Submission date: 13/11/2025
Resubmission date: 17/03/2026
Camera ready submission date: 18/06/2026

1st round notification: 09/02/2026
2nd round notification: 18/05/2026
Available online: 30/07/2026
Publication date: 30/07/2026

Section: Extended versions of papers presented at the XXI Brazilian Symposium on Information Systems (SBSI 2025)

Uncovering Political Alliances and Their Temporal Evolution in the Brazilian Congress Voting Network

Pedro N. F. da Silva[1], **Dimas Cassimiro Nascimento**[2], **Bruno Nogueira**[3]

[1]Programa de Pós Graduação da Universidade de Pernambuco – UPE
Garanhuns, PE – Brazil

[2]Universidade Federal do Agreste de Pernambuco – UFAPE
Garanhuns, PE – Brazil

[3]Universidade Federal de Alagoas – UFAL
Maceió, AL – Brazil

pnfs@ecomp.poli.br, dimas.cassimiro@ufape.edu.br, bruno@ic.ufal.br

Abstract. *Network analysis reveals alliances, ideological shifts, and party dynamics, and can support transparency and evidence-based decision-making in democratic governance. Many studies ignore non-polarized filtering and edge pruning, which lowers modularity and clarity. This research applies the Leiden algorithm to polarized votes in Brazil's Congress, using edge pruning to detect cohesive communities. Grounded in Social Network Theory and Institutional Theory, it examines how alliances form under institutional norms. Using public roll-call voting data (formal legislative outcomes), we filtered polarized votes and optimized the network. To quantify temporal stability and evolution, we introduce three temporal metrics: the Herfindahl-Hirschman Index (HHI) for community concentration, the Adjusted Rand Index (ARI) for community similarity across years, and Edge Volatility for network structural changes. Additionally, we employ backbone comparison using the backbone extraction to identify statistically significant network connections. This work advances Information Systems by demonstrating how filtering and optimization expose hidden structures and political dynamics, while providing quantitative measures of temporal community evolution. Results show a 10.95% modularity gain compared to the backbone extraction and reveal clear ideological blocs and a "swing" community, which can be used by journalists, oversight institutions, and citizens to monitor coalition shifts and accountability over time.*

Keywords. *Community Detection; Network Modularity; Political Polarization; Graph Theory; Temporal Analysis; Decision Support; Transparency; Governance.*

1. Introduction

Political voting patterns are crucial for understanding the dynamics of legislative decision-making, especially in multiparty systems like Brazil, where alliances frequently shift, revealing significant changes in political power and influence. The availability of open data,

Cite as: Silva, P. N. F., Nascimento, D. C. & Nogueira, B. (2026). Uncovering Political Alliances and Their Temporal Evolution in the Brazilian Congress Voting Network. iSys: Revista Brasileira de Sistemas de Informação (iSys - Journal of Information Systems), 19(1), 7:1-7:32. DOI: 10.5753/isys.2026.6878

combined with increased computing power and technological advancements, has led to a surge in research utilizing complex networks [Brito et al. 2020]. Network analysis offers powerful tools to interpret the dynamics of complex systems, such as a country's political landscape. However, capturing the intricacies of political systems is challenging due to their inherent complexity and the multifaceted nature of political interactions [Maso et al. 2014].

Voting records should, in theory, reflect the ideological positions of elected representatives, providing a window into the underlying political dynamics. By modeling relationships based on voting behavior, we can construct a system that visually represents these connections, as supported by previous works [Moody and Mucha 2013, Cherepnalkoski et al. 2016]. The growing application of network theory in legislative studies thus offers a structured method to capture and visualize these dynamics, contributing to more informed predictions of legislative outcomes. Moreover, such studies have the potential to benefit not only political analysts but also the general population, as they can inspire the development of user-friendly tools to help citizens better understand political processes and improve the quality of their participation and voting decisions.

From an IS perspective, the beneficiaries of this work include (i) *journalists* and *fact-checkers* who need evidence to explain coalition behavior, (ii) *oversight institutions* (e.g., courts of accounts and anti-corruption agencies) and *legislative observatories* that monitor accountability, (iii) *parties and parliamentary staff* who require decision support to anticipate negotiation scenarios, and (iv) *citizens and civil-society organizations* who seek interpretable signals of alignment and realignment in representative democracy [Janssen et al. 2012, Meijer 2013]. Concretely, the detected communities and temporal metrics can support decisions such as prioritizing which deputies or parties to interview when a coalition shifts, identifying "swing" blocs that are pivotal for agenda formation, monitoring abrupt changes in cohesion as potential governance risks, and guiding the design of transparency dashboards that highlight unusual cross-bloc voting.

To build a more concrete bridge between data and political reality, we treat roll-call votes as *institutionally consequential* outcomes (formal decisions recorded under procedural rules). The network edges encode behavioral similarity in these outcomes, not private preferences: they operationalize observable alignment given the institutional agenda, party discipline, and strategic voting. Therefore, our results are best interpreted as *data-driven signals* of coalition structure that should be combined with contextual knowledge (e.g., government/opposition status, leadership negotiations, and party switching) to support explanation and accountability, rather than as a complete substitute for qualitative political analysis [Bovens 2007, Scott 2014].

The primary challenge in analyzing Brazilian Congress voting records, however, lies in the inconsistency in voting behavior, which often masks the formation of clear ideological divides. Many propositions lack polarization, while weak edges in the voting network dilute the community detection process, complicating the identification of robust political alliances. This weakens the modularity of detected communities, making it difficult to accurately analyze true political dynamics.

This study proposes new network science methods to address these issues, building on established methodologies in the field [Newman 2018]. Our approach involves using the Leiden algorithm [Traag et al. 2019], enhanced by selective filtering of polarized propositions and edge pruning to improve community detection within the Brazilian political network. By focusing on polarized votes and optimizing edge pruning, we aim to improve network modularity, thereby uncovering distinct patterns of party behavior over the years [Brito et al. 2020]. Specifically, this study aims to:

1. (RQ1) How can polarized proposition selection and iterative edge pruning improve the *interpretability* and modularity of voting-based communities, supporting transparency-oriented analysis?
2. (RQ2) What coalitional communities emerge in Brazil's Chamber of Deputies from 2004 to 2023, and how do they relate to institutional events (e.g., government transitions and impeachment periods)?
3. (RQ3) How can temporal metrics (HHI, ARI, and Edge Volatility) quantify stability and structural change in ways that are actionable for decision support and accountability practices?

We offer four main contributions in this paper. First, we present a methodological framework for improving community detection in voting networks by focusing on polarized propositions. Second, we achieve significant improvements in modularity scores compared to previous approaches, demonstrating the effectiveness of our methodology. Third, we introduce quantitative temporal metrics (HHI, ARI, and Edge Volatility) to systematically measure community stability, similarity, and network structural changes over time. Fourth, we provide empirical evidence of the dynamic nature of political alliances in Brazil, offering new insights into party cohesion and ideological shifts through both qualitative analysis and quantitative temporal measurements.

This article is an **extended version** of our prior conference manuscript [da Silva et al. 2025], and it explicitly adds: (i) an automated search over polarization bounds and pruning levels to select yearly configurations, (ii) a longitudinal layer with three temporal metrics (HHI, ARI, and Edge Volatility), (iii) a procedure for temporal consistency of community labels, (iv) a baseline comparison against disparity-filter backbone extraction, and (v) an expanded IS-oriented discussion on transparency, accountability, governance, and interpretability (including the interaction between data quality, institutional context, and human judgment).

This paper is structured as follows: Section 2 discusses the theoretical foundations, emphasizing social network theories applied to the political context. Section 3 reviews related studies, focusing on previous analyses of voting networks and legislative community detection. In Section 4, we detail the methodology, including the application of the Leiden algorithm, edge pruning techniques to optimize network modularity, temporal metrics calculation (HHI, ARI, and Edge Volatility), and backbone comparison using the disparity filter method. Section 5 presents the results, analyzing the temporal evolution of political communities, the effects of polarization on party alliances, and quantitative temporal metrics findings. Finally, Section 6 summarizes the contributions, discusses the limitations of the study, and suggests directions for future research.

2. Theoretical Foundation

This research draws upon Social Network Theory, which is well-established in the study of interactions within large organizations and political systems [Newman 2018]. Social Network Theory provides a framework for analyzing the relationships among entities (e.g., parliamentarians), focusing on the structural properties of the network, such as cohesion, centrality, and community detection.

In the context of political systems, Social Network Theory has been widely used to examine voting patterns and alliances within legislative bodies. Previous studies have shown that voting networks can reveal ideological alignments and shifts in political behavior, making it an effective tool for understanding complex political interactions [Moody and Mucha 2013, Cherepnalkoski et al. 2016]. By modeling these networks, researchers can visualize the formation of coalitions, party cohesion, and the evolution of political alliances over time.

This research also leverages concepts from network science to analyze community detection within voting networks. Community detection is a fundamental aspect of Social Network Theory, aimed at identifying clusters or groups of nodes (e.g., deputies) that are more densely connected to each other than to other nodes in the network [Levorato and Frota 2017, Ferreira et al. 2018]. The identification of communities in political networks helps to reveal patterns of political behavior, such as ideological divides, cohesive voting blocks, and shifting alliances. In this study, we employ the Leiden algorithm, which is known for its efficiency and accuracy in detecting communities in large networks [Newman 2018]. This algorithm iteratively refines communities to maximize modularity, a measure of the strength of division within a network.

Another key concept integrated into this study is polarization, which has been central to understanding ideological alignment within political networks. Polarization refers to the degree to which nodes (deputies) in a network are divided into distinct ideological camps, often resulting in reduced interaction between opposing groups. Studies like [Cherepnalkoski et al. 2016] and [Maso et al. 2014] have demonstrated that voting similarity can be used to establish edges in the network, facilitating the analysis of polarization levels. In addition, [Levorato and Frota 2017] defined political polarization by clustering deputies based on their voting patterns, identifying ideological shifts over time and exploring the impact of these shifts on party cohesion.

The methodological framework of this study also incorporates techniques such as edge pruning and filtering of polarized propositions, aimed at improving the modularity and clarity of detected communities. This approach is supported by previous research, which has shown that removing weaker connections and focusing on polarized votes can enhance the identification of ideological divides and the visualization of party alliances [Faustino et al. 2019, Ferreira et al. 2018]. By applying these techniques to the Brazilian Congress voting network from 2004 to 2023, this study aims to provide a more detailed understanding of political dynamics and the temporal evolution of party behavior in Brazil.

To quantify and analyze the temporal dynamics of political communities, this study draws upon theoretical foundations from longitudinal network analysis and cluster validation. Temporal network analysis, as conceptualized in network science, focuses on understanding how network structures evolve over time, capturing both structural stability and change patterns [Newman 2018]. Model-based approaches, such as the Dynamic Degree Corrected Stochastic Block Model proposed by Wilson et al. [Wilson et al. 2019], provide probabilistic frameworks for detecting structural changes in temporal networks. While such models offer sophisticated mechanisms for change detection, our approach complements these methods by employing metric-based analysis that provides interpretable quantitative measures of temporal dynamics. The Herfindahl-Hirschman Index (HHI), originally developed in economics to measure market concentration, is adapted here to assess the concentration of deputies within communities, providing insights into power distribution and community consolidation over time. The Adjusted Rand Index (ARI), a well-established metric in cluster validation and pattern recognition, measures the similarity between community partitions across time periods, correcting for chance agreement and enabling robust comparisons of community stability [Hubert and Arabie 1985]. Edge Volatility extends concepts from dynamic network analysis, quantifying the structural changes in network connections between time periods, which reflects the fluidity of political alliances and voting alignments.

