The Journal of Web Science, 2019, 7

Modeling Dynamic Ideological Behavior in Political Networks

Carlos H. G. Ferreira[1], Fabricio Murai[1], Breno de Souza Matos[1] and Jussara M. de Almeida[1]

[1]*Department of Computer Science - Universidade Federal de Minas Gerais, Brazil*

{chgferreira, murai, brenomatos, jussara}@dcc.ufmg.br

ABSTRACT

In this article, we model and analyze the dynamic behavior of political networks, both at the individual (party member) and ideological community levels. Our study relies on public data covering 15 years of voting sessions of the House of Representatives of two diverse party system, namely, Brazil and the United States. While the former is an example of a highly fragmented party system, the latter illustrates the case of a highly polarized and non-fragmented system. We characterize the ideological communities, their member polarization and how such communities evolve over time. Also, we propose a temporal-ideological space model, based on temporal vertex embeddings, which allows us to assess the individual changes in ideological behavior over time, as expressed by the party members' voting patterns. Our results unveil very distinct patterns across the two case studies, both in terms of structural and dynamic properties.

Keywords: dynamic behavior model, temporal node embedding, community evolution, political networks

ISSN 2332-4031; DOI 10.34962/jws-80

1 Introduction

Party systems can be classified with respect to fragmentation and polarization [58]. Fragmentation corresponds to the number of parties existing in a political system (e.g., a country), while polarization is related to the multiple opinions that lead to the division of members into groups with distinct political ideologies [23, 58]. In countries where the party system has low fragmentation, the polarization of political parties can be more clearly observed since one party tends to occupy most seats supporting the government than the others opposing it [40]. Conversely, in fragmented systems, the many political parties often create coalitions, a inter-party alliance, to raise their influence in the political system and reach a common end [2, 11]. Thus, a great deal of ideological similarity, as expressed by voting decisions, is often observed across different parties.

Previous work has analyzed the ideological behavior of political party members by the modeling of voting data in signed and weighted networks [3, 4, 13, 34, 42] These prior efforts tackled topics such as community detection, party cohesion and loyalty analysis, governance of a political party and member influence in such networks. Yet, the identification and characterization of ideological communities, particularly in fragmented party systems, require observing some issues, such as: (i) presidents may define coalitions to strengthen the implementation of desired public policies, which may be ruptured after some time [11, 39]; (ii) political members have different levels of partisanship and loyalty, and their political preferences may change over time [3, 6]; and (iii) different parties may have the same political ideology, being redundant under a party system [62].

In such context, we here study the dynamic behavior of political party members aiming at identifying how ideological communities are created and evolve over time and how individual members change their ideological behavior with respect to others. To that end, we consider two case studies, Brazil and US, which are representatives of distinct party systems: the former is highly fragmented and redundant [62], while the latter is not fragmented but rather polarized with two major parties, although some party members are less polarized [17].

Using public voting data of the House of Representatives of both countries, covering a 15-year period, we model and characterize the emergence and evolution of communities of House members with similar political ideology and ideological changes of individual members over time. We study group and individual ideological behavior, as captured by their voting decisions, aiming at tackling four research questions (RQs):

- **RQ1: How are ideological communities in governments with different (fragmented and non-fragmented) party systems characterized?** We model the voting behavior of each House member during a given time period using a network, where nodes represent members of the same House of Representatives, and weighted edges are added if two members voted similarly. We use the Louvain algorithm [10] to detect communities in each network and characterize structural properties of such communities. Unlike the aforementioned prior analyses in the political domain, we compare the properties of these communities in fragmented and in non-fragmented party systems.

- **RQ2: How can we identify polarization in the ideological communities?** We use neighborhood overlap to estimate the tie strength associated to each network edge, characterizing it as either strong or weak [20]. This approach to estimate tie strength has been employed in several contexts [30, 41, 61, 66] and to a short extent in the political domain [64]. But unlike prior studies, we use strong ties to identify polarized communities in each network, comparing distinct political systems with respect to polarization.

- **RQ3: How do polarized communities evolve over time?** We analyze how polarized communities evolve over the years of a government, characterizing how the membership of such communities change over time.

