visão mais refinada da segmentação interna dos grandes blocos. Ambas são válidas, e sua escolha deve considerar o objetivo da análise. No presente trabalho, o uso integrado das três abordagens — visualização, métricas e agrupamento — permitiu uma interpretação mais precisa e assertiva das redes parlamentares, e pode servir de base para futuras investigações com dados similares.

Como limitação deste trabalho, destaca-se a impossibilidade de identificar as causas dos comportamentos observados nas redes. Como não foram incorporadas variáveis externas ao modelo, não foi possível realizar uma análise estatística que indicasse os principais fatores associados à dinâmica e às variações no comportamento das redes complexas ao longo do tempo.

6.1 Principais Contribuições

Foi possível construir redes complexas capazes de representar como os parlamentares brasileiros se relacionam entre si. Ademais, o estudo evidenciou vantagens e limitações de dois métodos distintos de agrupamento, demonstrando que ambos podem ser úteis, depen-

dendo da natureza do problema investigado — especialmente se há ou não conhecimento prévio sobre o comportamento dos dados. Além disso, o trabalho permite a elaboração de um manual de boas práticas para análise baseada em redes complexas, que pode seguir uma sequência estruturada: iniciar com uma análise descritiva, construir a rede, e então realizar uma investigação detalhada das métricas de rede aliada à análise visual.

6.2 Trabalhos Futuros

Como proposta para trabalhos futuros, seria interessante replicar o presente estudo incorporando variáveis externas que representem o contexto socioeconômico do Brasil em cada ano, a partir de critérios bem definidos. Uma possibilidade seria considerar os principais acontecimentos nos cinco setores que mais contribuem para o Produto Interno Bruto (PIB) do país em cada período e analisar de que forma tais eventos impactam as redes formadas.

6.3 Contribuições em Produção Bibliográfica

O artigo intitulado “Relationships among Brazilian Politicians: A Study Using Network Science and Machine Learning” foi submetido ao *Journal of Complex Networks*.

Referências

AMARAL, O. O que sabemos sobre a organização dos partidos políticos: uma avaliação de 100 anos de literatura. **Revista Debates**, v. 7, n. 2, p. 11–32, 2013. DOI: 10.22456/1982-5269.38429.

ANDERBERG, M. R. The broad view of cluster analysis. **Cluster analysis for applications**, Elsevier, v. 1, n. 1, p. 1–9, 1973. DOI: 10.1016/B978-0-12-057650-0.50007-7.

BARABÁSI, A.-L. **Linked: The new science of networks**. [S.l.]: American Association of Physics Teachers, 2003. DOI: 10.1119/1.1538577.

BARABÁSI, A.-L. **Network Science**. Cambridge University Press, 2016. Disponível em: <<http://networksciencebook.com>>.

BARABÁSI, A.-L.; ALBERT, R. Emergence of scaling in random networks. **Science**, v. 286, n. 5439, p. 509–512, 1999. DOI: 10.1126/science.286.5439.509.

BASTIAN, M.; HEYMANN, S.; JACOMY, M. **Gephi: An Open Source Software for Exploring and Manipulating Networks**. 2009. DOI: 10.1609/icwsm.v3i1.13937. Disponível em: <<http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>>.

BISHOP, C. M. **Pattern Recognition and Machine Learning**. [S.l.]: Springer, 2006.

BLONDEL, V. D. et al. Fast unfolding of communities in large networks. **Journal of statistical mechanics: theory and experiment**, IOP Publishing, v. 2008, n. 10, p. P10008, 2008. DOI: 10.1088/1742-5468/2008/10/P10008.

BRAGA, M. d. S. S.; Pimentel Jr, J. Os partidos políticos brasileiros realmente não importam? **Opinião pública**, SciELO Brasil, v. 17, p. 271–303, 2011. DOI: 10.1590/S0104-62762011000200001.

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. DOI: 10.1371/journal.pone.0229928.

COHEN, J. A coefficient of agreement for nominal scales. **Educational and psychological measurement**, Sage Publications Sage CA: Thousand Oaks, CA, v. 20, n. 1, p. 37–46, 1960. DOI: 10.1177/001316446002000104.

COHEN, R.; HAVLIN, S. **Complex networks: structure, robustness and function.** [S.l.]: Cambridge university press, 2010. DOI: 10.1017/CBO9780511780356.

ERDŐS, P.; RÉNYI, A. On random graphs i. **Publicationes Mathematicae**, v. 6, p. 290–297, 1959. DOI: 10.5486/PMD.1959.6.3-4.12.

ESTER, M. et al. A density-based algorithm for discovering clusters in large spatial databases with noise. In: **kdd**. [S.l.: s.n.], 1996. v. 96, n. 34, p. 226–231.

FÁVERO, L. P.; BELFIORE, P. **Manual de análise de dados: estatística e modelagem multivariada com Excel®, SPSS® e Stata®.** [S.l.]: Elsevier Brasil, 2017.

FREEMAN, L. C. A set of measures of centrality based on betweenness. **Sociometry**, JSTOR, p. 35–41, 1977. DOI: 10.2307/3033543.