The backbone comparison component of this study is grounded in statistical network analysis and null model theory. The disparity filter method [Serrano et al. 2009] employs statistical significance testing to identify edges that are statistically significant given the local context of each node, based on the null hypothesis that edge weights are uniformly distributed. This approach aligns with broader theoretical frameworks in network science that seek to distinguish meaningful structural patterns from random configurations, providing a principled method for network simplification while preserving essential topological features. The comparison between our proposed method (polarized filtering with edge pruning) and the disparity filter backbone extraction enables validation through established statistical network analysis principles, ensuring that improvements in modularity and community detection reflect genuine structural insights rather than methodological artifacts.

Overall, this study integrates theoretical concepts from Social Network Theory with empirical methods from network science, longitudinal analysis, and statistical network analysis to offer a novel approach to analyzing political networks. By focusing on community detection, modularity optimization, polarization, temporal dynamics quantification, and statistical validation, this research demonstrates how established network theories and temporal analysis frameworks can be applied to complex political systems, providing both qualitative insights and quantitative measures of political alliance evolution.

3. Related Work

Several studies have utilized parliamentary voting data to create political networks, contributing to a better understanding of political dynamics [Maso et al. 2014, Cherepnalkoski et al. 2016, Levorato and Frota 2017, Ferreira et al. 2018]. These works focus on detecting communities within the network and explore additional concepts, such as party cohesion, polarization, and ideological shifts, which help in comprehending the structure of political networks.

Cherepnalkoski et al. [Cherepnalkoski et al. 2016] applied network analysis to study parliamentary dynamics in the European context by combining voting similarity with social media data. Their findings revealed political coalitions, demonstrated collaboration between opposing groups, and highlighted the consistency between voting patterns and social media alignments. Dal Maso et al. [Maso et al. 2014] conducted a similar study in Italy, analyzing voting similarity across three categories: in favor, against, and abstention. Their study emphasized party cohesion and identified parliamentarians whose voting behaviors were more aligned with parties other than their own.

In Brazil, Levorato and Frota [Levorato and Frota 2017] examined propositions from 2011 to 2016, reporting a decrease in network polarization over time and highlighting shifting alignments between centrist and right-wing parties. Ferreira et al. [Ferreira et al. 2018] extended this analysis from 2004 to 2017 by introducing the concept of Party Discipline, showing that community-based clustering was stronger than clustering based solely on party affiliation. In a subsequent work, Ferreira et al. [Ferreira et al. 2019] proposed methods for modeling dynamic ideological behavior in political networks, focusing on capturing temporal shifts in political alignments and ideological positions. Faustino et al. [Faustino et al. 2019] analyzed shifts in party affiliations by constructing networks based on deputy migrations. Studies in the United States, such as Bryden and Silverman [Bryden and Silverman 2019] and Melo et al. [Melo et al. 2018], explored the impact of user behavior and resource influence on voting, respectively.

Beyond political networks, research on temporal network analysis has advanced methods for detecting and modeling change in dynamic networks. Wilson et al. [Wilson et al. 2019] introduced a Dynamic Degree Corrected Stochastic Block Model (DDC-SBM) for modeling and detecting change in temporal networks, providing a probabilistic framework for identifying structural changes over time. While their approach focuses on model-based change detection, our methodology employs metric-based temporal analysis to quantify community stability and network evolution.

Prior studies leveraging voting data to form political networks have enhanced our understanding of legislative dynamics, but they often fall short in clearly distinguishing ideological divides. Methods like those by Dal Maso et al. [Maso et al. 2014], which examine voting similarity using broad classifications, face challenges in capturing cohesive political alliances, particularly in multiparty systems with fluid coalitions. Ferreira et al. [Ferreira et al. 2018], while introducing concepts like Party Discipline, noted that community detection can be weakened by the inclusion of less ideologically distinct propositions, resulting in less cohesive clusters. Brito et al. [Brito et al. 2020], who used voting agreement to identify communities, provided valuable insights but lacked network refinement techniques to eliminate weaker connections that dilute modularity.

Our approach overcomes these limitations by implementing the Leiden algorithm in conjunction with selective edge pruning and filtering for polarized propositions. This combination enhances the clarity of detected communities and improves modularity, ensuring that the resulting communities reflect stronger ideological alignments. By optimizing network structure through focused proposition selection and edge trimming, our methodology not only provides a more precise view of political alliances but also reveals temporal dynamics within a complex multiparty landscape, advancing the understanding of legislative behavior beyond prior methodologies.

4. Proposed Methodology

The methodology for this study follows a structured sequence of automated operations designed to optimize community detection within the Brazilian Congress voting network, as illustrated in Figure 1. The code runs automatically, requiring only the input data, and provides the final results at the end of the process. The process begins with Data Acquisition, where voting records are collected, followed by Data Pre-processing to organize and standardize the data for network analysis.

Next, an iterative loop is employed, comprising four main steps: Polarized Proposition Selection, Network Construction, Edge Pruning, and Community Detection. The selection of polarized propositions is an integral part of this loop, refining the dataset to highlight significant ideological divides. This refined dataset then supports the construction of a voting similarity network. Through edge pruning, weaker connections are incrementally removed to strengthen the modularity and clarify community boundaries. Community detection, performed using the Leiden algorithm, evaluates the resulting modularity to determine the effectiveness of each iteration. The loop repeats automatically until the configuration that achieves the highest modularity with the fewest communities is identified, delivering the final results without further user intervention.

Because polarization bounds and pruning levels directly affect interpretability, we explicitly treat them as *decision parameters* in the pipeline. Polarization bounds are explored in progressively stricter intervals (e.g., 0–100%, 10–90%, 20–80%), which operationalizes the trade-off between *signal* (more ideologically divisive propositions) and *coverage* (enough propositions to avoid fragmentation). Edge pruning is explored from 0% to 98% in 2% increments, which balances computational cost with a sufficiently fine-grained search over network sparsification. Selecting the yearly configuration that maximizes modularity while minimizing the number of communities favors partitions that are both structurally strong and easier to communicate to non-technical stakeholders; however, this choice can trade off against detecting smaller, niche communities that may be politically meaningful in specific institutional contexts.

After identifying the best model for each year, we compute temporal metrics to evaluate community stability and evolution over time. These metrics include the Herfindahl-Hirschman Index (HHI) for measuring community concentration, the Adjusted Rand Index (ARI) for assessing similarity between consecutive years, and Edge Volatility for quantifying structural stability. Following the computation of temporal metrics, we ensure temporal consistency of community labels across years, enabling meaningful longitudinal analysis. Finally, we perform a backbone comparison using the disparity filter method [Serrano et al. 2009] to provide a baseline for evaluating the effectiveness of our proposed approach.

The dataset used in this study comprises voting records from the Brazilian Chamber of Deputies, spanning from 2004 to 2023. This data was obtained from publicly avail-

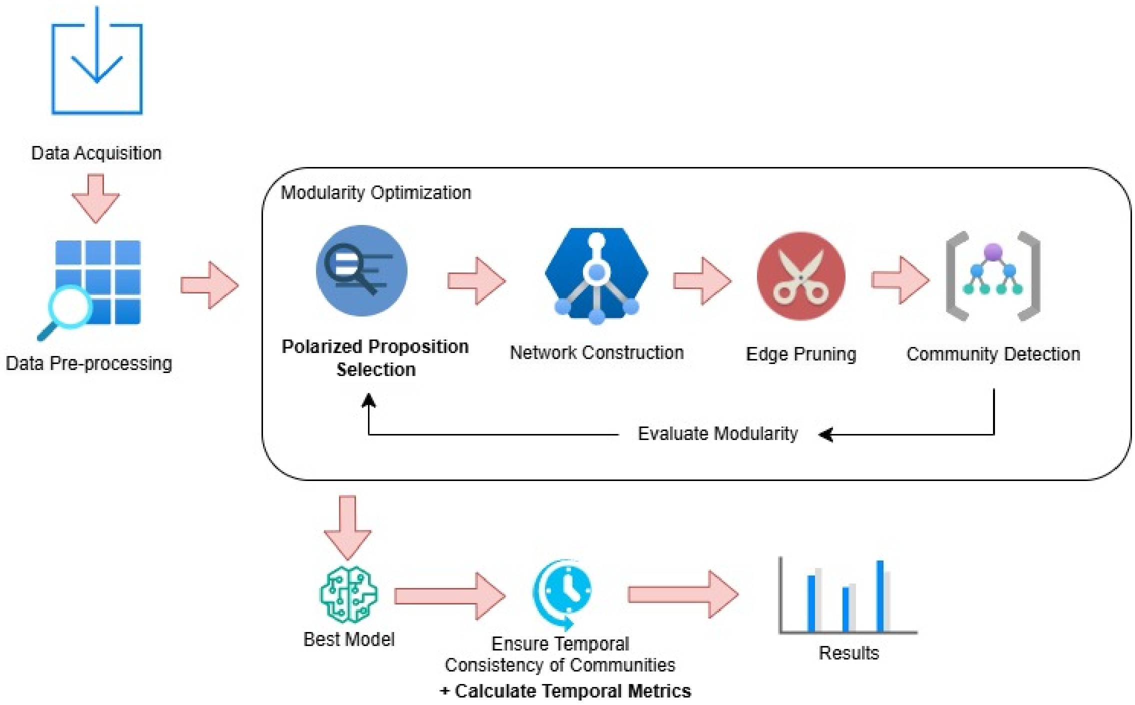


Figure 1. Flowchart of the proposed methodology

able sources[1], ensuring transparency and replicability in our analysis. The voting records cover 1,579,874 votes across 4,894 unique propositions, providing a comprehensive basis for examining political alliances and shifts within the Brazilian legislative context.

The dataset features a wide range of political parties, with prominent representation from major parties such as the Workers' Party (PT), the Progressives (PP), and the Brazilian Social Democracy Party (PSDB). This diversity of party representation enables a robust analysis of Brazil's multiparty system and its impact on political alliances.

Further details on the dataset, including annual voting trends and the distribution of votes across propositions, will be discussed in Section 5.1. The comprehensive scope of this dataset enables a longitudinal analysis of political dynamics and enhances the reliability of community detection across different legislative periods.

4.1. Data Pre-Processing

The data pre-processing phase ensures that the dataset is clean, consistent, and ready for network analysis. The process begins with renaming columns in the dataset using a predefined mapping. This step standardizes column names across multiple datasets, ensuring consistency and facilitating seamless merging during subsequent steps.

To consolidate information, datasets are merged on common columns using a systematic approach. This integration brings together all relevant data into a single unified structure, which is essential for comprehensive analysis. Following this, rows with null

[1] `https://basedosdados.org/dataset/3d388daa-2d20-49eb-8f55-6c561bef26b6?table=7a5bb339-fd52-4376-93a6-fa0807981fc6`

values in the "aprovacao" column are filtered out. This ensures that only valid voting records are retained, eliminating incomplete or unreliable data that could skew results.

Another critical step involves calculating the percentage of "Yes" votes for each proposition. A new column, "yes_vote_percentage" is added to the dataset, representing the proportion of "Yes" votes relative to the total votes cast, including "Yes", "No" and other categories. This metric is vital for identifying polarized propositions and understanding voting dynamics.

Finally, a filtering process is applied to focus on polarized propositions. Only those propositions where the percentage of "Yes" votes falls within predefined lower and upper bounds are retained. This step refines the dataset to highlight significant ideological divides, removing less informative data and enhancing the clarity of subsequent network analyses.

4.2. Polarized Proposition Selection

We introduce Polarized Proposition Selection as a key contribution of this study to enhance the effectiveness of community detection. In this step, we refine the dataset by selecting propositions that exhibit a higher level of ideological division. We filter propositions based on the percentage of "Yes" votes, ensuring that only those with significant polarization are included. Rather than analyzing all propositions indiscriminately, we progressively apply stricter filters, narrowing the range of "Yes" votes to thresholds such as 0-100%, 10-90%, 20-80%, and so on.

At each filtering level, we evaluate the modularity of the resulting communities to determine the threshold that maximizes modularity, a critical measure of community quality in network analysis [Blondel et al. 2008]. Through this iterative filtering process, we enable the network to capture ideological divides more effectively by emphasizing polarized propositions that reflect meaningful voting distinctions. By focusing on propositions with a balanced distribution of votes, we improve the clarity of detected communities and uncover significant political alliances within the Brazilian Congress.