- **RQ4: How can we assess the ideological changes of individual House members over time?** We capture ideological changes, as expressed by members' voting behavior, by mapping the network into a temporal latent ideological space. Building upon a recent work [67], we learn temporal *vertice embeddings* for consecutive networks (representing consecutive years) jointly, so that we can track individual members over time in the defined space. By doing so, we are able to analyze how the locations of individual members in this space change and thus measure ideological shifts over time. Unlike prior studies that use contextual information (e.g., topics of voting sessions [31, 47], prior speeches of individual members [21, 33, 57]) to build an ideological space, we use only the topological structure of the networks (which come from the voting data itself) to build such space, being thus a more general approach.

In sum, the contributions of our work are: (i) a methodology to identify and analyze dynamic ideological communities and their polarization in party systems based on complex network concepts; (ii) a temporal node embedding approach to assess ideological temporal changes; and (iii) two case studies covering very different party systems over a long time period.

Our study shows that in fragmented party systems, such as Brazil, despite party redundancy, some ideological communities exist and may, indeed, be polarized. However, such polarized communities are highly dynamic, greatly changing their membership over consecutive years. Indeed, by exploiting our temporal embedding approach, we show that most party members do consistently change their ideological behavior over consecutive years, either individually or in group (party-level). This is true, though to a lesser degree, for members of polarized communities as well. In the US, on the other hand, despite the strong and temporally stable party polarization, there are members, within each party, that exhibit different levels of polarization. Yet, in sharp contrast to the Brazilian case, most members have stable ideological behavior with practically no changes over successive years in the defined ideological space.

A preliminary version of this study was first presented in [22]. We here extend that prior effort by proposing a method to model dynamic ideological spaces based on temporal node embedding and by using this method to characterize how party members ideological behavior changes over time in both case studies. The rest of this article is organized as follows. Section 2 briefly discusses related work, whereas Section 3 describes our modeling methodology and case studies. We then present our main results, tackling RQ1-RQ4 in Sections 4-7. Conclusions and future work are offered in Section 8.

2 Related Work

Our focus in this article is on modeling and analyzing the dynamics of ideological behavior in the political domain, relying on complex network concepts to drive a great part of our effort. Thus, there are mainly three large bodies of work that relate to our present study: (1) modeling user interactions using complex network concepts; (2) modeling and analyzing political behavior and ideologies; and (3) using temporal embeddings to model changes and dynamic behavior.

2.1 Modeling User Interactions

Complex networks constitute a set of theoretical and analytical tools to describe and analyze phenomena related to interactions occurring in the real world [56]. Among the many properties of a network, the interactions between pairs of nodes can be used to define the strength of these links (or *tie strength*) [20]. Indeed, tie strength is a property that has been widely studied in several domains and defined in different ways. For example, the tie strength between pairs of people was studied in phone call and Short Message System (SMS) networks, where higher frequencies of SMS exchange and longer call durations characterize stronger ties [66]. The different types of interactions between Facebook users have also been used to define tie strength on that system [30]. Similarly, tie strength was used to build geolocation models based on Twitter data and exploited in the prediction of user location [41].

2.2 Modeling Political and Ideological Behavior

In the political context, the study of political ideologies has been largely accomplished through the analysis of roll call votes networks. In a roll call votes network, the nodes represent people (e.g., members), and two nodes are connected if they have voted similarly in one (or more) voting sessions. Using these networks, the authors of [3] studied committees' formation in the US House of Representatives, concluding that, despite the recent increase in polarization, there are moderate members in both parties who cooperate with each other. Similarly, authors of [53] studied the committees and subcommittees of the same chamber, exploiting the network connections that are built according to common membership. Analogously, the polarization in the US Senate was evaluated using a network defined by the similarity of senators' votes [44].

In [17], the authors studied the relations between members of the Italian parliament according to their voting behavior, analyzing the community structure with respect to political coalitions and government alliances over time. Similarly, the cohesiveness of members of the European parliament was investigated through the analysis of network models combining roll call votes and Twitter data [13]. Others studied the behavior of political members, modeling roll call votes using signed networks. For example, Levorato *et al.* used signed networks to evaluate aspects related to political governance and party behavior in the Brazilian House of Representatives [34]. The results revealed inefficient coalitions with the government as parties that make such coalitions have members distributed in different ideological communities over time. Mendonça *et al.* proposed an algorithm to evaluate signed networks using the European parliament network as case study [42]. Orthogonally, others have investigated the ideology of political members and users through profiles of social networks [1, 18].