FREEMAN, L. C. et al. Centrality in social networks: Conceptual clarification. **Social network: critical concepts in sociology.** Londres: Routledge, v. 1, p. 238–263, 2002.

GIRVAN, M.; NEWMAN, M. E. Community structure in social and biological networks. **Proceedings of the national academy of sciences**, The National Academy of Sciences, v. 99, n. 12, p. 7821–7826, 2002. DOI: 10.1073/pnas.122653799.

Governo Federal. **Os dados da Câmara a qualquer hora, ao seu dispor.** 2023. Website: <https://dadosabertos.camara.leg.br/>.

GRADY, D.; THIEMANN, C.; BROCKMANN, D. Robust classification of salient links in complex networks. **Nature communications**, Nature Publishing Group UK London, v. 3, n. 1, p. 864, 2012. DOI: 10.1038/ncomms1847.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction.** [S.l.]: Springer, 2009. DOI: 10.1007/978-0-387-84858-7.

JACOMY, M. et al. Forceatlas2, a continuous graph layout algorithm for handy network visualization designed for the gephi software. **PloS one**, Public Library of Science San Francisco, USA, v. 9, n. 6, p. e98679, 2014. DOI: 10.1371/journal.pone.0098679.

LANCICHINETTI, A.; FORTUNATO, S.; RADICCHI, F. Benchmark graphs for testing community detection algorithms. **Physical review E**, APS, v. 78, n. 4, p. 046110, 2008. DOI: 10.1103/PhysRevE.78.046110.

LINHARES, C. D. et al. Visual analysis for evaluation of community detection algorithms. **Multimedia Tools and Applications**, Springer, v. 79, p. 17645–17667, 2020. DOI: 10.1007/s11042-020-08700-4.

MACIEL, A. P. B.; ALARCON, A. de O.; GIMENES, É. R. Partidos políticos e espectro ideológico: parlamentares, especialistas, esquerda e direita no brasil. **Revista Eletrônica de Ciência Política**, v. 8, n. 3, 2018. DOI: 10.5380/recp.v8i3.54834.

MACQUEEN, J. Some methods for classification and analysis of multivariate observations. In: UNIVERSITY OF CALIFORNIA PRESS. **Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics.** [S.l.], 1967. v. 5, p. 281–298.

MCPHERSON, M.; SMITH-LOVIN, L.; COOK, J. M. Birds of a feather: Homophily in social networks. **Annual Review of Sociology**, Annual Reviews, v. 27, n. 1, p. 415–444, 2001. DOI: 10.1146/annurev.soc.27.1.415.

MITCHELL, T. M. **Machine learning**. [S.l.]: McGraw-Hill, 1997.

NEWMAN, M. E. Assortative mixing in networks. **Physical Review Letters**, APS, v. 89, n. 20, p. 208701, 2002. DOI: 10.1103/PhysRevLett.89.208701.

NEWMAN, M. E. Mixing patterns in networks. **Physical Review E**, APS, v. 67, n. 2, p. 026126, 2003. DOI: 10.1103/PhysRevE.67.026126.

NEWMAN, M. E. The structure and function of complex networks. **SIAM review**, SIAM, v. 45, n. 2, p. 167–256, 2003. DOI: 10.1137/S003614450342480.

NEWMAN, M. E. J. Power laws, pareto distributions and zipf's law. **Contemporary Physics**, v. 46, n. 5, p. 323–351, 2005. DOI: 10.1080/00107510500052444.

NEWMAN, M. E. **Networks: An Introduction**. Oxford: Oxford University Press, 2010. DOI: 10.1093/acprof:oso/9780199206650.001.0001.

PASTOR-SATORRAS, R.; VESPIGNANI, A. Epidemic spreading in scale-free networks. **Physical Review Letters**, APS, v. 86, n. 14, p. 3200–3203, 2001. DOI: 10.1103/PhysRevLett.86.3200.

PEARSON, K. Liii. on lines and planes of closest fit to systems of points in space. **The London, Edinburgh, and Dublin philosophical magazine and journal of science**, Taylor & Francis, v. 2, n. 11, p. 559–572, 1901. DOI: 10.1080/14786440109462720.

QUINLAN, J. R. Simplifying decision trees. **International journal of man-machine studies**, Elsevier, v. 27, n. 3, p. 221–234, 1987. DOI: 10.1016/S0020-7373(87)80053-6.

REIS, F. W. Identidade política, desigualdade e partidos brasileiros. **Novos estudos CEBRAP**, SciELO Brasil, p. 61–75, 2010.

RODRIGUES, L. M. Partidos, ideologia e composição social. **Revista Brasileira de Ciências Sociais**, SciELO Brasil, v. 17, p. 31–47, 2002. DOI: 10.1590/S0102-69092002000100004.

TAROUCO, G. d. S.; MADEIRA, R. M. Os partidos brasileiros segundo seus estudiosos: análise de um expert survey. **Civitas-Revista de Ciências Sociais**, SciELO Brasil, v. 15, p. e24–e39, 2015. DOI: 10.15448/1984-7289.2015.1.18077.

TUKEY, J. W. et al. **Exploratory data analysis**. [S.l.]: Springer, 1977. v. 2.