4.3. Network Construction

The generation of the network graphs in this study was based on the voting data for each year. The resulting network is a weighted undirected graph where vertices represent deputies and edges represent voting similarity relationships between pairs of deputies, with edge weights quantifying the level of agreement in their voting patterns. The process involved several key steps, as detailed below:

We encode votes as +1 (Yes) and -1 (No). Abstentions, obstructions, leadership releases, and absences are treated as missing and excluded from pairwise intersections; thus N_{ij} counts only propositions with Yes/No for both deputies. The filtered data was then transformed into a pivot table, where each row represented a deputy and each column represented a vote. This table served as the basis for calculating the adjacency matrix, representing the similarity of voting patterns between deputies.

The pairwise level of agreement between two deputies i and j is defined as:

$$w_{ij} = \frac{1}{N_{ij}} \sum_{k=1}^{N_{ij}} v_i^{(k)} v_j^{(k)} = \frac{\mathbf{v}^{(i)} \cdot \mathbf{v}^{(j)}}{N_{ij}}, \tag{1}$$

where $\mathbf{v}^{(i)} \cdot \mathbf{v}^{(j)}$ is the dot product between the voting vectors of deputies i and j, and N_{ij} denotes the number of propositions for which both deputies cast a recorded Yes/No vote (the pairwise intersection). Edges are defined only for $N_{ij} \geq 1$.

This approach ensures that weights reflect the similarity in voting behavior, with stronger edges indicating higher similarity.

4.3.1. Normalization of Weights

To ensure comparability of edge weights across different graphs, the weights were normalized to the range [0, 1] using min-max normalization. The normalization formula is:

$$\tilde{w}_{ij} = \frac{\mathbf{v}^{(i)} \cdot \mathbf{v}^{(j)} - w_{\min}}{w_{\max} - w_{\min}}, \tag{2}$$

where $w_{\max} = N_{ij}$ (maximum possible agreement when all votes are identical) and $w_{\min} = -N_{ij}$ (maximum possible disagreement when all votes are opposite). This simplifies to:

$$\tilde{w}_{ij} = \frac{\mathbf{v}^{(i)} \cdot \mathbf{v}^{(j)} + N_{ij}}{2N_{ij}}, \tag{3}$$

where this formula is defined only if $N_{ij} \geq 1$. Using N_{ij} instead of the total number of propositions N ensures that pairs with sparse overlap are not over-shrunk and pairs with dense overlap are not over-inflated.

This normalization ensures that the weights range from 0 (complete disagreement) to 1 (complete agreement), where $\tilde{w}_{ij} = 0.5$ indicates neutral similarity (zero dot product). No connection refers to cases where there is genuinely no edge (i.e., when $N_{ij} = 0$ or after post-threshold removal). The final network structure consists of deputies as vertices connected by weighted edges, where each edge weight $\tilde{w}_{ij}$ represents the normalized voting similarity between deputies i and j.

4.4. Edge Pruning

Edge pruning optimizes the network structure by systematically removing weaker connections, focusing on edges with the lowest weights. A key challenge in this process is ensuring that the pruning step enhances the interpretability of political alliances without artificially merging distinct communities. Instead of applying a fixed pruning threshold, we conducted an iterative evaluation of different thresholds to ensure that only weak or spurious connections were removed. This approach prevents artificial community merging by preserving strongly cohesive subgroups, thus maintaining the integrity of the network structure.

Starting with 0% removal, we progressively eliminate up to 98% of the weakest edges in 2% increments, prioritizing the weakest weights at each step. This approach concentrates the analysis on stronger and more significant voting relationships, enhancing network modularity and making political alliances within the network more distinct [Zhou et al. 2012]. By gradually pruning weaker edges, the method ensures that the resulting network reflects only the most cohesive and meaningful connections, enabling a clearer identification of ideological divides and community structures.

At each level of pruning, we assessed both the modularity and the number of detected communities, ultimately selecting the pruning threshold that maximized modularity while minimizing the number of communities. This balance prevents fragmentation of the network and ensures that detected communities represent meaningful political divisions.

To maintain network integrity, we retained the strongest edge of a node if its removal would disconnect that node from the network. This step preserved the representation of deputies' voting patterns, even for those with relatively low similarity to others, thus maintaining the accuracy of the analysis regarding peripheral or swing actors in political alliances. By balancing cohesion and interpretability, edge pruning allows for a focused representation of significant alliances, ensuring the network remains a robust reflection of voting behaviors within the Brazilian Congress.

4.5. Community Detection

Community detection is a central aspect of this research, aiming to identify clusters of deputies with similar voting patterns within the Brazilian Congress. In this context, communities are defined as groups of deputies who vote similarly across a range of legislative propositions, reflecting potential political alliances or ideological alignment.

To evaluate the quality of the detected communities, this study uses the **modularity** metric. Modularity measures the number of edges within communities that exceed the expected number in a random model, providing a quantifiable metric for assessing the strength of division within the network [Newman 2018, Blondel et al. 2008]. However, we acknowledge that modularity alone does not necessarily imply an accurate representation of real-world political alliances, as it is a structural measure that prioritizes network partitioning based on connectivity patterns rather than external political factors.

The modularity Q of a network is given by:

$$Q = \frac{1}{2m} \sum_{ij} \left[a_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j), \tag{4}$$

where:

- a_{ij} is 1 if nodes i and j are connected, and 0 otherwise,
- k_i and k_j are the degrees of nodes i and j, respectively,

- m is the total number of edges in the network,
- $\delta(c_i, c_j)$ is 1 if nodes i and j belong to the same community, and 0 otherwise.

This study employs the **Leiden algorithm** for community detection, which is an improved version of the Louvain algorithm [Traag et al. 2019]. The Leiden algorithm iteratively refines the community structure to maximize modularity, ensuring that communities are cohesive and well-separated. It starts by considering each node as an individual community and evaluates the gain in modularity when moving nodes between communities. Nodes are reassigned to communities that result in the highest increase in modularity. If moving a node results in a decrease in modularity, it remains in its original community. This process continues until no further gains in modularity can be achieved.

While modularity is a widely used metric for evaluating community structure, it does not inherently capture the underlying political meaning of the detected communities. Thus, although a high modularity score suggests well-separated groups, it does not confirm that these groups correspond to known political alliances or historical voting blocs. Future studies could explore external validation techniques to assess how well the detected communities align with real-world political structures.

The use of the Leiden algorithm in this research is motivated by its ability to identify well-connected communities and improve the clarity of political alliances detected in the network. By maximizing modularity, the algorithm ensures that the identified communities reflect structural voting similarities within the Brazilian Congress, providing insights into the dynamics of political blocs based on network connectivity patterns.

4.6. Temporal Consistency of Communities

When detecting communities year by year, a significant challenge arose: the labels assigned to communities by the *Leiden algorithm* changed with each run, even when the community structures themselves remained largely similar. This inconsistency in labeling made it difficult to determine whether a community from one year corresponded to the same or a similar community in another year. A similar issue is addressed in *OLCPM (Online Label Propagation and Clique Percolation Method)*, proposed by Boudebza et al. [Boudebza et al. 2022], which introduces a dynamic approach to tracking communities over time by leveraging *label propagation techniques*. Related work on modeling dynamic ideological behavior in political networks [Ferreira et al. 2019] has also addressed the challenge of tracking community evolution over time, emphasizing the importance of maintaining temporal consistency when analyzing political alignments.

To address this issue in our study, we implemented a correction method inspired by ideas from OLCPM, but adapted to our specific needs. Rather than relying on *continuous label propagation*, we adopt a *reference-based alignment approach*:

1. We select a reference year and identify the largest community within that year to serve as an anchor.
2. For each subsequent and previous year, we compare community compositions and identify the group with the **highest overlap** in party composition relative to the reference.
3. Labels are then adjusted to **ensure consistency** across the entire timeline.

While OLCPM focuses on maintaining dynamic coherence using an online update mechanism, our approach ensures **stability in community identification** by enforcing **explicit label alignment** rather than relying on implicit label propagation. This methodology is particularly crucial for accurately tracking *political alliances*, where stability in group identification is key to understanding long-term trends in voting behavior.

By maintaining **consistent labeling**, we effectively analyze temporal dynamics and gain deeper insights into the behavior of political parties and their deputies over time, facilitating a more comprehensive understanding of the Brazilian Congress's political landscape.

4.7. Temporal Metrics

To quantitatively assess the stability and evolution of detected communities over time, we compute three key temporal metrics: the Herfindahl-Hirschman Index (HHI), the Adjusted Rand Index (ARI), and Edge Volatility. These metrics provide complementary perspectives on how political communities change across different legislative periods. While model-based approaches, such as the Dynamic Degree Corrected Stochastic Block Model [Wilson et al. 2019], offer probabilistic frameworks for detecting structural changes in temporal networks, our metric-based approach provides interpretable quantitative measures that directly quantify community concentration, similarity, and structural stability. This complements existing temporal network analysis methods by offering a transparent and computationally efficient alternative for understanding network evolution.

4.7.1. Herfindahl-Hirschman Index (HHI)

The Herfindahl-Hirschman Index measures the concentration of community sizes within each year. It quantifies how evenly distributed deputies are across communities, with higher values indicating more concentrated (uneven) distributions and lower values indicating more balanced distributions. The HHI is calculated as:

$$HHI = \sum_{c=1}^{C} \left(\frac{|c|}{N} \right)^2, \tag{5}$$

where $|c|$ is the size of community c, C is the total number of communities, and N is the total number of deputies in the network. The HHI ranges from $1/C$ (most balanced distribution) to 1 (all deputies in a single community). This metric helps identify periods when political power is concentrated in fewer communities versus when it is more evenly distributed across multiple blocs.

4.7.2. Adjusted Rand Index (ARI)

The Adjusted Rand Index measures the similarity of community structures between consecutive years, accounting for the expected similarity under random assignment. ARI is calculated using only overlapping deputies present in both years, ensuring that comparisons are based on deputies who remained in office across the transition. The ARI ranges from -1 to 1, where:

- ARI = 1 indicates identical community structures
- ARI = 0 indicates similarity no better than random assignment
- ARI $<$ 0 indicates worse than random agreement

This metric enables us to identify periods of political stability (high ARI) versus periods of significant realignment (low ARI), providing insights into when major shifts in political alliances occurred within the Brazilian Congress.

4.7.3. Edge Volatility

Edge Volatility measures the structural stability of the network by quantifying the percentage of edges that change their status between communities across consecutive years. Specifically, it tracks whether edges that were previously *intra-community* (connecting deputies within the same community) become *inter-community* (connecting deputies across different communities), or vice versa. The volatility is calculated as:

$$Volatility_{t,t+1} = 100 \times \frac{NumberOfEdgesWithChangedStatus}{NumberOfCommonEdges}, \tag{6}$$

where only edges present in both years t and $t+1$ are considered. High volatility indicates significant structural changes in the network, suggesting shifts in voting patterns and political relationships. Low volatility, conversely, indicates greater stability in the underlying network structure, suggesting consistent voting behavior and maintained alliances.

Together, these three metrics provide a comprehensive view of community dynamics: HHI captures concentration patterns, ARI measures structural similarity, and Edge Volatility quantifies network stability. This multi-dimensional analysis enables a deeper understanding of how political communities evolve over time within the Brazilian Congress.

4.8. Backbone Comparison

To evaluate the effectiveness of our proposed method (polarized proposition filtering combined with iterative edge pruning), we compare our results against a baseline method using backbone extraction. The backbone extraction approach employs the **disparity filter** proposed by Serrano et al. [Serrano et al. 2009], which extracts statistically significant edges from weighted networks based on the strength of connections relative to their local context.