Unlike the aforementioned studies, our focus is on identifying and characterizing ideological communities in *both* fragmented and non-fragmented party systems. We also propose the use of tie strength, computed based on neighborhood overlap, to identify polarized communities, evaluating their evolution over time. To our knowledge, such analyses on *diverse* party systems are novel contributions of this work.

A closely related body of work has used roll call votes to measure *latent ideological patterns*. One such family of procedures is known as NOMINATE, whose variants are D-NOMINATE (originally called 'NOMINATE') [49], W-NOMINATE [50] and DW-NOMINATE [51]. NOMINATE procedures assume a spatial model where each member has an ideal position in a space, while 'yea' and 'nay' votes on each roll call take on two positions in that space. Both D-NOMINATE and W-NOMINATE assume a multidimensional space (typically bidimensional), where errors (i.e., a member closer to a certain vote decides to vote the opposite way) follow a logit model. Unlike the former, W-NOMINATE assumes a distance model where dimensions are weighted differently, allowing for more flexible utility functions. DW-NOMINATE builds and improves upon W-NOMINATE by letting errors be normally distributed.

In [9], the authors discussed some key shortcomings of methods based on ideal positions such as DW-NOMINATE and why they are not used more often in the American Political Development literature. One such limitation is the assumption of linear change in a member's ideal position over time. Moreover, these methods disregard important data both in the non-fragmented and fragmented political system scenarios. For instance, such methods cannot leverage information from unanimous votes – a typical situation in less polarized and fragmented systems – which are discarded before parameter estimation [49]. Similarly, the identities of members who have changed parties during the period of analysis are also disregarded. For instance, in [52], the voting behavior of a member who changes parties once is considered as two independent sequences. While this is not a severe issue in non-fragmented systems, it can introduce a large amount of noise when analyzing fragmented systems, such as the case of Brazil's party system, where such changes occurs with greater frequency. In contrast, our approach allows us to identify members who switched parties over time by observing member displacement in the latent representation obtained.

Clinton *et al.* [14] proposed a Bayesian simulation approach that improves existing methods by allowing the inclusion of ancillary information (e.g., the location of extremist members, member-specific covariates, or the evolution of the legislative agenda) in the model. The proposed framework also allows estimating changes of ideal positions over time by modeling the process associated with that change (e.g., members switching political parties). Although this approach offers a number of advantages over the aforementioned point estimate models, it also retains some statistical issues in relation to Bayesian ideal point estimation, such as proper variance estimates, scale and translation invariance, reflection invariance and outliers [5].

In contrast, our approach is based on a state-of-the-art node embedding technique for networks. This technique allows us to represent each node (member) as a point in a vector space of arbitrary dimension based on the way it is connected to the network in a structural sense, i.e., going beyond its connections to immediate neighbors. Moreover, we consider a sequence of legislatures/congresses as a dynamic network and, hence, we represent a node in different snapshots as different vectors. To that end, we build upon recent improvements on dynamic word embeddings [67] that allow for much more flexible dynamics than the linear changes assumed by DW-NOMINATE. As a positive side effect, this enables us to compare House members even if they have never voted simultaneously. Also, unlike methods based on ideal positions, our proposed method can leverage information from unanimous votes and party migration. Lastly, our approach can be easily extended to account for more than two positions (i.e., 'yea' and 'nay'), eliminating the need to collapse 'absence' and 'nay' as a single opposition category, as is the case for the United Nations General Assembly, where abstention is a milder form of disapproval than a 'nay' vote [55].