The disparity filter method works by:

1. For each node, calculating the normalized weight of each edge relative to the node's total strength

2. Determining the statistical significance of each edge based on its probability under a null model
3. Retaining only edges that are statistically significant at a given significance level α
4. Adjusting α automatically to maintain network connectivity and achieve a target edge reduction (approximately 10% of original edges while preserving at least 80% of nodes in the giant component)

After applying the disparity filter, we construct the network backbone and perform community detection using the Leiden algorithm, following the same procedure as our proposed method. This allows for a direct comparison of modularity scores, number of detected communities, and the quality of political alliance detection between the two approaches.

The comparison enables us to:

- Assess whether polarized filtering and iterative edge pruning provide advantages over statistical backbone extraction
- Evaluate the trade-offs between the two methods in terms of modularity and community detection quality
- Validate the effectiveness of our proposed approach in capturing meaningful political alliances

Results from both methods are saved separately, allowing for quantitative comparison and statistical analysis of their relative performance across the entire study period (2004-2023).

4.9. Data Quality, Interpretation, and Institutional Context

While roll-call votes are reliable *records of formal outcomes*, the link between voting similarity and political meaning depends on data quality and institutional context. In our preprocessing and modeling choices, we explicitly acknowledge the following interactions:

- **Missingness and procedural categories**: abstentions, absences, leadership releases, and obstructions are treated as missing to avoid forcing similarity where there is no comparable Yes/No choice. This improves measurement validity but can reduce overlap N_{ij} for some pairs, increasing uncertainty for peripheral deputies.
- **Agenda and institutional incentives**: communities reflect alignment *given what is voted on* in each year (agenda control, coalition agreements, and strategic coordination). Therefore, shifts in ARI/volatility can indicate institutional realignments, but interpretation should consider contemporaneous events and government/opposition configurations [Scott 2014].
- **Human interpretation as a complement**: the pipeline is best used to generate interpretable indicators and *audit trails* (e.g., when cohesion drops sharply), which stakeholders can triangulate with qualitative sources such as news coverage, party leadership statements, and legislative negotiation records [Meijer 2013, Bovens 2007].

- **Trade-offs in model transparency**: modularity-optimized partitions simplify communication but may hide smaller sub-blocs; conversely, keeping more edges/propositions can preserve nuance but reduces clarity. We address this by reporting temporal metrics and by comparing against a backbone baseline.

5. Results and Discussion

In this section, we present the findings of our analysis on the Brazilian Congress voting network from 2004 to 2023. We explore the effects of polarized proposition selection and edge pruning on community detection, evaluate the modularity improvements, and discuss the temporal evolution of political alliances.

The section is organized as follows. We begin with a **Dataset Overview** that describes the voting records and their distribution across parties and years. Next, we present a **Polarization Analysis** examining how different polarization thresholds affect modularity, followed by an **Edge Pruning Optimization** section that explores the relationship between pruning percentages and community detection quality. We then report the **Best Results** achieved by combining optimal polarization bounds and pruning percentages. The **Community Detection Results** section analyzes the number of members in each community over time, revealing the structure and dynamics of political blocs. We then examine the **Temporal Evolution of Communities**, tracking how party compositions shift within communities across the 20-year period, with particular attention to changes following significant political events such as the 2016 impeachment. The **Modularity Improvement** section compares our methodology against baseline approaches, demonstrating a 10.95% improvement over backbone extraction methods. Finally, we present a comprehensive **Temporal Metrics Analysis** using the Herfindahl-Hirschman Index (HHI), Adjusted Rand Index (ARI), and Edge Volatility to quantitatively assess community concentration, stability, and structural changes over time, providing a multi-dimensional view of political alliance dynamics in the Brazilian Congress.

All experiments in this section were conducted on an Acer Aspire Nitro 5 AN515-55-705U notebook, equipped with an Intel Core i7-10750H processor, an NVIDIA GTX 1650 Ti GPU, and 16GB of RAM, running Windows 10. The implementation utilized Python version 3.9-slim, with several key libraries supporting the analysis, including 'leidenalg' (version 0.10.0) for community detection, 'basedosdados' for data acquisition[2], 'networkx' (version 2.8.8) for network construction and manipulation[NetworkX Developers 2024], 'numpy' (version 1.21.4) for numerical operations, and 'pandas' (version 1.3.4) for data processing and management. These tools provided the computational framework necessary for implementing the methodologies and conducting the experiments efficiently. The repository is available anonymously at the following link: `https://github.com/pedronatanaelfs/depcom`.

Beyond reporting metrics, we interpret the outputs as an IS artifact for transparency and decision support. In practice, yearly communities and their temporal evolution can be used to (i) *explain* coalitional structure to the public in a traceable way

[2] `https://basedosdados.org/dataset/3d388daa-2d20-49eb-8f55-6c561bef26b6?table=7a5bb339-fd52-4376-93a6-fa0807981fc6`

(transparency), (ii) *flag* periods where accountability questions are more pressing (e.g., abrupt changes in cohesion or volatility), and (iii) *support decisions* about where to focus journalistic investigation or institutional monitoring (e.g., on pivotal swing blocs or on years with low ARI and high volatility).

5.1. Dataset Overview

The voting data analyzed is derived from the Brazilian Chamber of Deputies, covering the period from 2004 to 2023. The dataset consists of a total of 1,579,874 votes, corresponding to 4,894 unique propositions.

Table 1 shows the most frequent political parties appearing in the dataset, including the Workers' Party (PT), the Progressives (PP), the Brazilian Social Democracy Party (PSDB), the Liberal Party (PL), and the Social Democratic Party (PSD), among others. Specifically, the Workers' Party (PT) registered the highest number of votes, totaling 204,583.

Table 1. Most frequent political parties and their number of votes in the dataset

Political Party	Number of Votes
PT	204,583
PP	128,404
PSDB	113,733
PL	101,599
PSD	96,892
PMDB	92,356
PSB	89,623
PDT	72,652
PSL	71,838
DEM	67,249

Figure 2 shows the distribution of votes and propositions per year from 2004 to 2023. The data reveals a substantial increase in legislative activity over time, with peaks in 2015, 2019-2021, indicating periods of heightened legislative activity in the Brazilian Congress.

5.2. Polarization Analysis

To improve community detection, we focused on selecting polarized propositions, which are votes where deputies are more evenly divided, revealing clearer ideological differences.

Although the method automatically selects optimal polarization thresholds, Figure 3 presents a detailed study of how different polarization limits affect modularity. We evaluated different polarization thresholds by varying the lower and upper bounds of the percentage of "Yes" votes for each proposition. The **upper bound** represents propositions with the highest percentage of "Yes" votes, such as cases where 80% of the votes

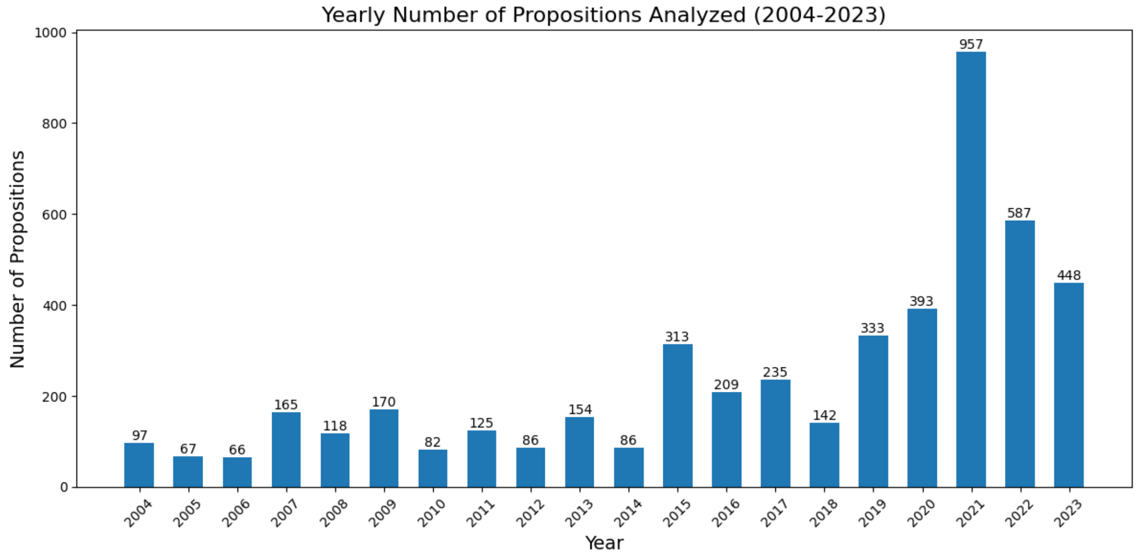


Figure 2. Yearly Number of Votes and Unique Propositions Analyzed (2004-2023).

were "Yes." Conversely, the **lower bound** corresponds to propositions with the lowest percentage of "Yes" votes, such as those where only 20% of the votes were "Yes." For each threshold, we calculated the modularity of the resulting communities.

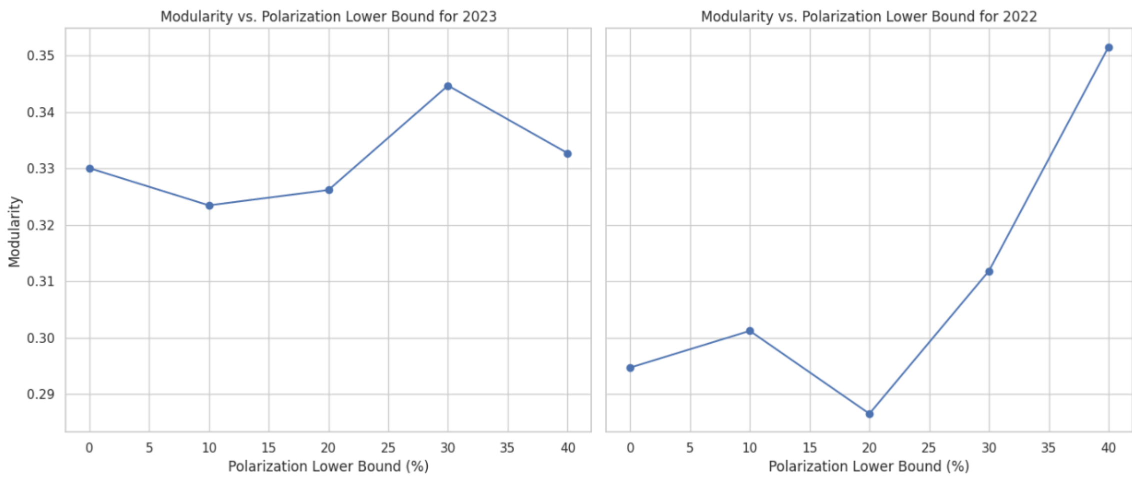


Figure 3. Modularity vs. Polarization Lower Bound for 2023 and 2022.

Figure 3 illustrates how modularity changes with different polarization lower bounds for 2023 and 2022. The results indicate that filtering polarized propositions strengthens community detection by highlighting clearer ideological divides, though excessive filtering can fragment communities when too few propositions remain. The different patterns between 2022 and 2023 suggest varying levels of ideological division across different legislative periods.

5.3. Edge Pruning Optimization

Edge pruning is a technique used to enhance community detection by systematically removing weaker connections from the network. By focusing on the most significant edges that represent strong voting alignments, this process refines the network's structure and highlights cohesive communities more effectively.

To identify the optimal pruning level, we evaluated how different pruning percentages impact both modularity and the number of detected communities. The ideal pruning percentage is selected where modularity reaches its peak while maintaining a minimal number of communities, thus avoiding excessive fragmentation.