2.3 Temporal Embeddings to Model Dynamic Behavior

As mentioned, one key contribution of this article is to model and quantify changes in ideological behavior of individual party members. Thus, a third body of work that relates to our present effort is the use of temporal embeddings to model changes and dynamic behavior. Despite the rich literature on the use of embeddings to extract latent signals in various domains (e.g., word embeddings in textual documents [7, 32] and node embeddings in networks [16]), the study of *temporal* embeddings is relatively new. Some efforts have proposed temporal latent space models by exploiting network embeddings [46, 69], in some cases jointly with node attributes [29, 36]. However, these prior approaches learn the embeddings for separate time windows independently, which ultimately leads to an "alignment problem". Put simply, this means that it may not be possible to place all learned embeddings in the *same* latent space, and thus it may be hard to track an element across time. One challenge in the "alignment problem" is to preserve similarities and to reveal differences of the neighborhood across time in the *same* latent space.

Recently, the authors of [67] proposed a method to model word semantic evolution which simultaneously learns time-aware word vector representations and effectively solves the aforementioned "alignment problem". To capture changes in ideological behavior of individual party members, we here adapt this technique to the *network* embedding domain applying it to node2vec [27], a popular graph embedding method. We use the adapted technique to map all party members onto a single stable latent ideological space covering multiple years, and then track member's locations in this single space over time.

A similar problem was addressed in [26], but the proposed approach has some key properties [25, 38] that are not desired given our target problem. Specifically, it assumes that the temporal changes in the networks are of short duration since it only considers the network of the previous time window to learn the next time embedding. Also, it uses the learned embedding from the previous time window to initialize the new one. These two properties implicitly keep the new embedding (time t) close to the immediately previous one (time t-1). Thus, the approach is unstable in sparse networks, when not all nodes are present in all time windows. In contrast, the method proposed in [67], which will be further discussed in Section 3.4, learns time-aware embeddings using all time windows simultaneously, reaching robustness for scenarios with both smooth and rough changes, as required in the political context. It is also more robust to data sparsity and more scalable in terms of memory usage, which is important for large networks.

3 Methodology

This section describes the methodology used in our study, starting with basic concepts (Section 3.1) and our case studies (Section 3.2). We then present our modeling of voting behavior (Section 3.3) and the time-aware node embedding approach used to model an ideological latent space (Section 3.4).

3.1 Basic Concepts

The House of Representatives is composed of several members who occupy the seats during each government period. House members participate in a series of voting sessions, when bills, amendments, and propositions are discussed and voted. Thus, attending such sessions is the most direct way for members to express their ideologies and opinions. When these members are associated with a large number of political parties, the party system in question is regarded as fragmented. In this case, during a term of office, coalitions are often established, leading political parties to organize themselves into ideological communities, defending together common interests during voting sessions [39, 58].

One can evaluate the behavior of parties and their members in terms of how cohesive they are as an ideological community by analyzing voting data using widely disseminated metrics, such as Rice's Index [54]. Yet, the use of Rice Index has been shown to be problematic when there are more than two voting options (other than only *yes* and *no*) [28], as, for example, in the European parliament and in our study, as we will see.

Instead, we here employ the *Partisan Discipline* and *Party Discipline* metrics, proposed in [62]. The former captures the ideological alignment of a member to her party (estimated by the behavior of the majority), and the latter expresses the ideological cohesiveness of a party. Given a member m, belonging to party p_m, the *Partisan Discipline* of m, pd_m, is given by the fraction of all voting sessions to which m attended and voted similarly to the majority of p_m's members. That is, let n be the number of voting sessions attended by m and $I(m, p_m, i)$ be 1 if m voted similarly to the majority of members of p_m in voting session i ($i = 1..n$) and 0 otherwise. Then,

$$pd_m = \frac{\sum_{i=1}^{n} I(m, p_m, i)}{n} \tag{1}$$

We note that pd_m ranges from 0 to 1, where 1 indicates that member m voted similarly to the majority of p_m's members in all voting sessions, and 0 indicates the opposite behavior. We note also that the *Partisan Discipline* can be generalized to assess the discipline and ideological alignment of a member to any community (not only his original party).

The *Party Discipline* of a party p is computed as the average *Partisan Discipline* of all of its members, that is,

$$PD(p) = \frac{\sum_{m=1}^{M} pd_m}{M} \tag{2}$$

where M is the number of members of p. *Party Discipline* captures how cohesive a party (or community) is in a set of votes. That is, a $PD(p)$ value of 1 (maximum) indicates that party p is totally disciplined (or cohesive).