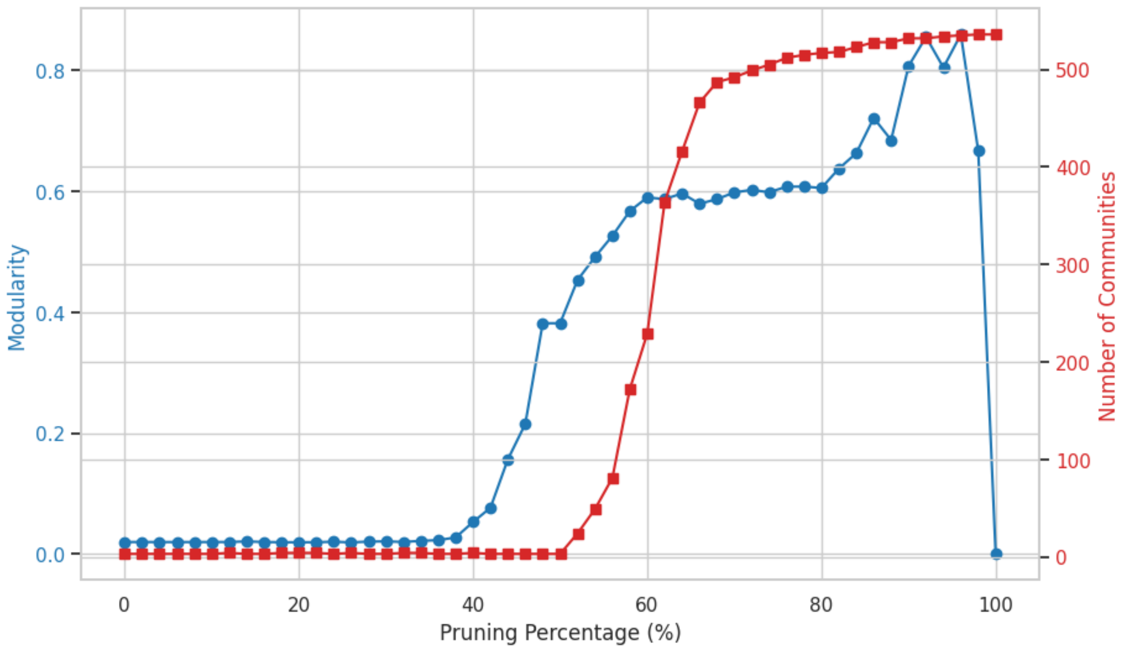


Figure 4. Modularity and Number of Communities vs. Pruning Percentage for 2023.

Figure 4 demonstrates the relationship between pruning percentage, modularity, and the number of communities for 2023. The results show that moderate pruning improves modularity by removing weaker edges and clarifying community boundaries, but excessive pruning fragments the network and reduces modularity. The optimal pruning level balances cohesive communities with minimal fragmentation.

5.4. Best Results

Table 2 presents the optimal polarization bounds, pruning percentages, number of communities, and modularity values for each year, resulting from the iterative optimization process described in the methodology.

5.5. Community Detection Results

Using the optimal polarization thresholds and pruning percentages, we applied the Leiden algorithm iteratively within the optimization loop to detect communities in the voting network. The analysis reveals a consistent structure across the study period: two dominant

Table 2. Optimal polarization bounds, pruning percentages, number of communities, and modularity values for each year

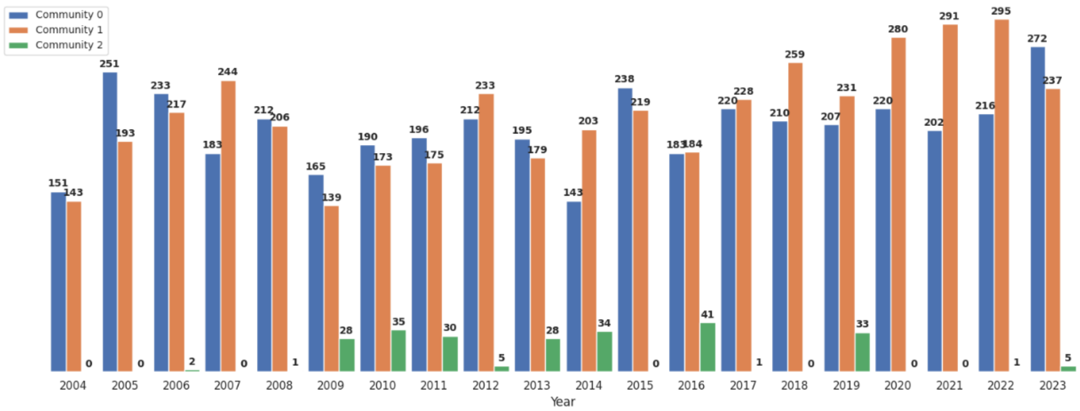
Year	Prun. (%)	Comm.	Modul.	Low Bd. (%)	Up Bd. (%)
2004	26	2	0.49	40	60
2005	42	2	0.35	30	70
2006	46	3	0.40	40	60
2007	38	2	0.39	40	60
2008	40	3	0.36	40	60
2009	42	3	0.29	40	60
2010	44	3	0.23	30	70
2011	42	3	0.34	40	60
2012	42	3	0.32	40	60
2013	42	3	0.31	40	60
2014	44	3	0.29	40	60
2015	44	2	0.29	40	60
2016	46	3	0.27	40	60
2017	46	3	0.32	40	60
2018	42	2	0.19	30	70
2019	42	3	0.36	40	60
2020	48	2	0.36	40	60
2021	42	2	0.34	40	60
2022	44	3	0.28	40	60
2023	48	3	0.38	40	60

ideological blocs representing distinct political alignments, with a third smaller community acting as a swing faction that can influence legislative outcomes. Community 0 is consistently dominated by the Workers' Party (PT), while Community 1 is primarily composed of parties such as PSDB, PP, PL, and PSD, with significant realignments following the 2016 impeachment of Dilma Rousseff.

Figure 5 shows the number of members in each community over time, illustrating the relative sizes and stability of the three communities.

5.6. Temporal Evolution of Communities

Figures 6, 7, and 8 illustrate the temporal evolution of community composition, revealing three key periods of transformation. Prior to 2016, Community 0 maintained strong PT dominance, while Community 1 was primarily led by PSDB, and Community 2 served as a cohesive swing bloc with significant representation from PSDB, PSB, PFL, and PDT. The 2016 impeachment period marked a major realignment, with Community 1 showing increased prominence of PP, PSD, and PL (parties that supported the impeachment)



s

Figure 5. Number of deputies from the top 5 parties in Communities 0, 1, and 2 over time (2004-2023).

reflecting the consolidation of anti-PT alliances. During the Bolsonaro administration (2019-2022), Community 1 continued to strengthen with these parties, while Community 2's representation significantly diminished, suggesting a shift toward a more polarized two-bloc structure.

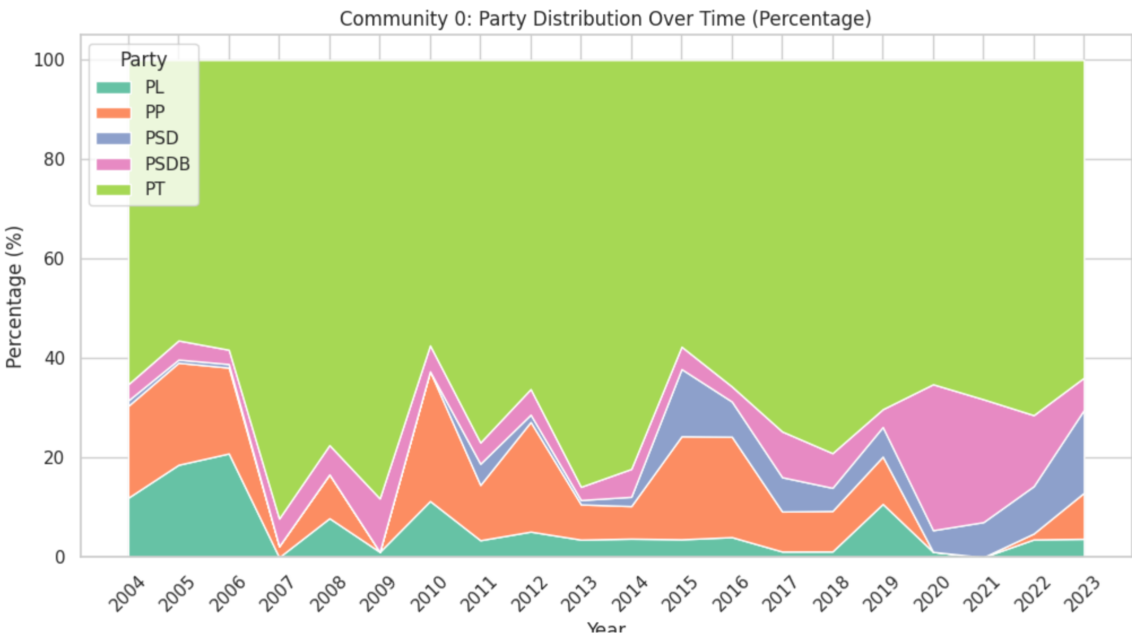


Figure 6. Temporal distribution of the top 5 parties in Community 0, showing the percentage of deputies from each party from 2004 to 2023.

5.7. Modularity Improvement

Our methodology aimed to enhance modularity by using polarized proposition selection and edge pruning, optimizing the network to better reveal cohesive communities.

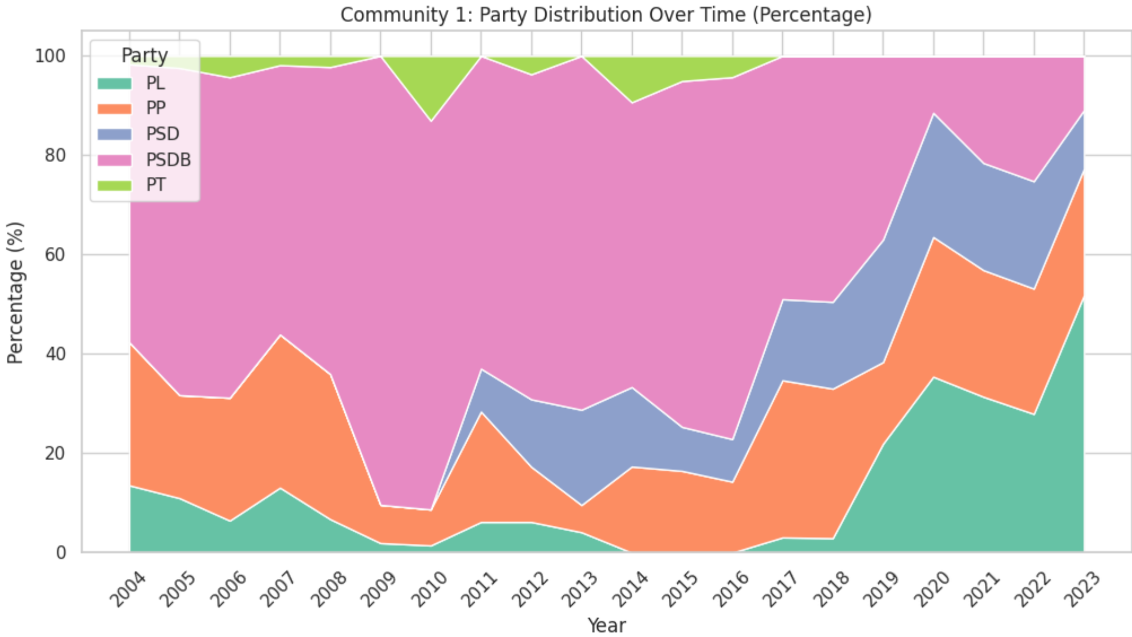


Figure 7. Temporal distribution of the top 5 parties in Community 1, illustrating the changing composition of party representation from 2004 to 2023.

5.7.1. Comparison with Previous Approaches

Figure 9 compares the modularity results of our approach with the ”No Filter” approach and the ”Backbone Extraction” method used by [Brito et al. 2020]. Our methodology achieved a mean modularity of 0.327, providing a 10.95% improvement over Backbone Extraction (0.295) and significantly outperforming the No Filter approach (0.008). Our method was the best-performing approach in 14 years, compared to 6 years for Backbone Extraction, with a lower standard deviation of 30.91%, indicating more consistent results.