3.2 Case Studies

We consider two case studies: Brazil and the United States (US). In Brazil, the House of Representatives consists of 513 seats. A member vote can be either *Yes*, *No*, *Obstruction* or *Absence* in each voting session. A *Yes* or *No* vote expresses, respectively, an agreement or disagreement with the given proposition. Both *Absence* and *Obstruction* mean that the member did not participate in the voting, although an *Obstruction* expresses the intention of the member to cause the voting session to be cancelled due to insufficient quorum. Similarly, the US House of Representatives includes 435 seats, and a member vote can be *Yes*, *No* or *Not Voting*, whereas the last one indicates the member was not present in the voting session. In our study, we disregard *Absence* and *Not Voting* votes, as they do not reflect any particular inclination of the members with respect to the topic under consideration. However, we do include *Obstructions* as they reflect an intentional action of the members and a clear opposition to the topic. Thus, for Brazil, three different voting options were considered.

For both case studies, we collected voting data from public sources. The plenary roll call votes of Brazil's House of Representatives are available through an application programming interface (API) maintained by the government[1]. We collected roll call votes from 2003 to 2017 (4 legislatures). US voting data covering the same 15-year period (i.e., between the 108^{th} and 115^{th} congresses) was collected through the ProPublica API[2]. Each dataset consists of a sequence of voting sessions; for each session, the dataset includes date, time and voting option of each participating member.

In a preliminary analysis of the datasets, we noted that some members had little attendance to the voting sessions, especially in Brazil. Thus, we chose to filter our datasets to remove members with low attendance as they introduce noise to our analyses. Specifically, we removed members that had not attended (thus had not associated vote) to more than 33% of the voting sessions during each year[3]. On average, 19% and 1.98% members were removed from the Brazilian and US datasets for each year, respectively.

Table 1 shows an overview of both (filtered) datasets, with Brazil on the top part of the table and the US on the bottom. The table presents, for each year, the acting president[4] and his/her party[5], total number of voting sessions, total number of member votes, as well as numbers of parties and members occupying seats in the House of Representatives during the year. The two rightmost columns, *Avg. PD* and *SD PD*, present the average and standard deviation of the *Party Discipline* computed across all parties. We show data for different Brazilian legislatures and US congresses in separate blocks of the table.

[1] http://www2.camara.leg.br/transparencia/dados-abertos/dados-abertos-legislativo (in Portuguese).

[2] https://projects.propublica.org/api-docs/congress-api/

[3] This threshold was chosen based on Article 55 of the Brazilian Constitution that establishes that a deputy or senator will lose her mandate if she does not attend more than one third of the sessions.

[4] Brazilian president Dilma Rousseff was impeached from Office in 2016 and, therefore, Brazil had two Presidents that year.

[5] For Brazil: Worker's Party (PT) and Democratic Movement Party (PMDB). For the US: Democratic (D) and Republican (R).

Table 1: Datasets overview (PD: party discipline, SD: st. dev.)

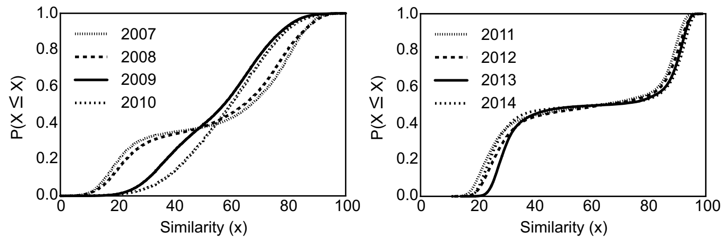
Brazil (52^{nd} to 55^{th} legislatures)							
Year	**President (Party)**	**# of Sessions**	**# of Votes**	**# of Parties**	**# of Members**	**Avg. PD(%)**	**SD PD**
2003	Lula (PT)	150	106755	23	435	88.23	0.08
2004	Lula (PT)	118	71576	23	377	87.43	0.08
2005	Lula (PT)	81	50616	24	382	88.91	0.07
2006	Lula (PT)	87	62358	24	419	91.12	0.05
2007	Lula (PT)	221	190424	31	478	92.45	0.07
2008	Lula (PT)	157	122482	31	452	92.34	0.07
2009	Lula (PT)	156	125759	30	465	91.87	0.06
2010	Lula (PT)	83	63255	29	452	92.46	0.05
2011	Dilma (PT)	98	78662	29	481	89.34	0.08
2012	Dilma (PT)	79	60219	28	454	89.56	0.05
2013	Dilma (PT)	158	115751	29	451	88.70	0.06
2014	Dilma (PT)	87	66154	28	451	92.93	0.04
2015	Dilma (PT)	273	231031	28	502	85.84	0.06
2016	Dilma (PT) Temer (PMDB)	218	156006	28	452	90.12	0.05
2017	Temer (PMDB)	230	159704	29	435	89.76	0.08