Higher modularity implies the formation of more cohesive communities within the network, which strengthens the reliability of findings related to political alliances. The increase in modularity achieved through our method allows a clearer identification of political blocs, facilitating a more precise analysis of legislative behavior and enabling better predictions of voting outcomes. With improved modularity, the detected communities reflect stronger internal connections, providing valuable insights into party cohesion and ideological alignments within Congress.

From a societal standpoint, clearer blocs improve the *communicability* of complex legislative behavior: stakeholders can compare years, identify when coalitions consolidate or fragment, and translate these signals into concrete accountability questions (e.g., Which deputies changed alignment during a high-volatility transition, and under which institutional incentives or negotiations did that occur?).

5.8. Temporal Metrics Analysis

To quantitatively assess the stability and evolution of detected communities over time, we computed three key temporal metrics across all year-to-year transitions from 2004 to

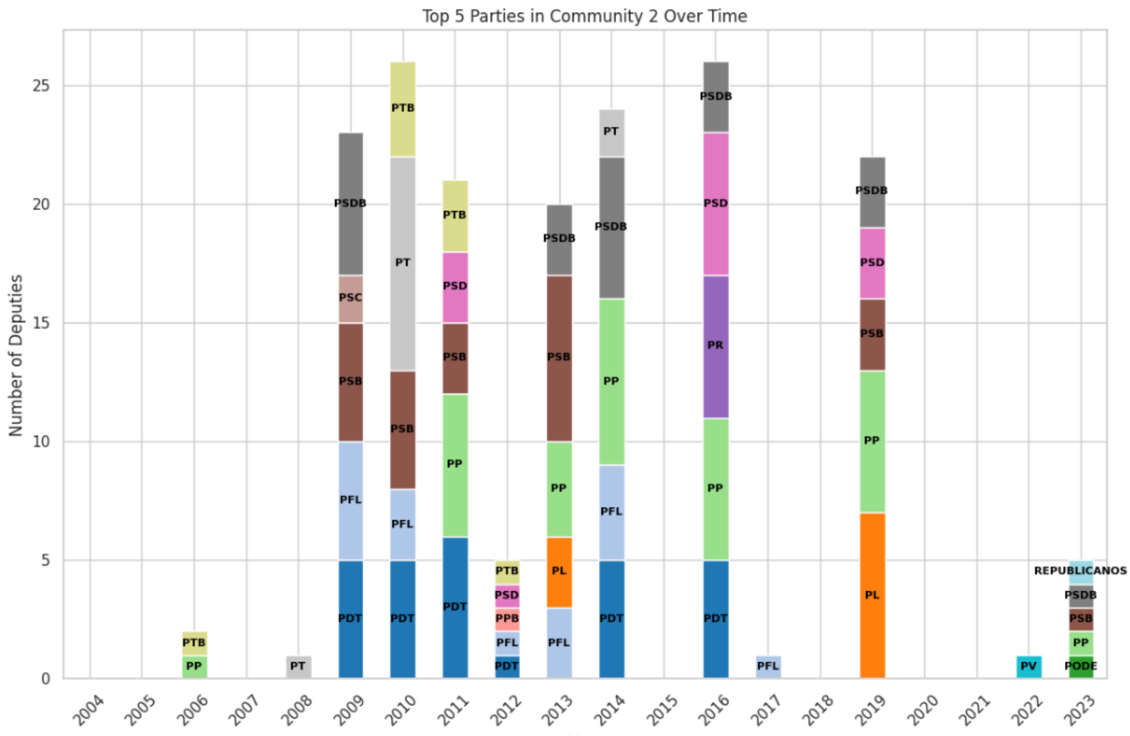


Figure 8. Top 5 parties in Community 2 over time, displaying the distribution of deputies by party from 2004 to 2023.

2023: the Herfindahl-Hirschman Index (HHI), the Adjusted Rand Index (ARI), and Edge Volatility. These metrics provide complementary perspectives on how political communities change across different legislative periods. While model-based approaches for temporal network analysis, such as the Dynamic Degree Corrected Stochastic Block Model [Wilson et al. 2019], provide probabilistic frameworks for detecting structural changes, our metric-based approach offers interpretable quantitative measures that directly assess community concentration, similarity, and structural stability, providing a transparent framework for understanding network evolution.

5.8.1. Herfindahl-Hirschman Index (HHI)

The HHI measures the concentration of deputies within communities for each year. Our analysis reveals an average HHI of 0.4541, with values ranging from 0.3386 to 0.5168 over the 20-year period. The highest concentration was observed in 2021 (HHI = 0.5168), indicating that communities were more concentrated during that year, with fewer, larger communities dominating the network structure. The average number of communities detected was 2.6, with a range of 2 to 3 communities per year, suggesting a relatively stable two-to-three bloc structure throughout the analysis period.

Figure 10 shows the evolution of HHI values over the study period, revealing fluctuations in community concentration that correspond to significant political events. Notable patterns include high concentration during Lula's presidency (2004-2008, HHI

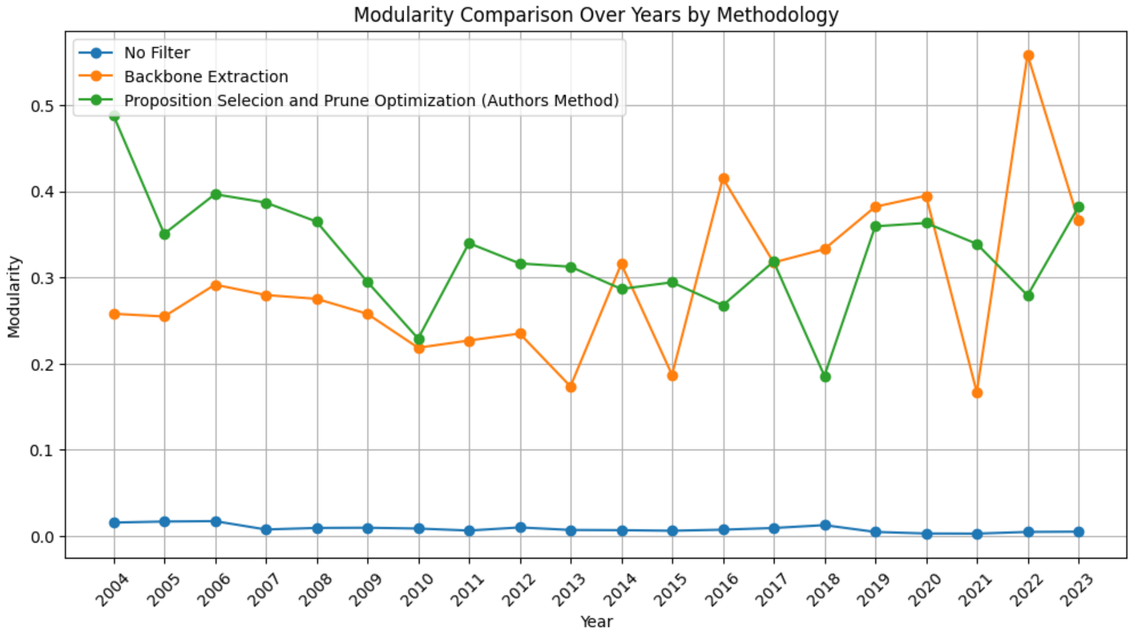


Figure 9. Modularity comparison over the years by methodology, showing the performance of "No Filter," "Backbone Extraction," and "Authors Method" (proposition selection and prune optimization) in achieving modularity from 2004 to 2023.

around 0.500), the lowest value in 2009 (HHI = 0.330) during political shifts, declines during the June 2013 protests and impeachment crisis (2013-2014, 2016), and peaks during Temer's presidency and Bolsonaro's administration (2017-2018, 2020-2022, reaching 0.517 in 2021), periods marked by clearer ideological divisions.

The HHI values indicate that the Brazilian Congress maintains a relatively concentrated community structure, with two main ideological blocs consistently present. The variation in HHI values over time reflects periods of greater or lesser consolidation, corresponding to major political events such as presidential elections, impeachment processes, or significant policy debates that strengthened or weakened party alignments.

5.8.2. Adjusted Rand Index (ARI)

The ARI quantifies the similarity between community partitions across consecutive years, providing a measure of community stability over time. Our results show an average ARI of 0.2096 across all year-to-year transitions, with values ranging from 0.0235 to 0.4621. The most stable transition occurred between 2017 and 2018 (ARI = 0.4621), indicating that the community structure remained relatively consistent during this period, with approximately 428 deputies on average overlapping between consecutive years.

Figure 11 presents the ARI values for each year-to-year transition, showing the degree of similarity between consecutive community structures. The relatively low average ARI value (0.2096) suggests that community structures are moderately fluid, with significant reorganization occurring between years. This finding aligns with the qualitative

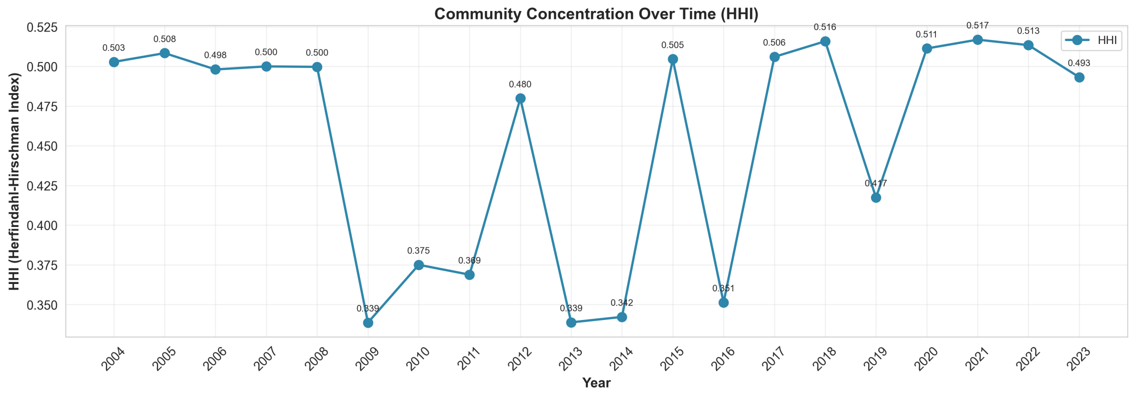


Figure 10. Community Concentration Over Time (HHI). The figure shows the Herfindahl-Hirschman Index values from 2004 to 2023, with each point labeled with its corresponding HHI value. Higher HHI values indicate more concentrated community structures.

observations of party shifts, with periods of low ARI corresponding to major political realignments such as the 2016 impeachment, while higher ARI values suggest more stable coalition configurations. Figure 12 shows the number of overlapping deputies between consecutive years, providing context for the ARI calculations.

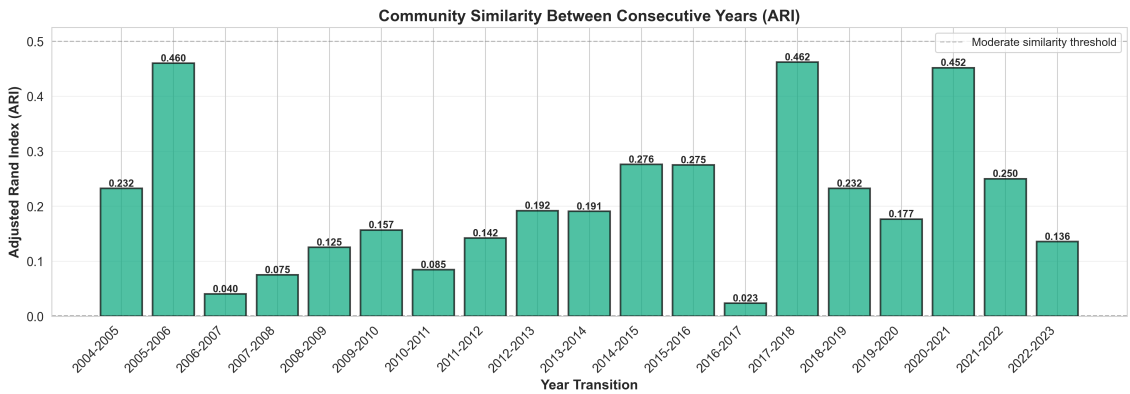


Figure 11. Community Similarity Between Consecutive Years (ARI). The bar chart displays the Adjusted Rand Index for each year-to-year transition from 2004 to 2023. Values closer to 1 indicate higher similarity, while values near 0 suggest substantial reorganization. The dashed line at ARI = 0.5 represents a moderate similarity threshold.

5.8.3. Edge Volatility

Edge Volatility measures the extent to which network connections change between consecutive legislative periods. Our analysis reveals an average edge volatility of 26.02% across all transitions, with values ranging from 8.41% to 40.48%. The least volatile transition occurred between 2004 and 2005 (8.41%), indicating a period of relatively stable

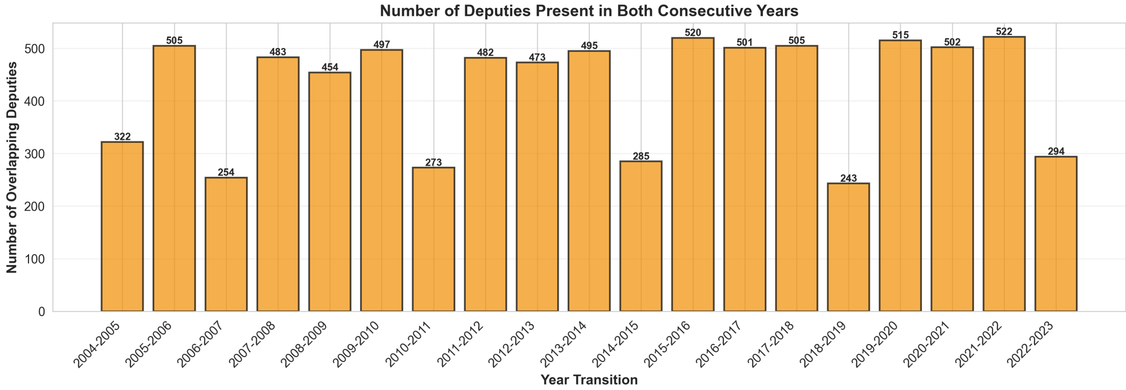


Figure 12. Number of Deputies Present in Both Consecutive Years. This figure shows the overlap in deputy membership between consecutive years, which serves as the basis for ARI calculations. Higher overlap values provide more reliable similarity assessments.

voting alignments. In contrast, the most volatile transitions reached up to 40.48%, suggesting significant shifts in voting patterns and deputy alignments.

Figure 13 displays the percentage of edges that changed their status between consecutive years. The high average volatility (26.02%) indicates that roughly one-quarter of network connections are modified annually, reflecting the fluid nature of political alliances. On average, approximately 6,674 edges changed per transition out of 26,404 common edges. Figure 14 provides additional context by showing both the number of changed edges and common edges for each transition. Periods of high volatility may coincide with low ARI values, indicating that structural network changes are accompanied by community reorganization.

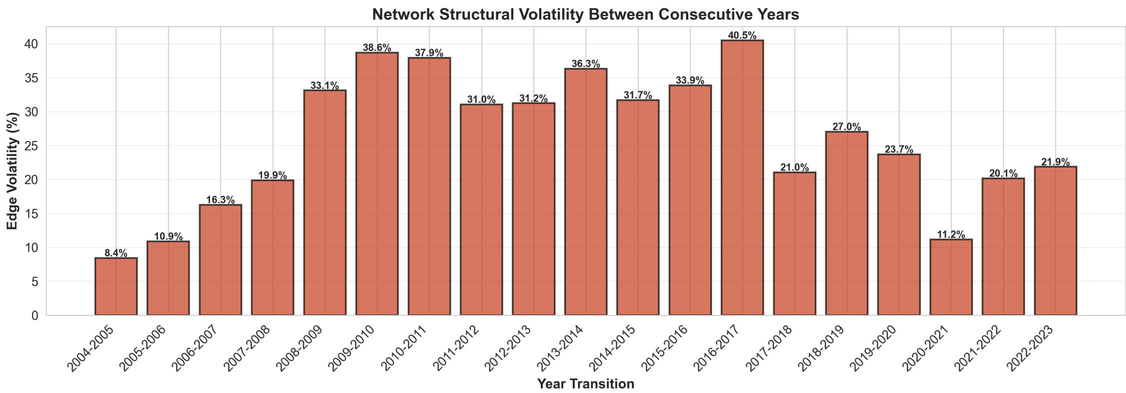


Figure 13. Network Structural Volatility Between Consecutive Years. The bar chart shows the percentage of edges that changed their community status (intra-community to inter-community or vice versa) for each year-to-year transition. Higher percentages indicate greater structural instability in the voting network.

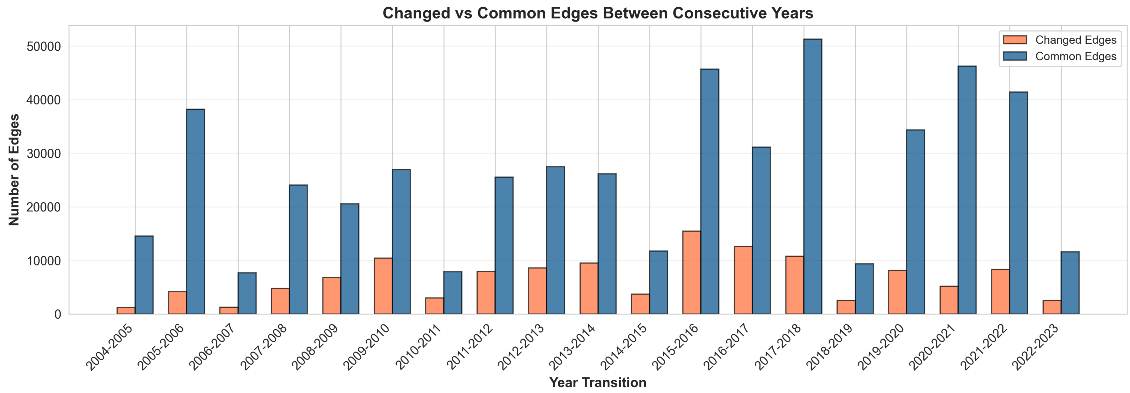


Figure 14. Changed vs Common Edges Between Consecutive Years. This figure compares the number of edges that changed their community status (orange bars) with the number of common edges present in both years (blue bars) for each transition. The comparison illustrates the relative magnitude of structural changes.

5.8.4. Integration of Temporal Metrics

The integration of HHI, ARI, and Edge Volatility provides a comprehensive quantitative framework for understanding temporal dynamics in the Brazilian Congress. The moderate HHI values indicate consistent community concentration, while the moderate-to-low ARI values reveal ongoing reorganization. The substantial edge volatility (26.02% average) confirms the dynamic nature of voting alignments, suggesting that while communities maintain a general structure (as reflected in HHI), the specific connections and memberships are subject to significant change (as reflected in ARI and Edge Volatility).

Figure 15 presents a scatter plot correlating ARI values with edge volatility, with color coding representing a stability score that identifies periods of high stability versus significant change. Figure 16 shows the stability score over time, directly quantifying periods of relative stability and instability in the political network structure.

These metrics complement the qualitative analysis of community evolution, providing quantitative validation of the observed patterns. For instance, the low ARI values align with the qualitative observations of party shifts in Communities 0, 1, and 2, while the edge volatility quantifies the extent of structural changes that underlie these shifts. Together, these metrics offer a more robust and nuanced understanding of political dynamics than either qualitative or quantitative approaches alone. This metric-based framework complements model-based approaches for temporal network analysis [Wilson et al. 2019] and methods for modeling dynamic ideological behavior [Ferreira et al. 2019], providing interpretable measures that directly quantify the stability and evolution of political communities over time.

5.9. Discussion of Key Findings

This subsection synthesizes the quantitative results and connects them to the broader interpretation of political alliances. Rather than treating communities as purely computational outputs, we discuss them as interpretable summaries of roll-call alignment that

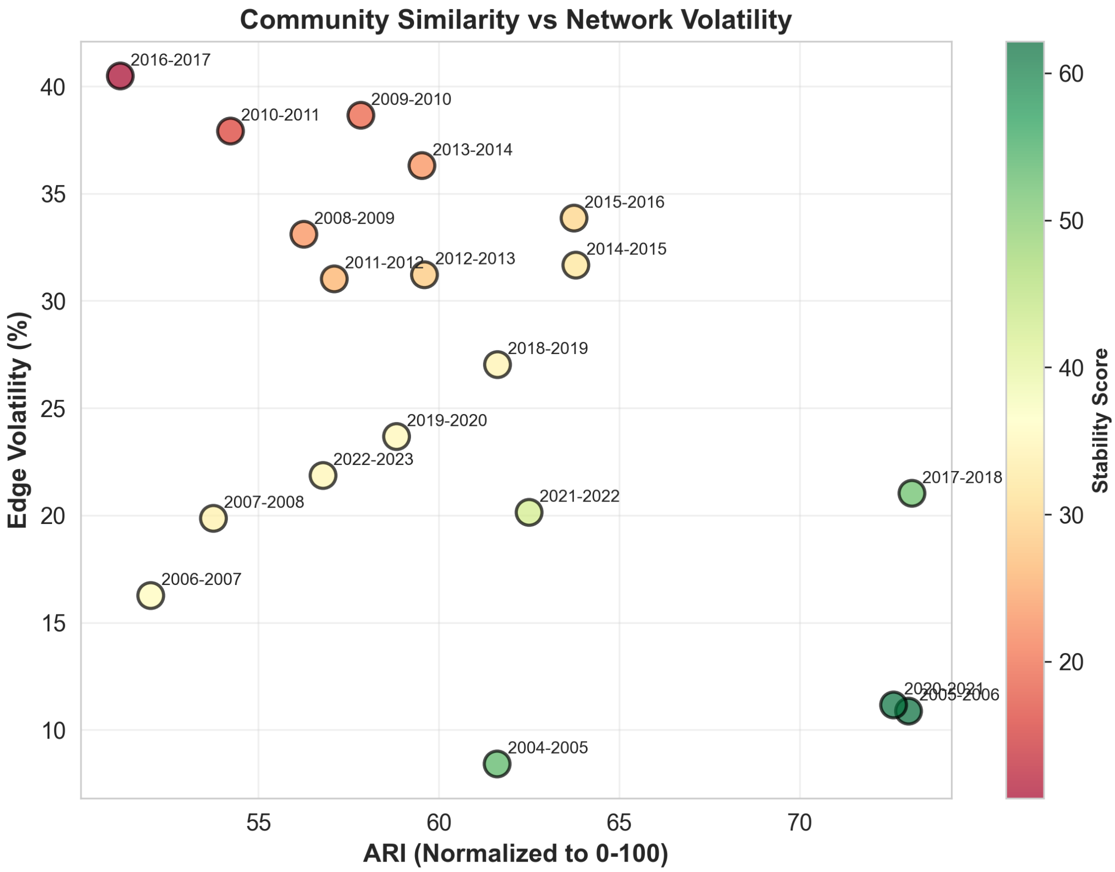


Figure 15. Community Similarity vs Network Volatility. This scatter plot shows the relationship between ARI (normalized to 0-100 scale) and Edge Volatility for each year-to-year transition. The color scale represents a stability score, with higher values (green) indicating more stable periods and lower values (red) indicating periods of significant change. Each point is labeled with its corresponding year transition.

can be compared across years and triangulated with institutional events (e.g., government transitions, agenda shifts, and coalition bargaining). We also highlight how modularity improvements and the proposed temporal metrics can be used to support transparency-oriented reporting and decision-making by different stakeholders.