United States (108^{th} to 115^{th} congresses)							
Year	**President (Party)**	**# of Sessions**	**# of Votes**	**# of Parties**	**# of Members**	**Avg. PD(%)**	**SD PD**
2003	Bush (R)	623	258867	3	432	95.76	0.03
2004	Bush (R)	502	203557	3	427	95.11	0.03
2005	Bush (R)	637	264735	3	432	95.02	0.03
2006	Bush (R)	511	210592	3	428	94.98	0.04
2007	Bush (R)	956	297957	2	414	92.23	0.04
2008	Bush (R)	605	244734	2	426	92.73	0.04
2009	Obama (D)	929	385344	3	431	93.78	0.02
2010	Obama (D)	631	253296	3	422	95.34	0.01
2011	Obama (D)	908	377601	2	428	91.98	0.01
2012	Obama (D)	621	253812	2	425	91.50	0.01
2013	Obama (D)	594	245430	2	427	93.04	0.01
2014	Obama (D)	531	217822	2	426	93.24	0.01
2015	Obama (D)	662	277732	2	432	94.87	0.01
2016	Obama (D)	588	241263	2	427	95.11	0.01
2017	Trump (R)	708	292503	2	427	95.99	0.00

Starting with the Brazilian dataset, we can see that the number of parties occupying seats has somewhat grown in recent years, characterizing an increasingly fragmented party system. Yet, in general, average *PD* values are very high (ranging from 85% to 92%), with small variation across parties, indicating that, despite the fragmentation, most party members have high partisan discipline. Regarding the American dataset, Table 1 shows that the number of voting sessions is much larger than in Brazil. This is because the API of the Brazilian House of Representative provides only data related to votes in plenary, while the US dataset covers all votes. Moreover, although the numbers of members are comparable to those in the Brazilian dataset, the number of parties occupying seats in each year is much smaller. Indeed, only two parties, namely Republican (R) and Democrat (D), fill all available seats since the 112^{th} Congress. Thus, unlike the Brazilian case, party fragmentation is not an issue in the US system. Nevertheless, parties have a high party discipline in both systems.

3.3 Network Model

We model the dynamics of ideological communities in voting sessions in each country using graphs as follows. We discretize time into non-overlapping windows of fixed duration. For each time window w analyzed, we create a weighted and undirected graph $G^w(V, A)$ in which $V = \{v_1, v_2, ...v_n\}$ is a set of vertices representing House members and each edge (v_i, v_j) is weighted by the similarity of voting positions of members v_i and v_j. Specifically, the weight of edge (v_i, v_j) is given by the ratio of the number of sessions in which both members voted



(a) Brazil (b) United States

Figure 1: CDF of similarity (edge weights) in percentage.

similarly to the total number of sessions to which both members attended, during window w. Since in Brazil, government coalitions are usually made every year, we choose one year as the time window for analyzing community dynamics.

After building each graph, we noted that all pairs of members voted similarly at least once in all years analyzed and in both countries and, therefore, all graphs built are complete. This reflects the fact that some voting sessions are not discriminative of ideology or opinion, as most members (regardless of party) voted similarly. Thereby, it is necessary to filter out edges that do not contribute to the detection of ideological communities. To that end, we started by analyzing the distributions of edge similarity for all networks. Representative distributions for specific years are shown in Figures 1a and 1b for Brazil and US, respectively. We note that while the distributions for the US have clear concentrations around very small (roughly 30%) and very large (around 85%) similarity values, the distributions for Brazil exhibit greater variability, which is consistent with the greater fragmentation of the party system.