5.9.1. Influence of Polarization and Pruning

Our results demonstrate that selecting polarized propositions and applying edge pruning optimization substantially enhance the effectiveness of community detection. By focusing on highly polarized propositions, we filter out noise and capture clearer ideological divides among deputies, allowing the formation of more cohesive and meaningful communities. Edge pruning further refines the network by removing weaker connections, which strengthens the internal structure of communities and emphasizes the most significant alliances within the network. This combined approach not only increases modularity

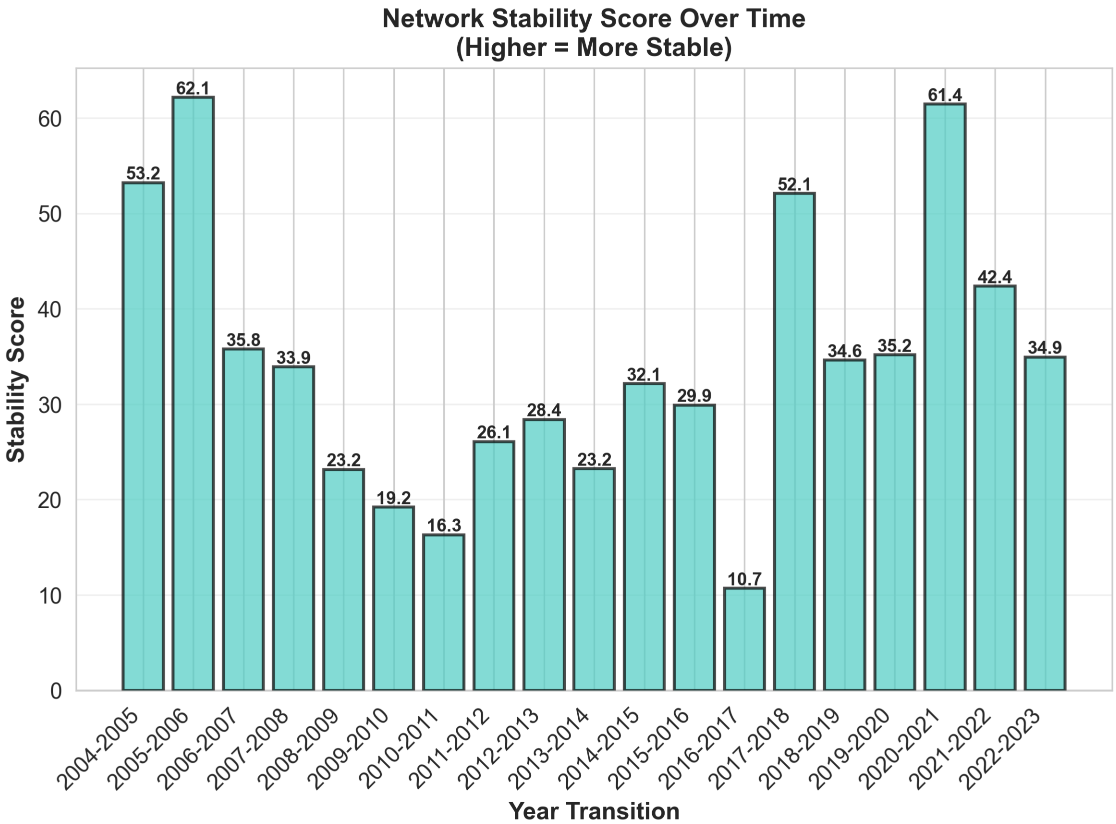


Figure 16. Network Stability Score Over Time. The bar chart displays a composite stability score for each year-to-year transition, calculated by combining ARI and Edge Volatility metrics. Higher scores indicate more stable periods with high community similarity and low structural volatility, while lower scores indicate periods of significant reorganization.

but also provides a more accurate reflection of political alliances in Congress, highlighting distinct ideological blocs.

5.9.2. Temporal Dynamics of Political Alliances

The analysis of temporal dynamics reveals that political alliances within the Brazilian Congress are fluid, with parties frequently shifting between communities over the years. From 2004 to 2023, we observed considerable variation in party representation within each community, as shown in the figures. For example, Community 1 shows increasing prominence of parties such as PSD, PP and PL in recent years, while Community 0 remains consistently dominated by PT. These shifts reflect changes in party alignment and may be associated with shifts in political leadership, evolving policy agendas, or responses to external events. Understanding these temporal dynamics is essential for accurately analyzing political behavior and forecasting potential shifts in alliances. Our findings align with research on modeling dynamic ideological behavior in political networks [Ferreira et al. 2019], which emphasizes the importance of capturing temporal shifts in political alignments and ideological positions to understand the evolving nature of legislative coalitions.

When connected to institutional context, the detected communities act as a compact representation of *who votes with whom* under the formal decision process of roll-calls. This can support transparency by providing interpretable, repeatable summaries for the public; it can support accountability by enabling longitudinal comparisons (e.g., identifying when a party systematically aligns with a bloc despite public commitments); and it can support governance by highlighting transitions where coalition structure becomes unstable (low ARI, high volatility), which may correlate with higher negotiation costs or policy uncertainty.

To make the bridge to practice explicit, the results can inform decisions such as:

- **Journalistic investigation**: prioritize interviews and fact-checking around deputies/parties that move between communities during low-ARI transitions, and explain coalitional shifts with quantitative evidence.
- **Institutional monitoring**: flag high-volatility years for deeper scrutiny (e.g., investigating whether procedural changes, coalition bargaining, or party switching coincide with structural network change).
- **Legislative strategy and negotiation**: identify persistent "swing" communities to anticipate pivotal blocs in agenda formation and coalition building.
- **Civic transparency tools**: build dashboards that show community membership, modularity, and temporal stability to help citizens navigate complex voting behavior without reducing it to single-issue narratives.

6. Conclusion

This study provides a longitudinal analysis of political alliances in Brazil's Chamber of Deputies by applying the Leiden algorithm enhanced with polarized proposition selection and iterative edge pruning. The resulting communities and temporal metrics capture two major ideological blocs and a recurring swing community, while improving modularity by 10.95% over a backbone-extraction baseline.

From an Information Systems perspective, we position the pipeline as a transparency- and accountability-oriented decision-support artifact [Hevner et al. 2004, Meijer 2013, Bovens 2007]. By producing interpretable yearly partitions and stability indicators (HHI, ARI, Edge Volatility), the approach can help journalists, oversight institutions, legislative observatories, parties, and citizens reason about coalition structure, ask better explanatory questions, and identify periods of heightened institutional reconfiguration.

Key limitations follow directly from the interaction between data and institutional reality. Roll-call votes reflect formal outcomes, but they do not capture behind-the-scenes negotiation, agenda-setting, or strategic abstention. Likewise, stronger modularity does not guarantee that communities map one-to-one to stable ideological identities; interpretation should be triangulated with context (government/opposition status, leadership releases, party switching, and salient events) [Scott 2014].

Future work can strengthen the societal value of this contribution by integrating the method into user-facing decision-support tools. This includes interactive dashboards for open-government portals, alerting mechanisms for sudden drops in cohesion or spikes in volatility, and co-designed evaluation with journalists and oversight professionals. Methodologically, future work can incorporate additional institutional signals (committee roles, bill topics, party leadership positions, media statements), explore dynamic community detection to reduce post-hoc label alignment, and validate communities against external ground truth where available. Exploring alternative partitioning strategies, including hierarchical community detection [Girvan and Newman 2002], may also reveal multi-level structures that are useful for governance analysis.

References

[Blondel et al. 2008] Blondel, V., Guillaume, J., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008.

[Boudebza et al. 2022] Boudebza, S., Boujenfa, K., Hamdaoui, B., and Kharoune, M. (2022). An approach for detecting dynamic communities in social networks. *arXiv preprint*, 2212.02383.

[Bovens 2007] Bovens, M. (2007). Analysing and assessing accountability: A conceptual framework. *European Law Journal*, 13(4):447–468.

[Brito et al. 2020] Brito, A. C. M., Silva, F. N., and Amancio, D. R. (2020). A complex network approach to political analysis: Application to the brazilian chamber of deputies. *PLoS ONE*, 15(3):e0229928.

[Bryden and Silverman 2019] Bryden, J. and Silverman, E. (2019). Underlying socio-political processes behind the 2016 us election. *PLoS ONE*, 14(4):e0214854.

[Cherepnalkoski et al. 2016] Cherepnalkoski, D., Karpf, A., Mozetič, I., and Grčcar, M. (2016). Cohesion and coalition formation in the european parliament: roll-call votes and twitter activities. *PLoS ONE*, 11(11):e0166586.

[da Silva et al. 2025] da Silva, P. N. F., Nascimento, D. C., and Nogueira, B. (2025). Community detection and analysis of political alliances in the brazilian congress voting network. In *Anais do XXI Simpósio Brasileiro de Sistemas de Informação (SBSI 2025)*, pages 535–544, Porto Alegre, RS, Brasil. SBC.

[Faustino et al. 2019] Faustino, J., Barbosa, H., Ribeiro, E., and Menezes, R. (2019). A data-driven network approach for characterization of political parties' ideology dynamics. *Applied Network Science*, 4(1):48.

[Ferreira et al. 2018] Ferreira, C., de Sousa Matos, B., and Almeira, J. (2018). Analyzing dynamic ideological communities in congressional voting networks. In *International Conference on Social Informatics*, pages 257–273. Springer.

[Ferreira et al. 2019] Ferreira, C., Murai, F., de Souza Matos, B., and de Almeida, J. (2019). Modeling dynamic ideological behavior in political networks. *The Journal of Web Science*, 1:1–14.

[Girvan and Newman 2002] Girvan, M. and Newman, M. E. J. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826.

[Hevner et al. 2004] Hevner, A. R., March, S. T., Park, J., and Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1):75–106.

[Hubert and Arabie 1985] Hubert, L. and Arabie, P. (1985). Comparing partitions. *Journal of Classification*, 2(1):193–218.

[Janssen et al. 2012] Janssen, M., Charalabidis, Y., and Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management*, 29(4):258–268.

[Levorato and Frota 2017] Levorato, M. and Frota, Y. (2017). Brazilian congress structural balance analysis. *Journal of Interdisciplinary Methodologies and Issues in Science*, 2.

[Maso et al. 2014] Maso, C. D., Pompa, G., Puliga, M., Riotta, G., and Chessa, A. (2014). Voting behavior, coalitions and government strength through a complex network analysis. *PLoS ONE*, 9(12):e116046.

[Meijer 2013] Meijer, A. (2013). Understanding the complex dynamics of transparency. *Public Administration Review*, 73(3):429–439.

[Melo et al. 2018] Melo, H., Reis, S., Moreira, A., Makse, H., and Andrade, J. J. (2018). The price of a vote: Diseconomy in proportional elections. *PLoS ONE*, 13(8):e0201654.

[Moody and Mucha 2013] Moody, J. and Mucha, P. (2013). Portrait of political party polarization. *Network Science*, 1(1):119–121.

[NetworkX Developers 2024] NetworkX Developers (2024). *NetworkX Documentation*. Accessed: 18-Oct-2024.

[Newman 2018] Newman, M. (2018). *Networks: an introduction*. Oxford University Press.

[Scott 2014] Scott, W. R. (2014). *Institutions and Organizations: Ideas, Interests, and Identities*. SAGE Publications, Thousand Oaks, CA, 4 edition.

[Serrano et al. 2009] Serrano, M., Boguñá, M., and Vespignani, A. (2009). Extracting the multiscale backbone of complex weighted networks. *Proceedings of the National Academy of Sciences*, 106(16):6483–6488.

[Traag et al. 2019] Traag, V., Waltman, L., and van Eck, N. (2019). From louvain to leiden: guaranteeing well-connected communities. *Scientific Reports*, 9.

[Wilson et al. 2019] Wilson, J., Stevens, N., and Woodall, W. (2019). Modeling and detecting change in temporal networks via a dynamic degree corrected stochastic block model. *Quality and Reliability Engineering International*, 35(5):1363–1378.

[Zhou et al. 2012] Zhou, F., Mahler, S., and Toivonen, H. (2012). Simplification of networks by edge pruning. In *Bisociative Knowledge Discovery*, pages 179–198. Springer.