A widely used approach to filter out less discriminative connections from networks is to define a similarity threshold. Following previous work [15, 59], we adopt thresholds defined by the network *context*. Specifically, we argue that the threshold should not be much smaller than the average partisan discipline of individual members. That is, two members that have much lower similarity than their partisan disciplines should not be considered as part of the same ideological community. On the other hand, the higher the similarity threshold chosen, the larger the number of edges removed and the more sparse the resulting graph is. Based on these observations, we chose to remove all edges with weights below the 90^{th} percentile of the similarity distribution for the Brazilian graphs. For the US, we removed edges with weights below the 55^{th} percentile of the similarity distribution. Both percentiles correspond roughly to a similarity value of 80%, which is not much smaller than the average partisan disciplines in both countries (see Table 1). We removed nodes that become isolated after the edge filtering. We found that these thresholds yield a good tradeoff between removing less discriminative connections and graph sparsity. Specifically, the fraction of nodes and edges removed from the Brazilian networks fall in the 0-24% and 86-93% ranges, respectively, across all years analyzed. For the US, the fractions are much lower, varying from 0% to 11% for nodes and from 54% to 56% for edges.

In sum, we model the voting sessions using two sets of networks, one set per country, one network per year. Then, we use the Louvain Method [10] to identify ideological communities in each network. This method has been extensively used to detect network communities in various domains [12, 24, 48]. It is based on the optimization of *modularity* [45], a metric to evaluate the structure of clusters in a network. *Modularity* is large when the clustering is good, with a maximum value of 1. We here use *modularity* and *party discipline* as main metrics to assess the cohesiveness of the communities found. The former captures the quality of the result with respect to the topological structure of the communities in the network, whereas the latter, computed for the communities (rather than for individual parties), captures quality in terms of context semantics.

3.4 Ideological Space Model

In order to model how the ideological behavior of individual party members evolves over time, we start from the networks defined in Section 3.3, which capture each individual's behavior in terms of how a member voted relative to others during a given time window. We then build a network representation that embeds vertices into a low-dimensional vector space, which preserves properties of the network's topological structure. Since the ideological behavior of an individual member is here captured by how she voted relatively to her peers (i.e., by her neighborhood in the network), we consider the low-dimensional latent space produced by the graph embedding technique an *ideological space*. One key challenge is how to track individual members over time in this ideological space so as to identify changes in their behavior. This is difficult because there are multiple networks (and thus network embeddings), one for each time window under consideration. In this section, we describe our approach to address this challenge and build a consistent time-aware ideological space.

We build upon node2vec, a popular graph embedding technique [27]. Node2vec learns low-dimensional representations for vertices in a single graph by performing biased random walks, and using them as input to word2vec [43], a widely used word embedding technique. Word2vec receives as input a textual corpus and produces as output a vector space. Each word in the input corpus is mapped into a point in the vector space such that words that share common contexts in the input corpus fall close to each other in the vector space. In the case of node2vec, assuming that random walks are input sentences and visited vertices represent individual words, vertices are mapped into the low-dimensional latent space so as to maximize the likelihood of preserving the network neighborhoods. Grover *et al.* defined a flexible notion of neighborhood [27], which can be instantiated differently by carefully choosing the parameters of the biased random walk procedure (see more details below).

However, like word2vec, node2vec also suffers from the "alignment problem" when applied to a temporal sequence of networks. That is, the embeddings generated by the successive application of node2vec to networks for consecutive time windows are *not* mapped onto the same latent space. Thus, a vertex representation in one embedding has no correspondence to its representation in the next embedding (i.e., the one generated from the next time window).

Yao *et al.* tackled the problem of inferring how word semantics evolve over time by proposing a dynamic statistical model for learning time-aware word embeddings [67]. The proposed solution, which we refer to as DynamicWord2Vec, effectively addresses the "alignment problem" in the context of word embeddings. Inspired by that work, we build a temporally-consistent ideological space to represent parties and their political members by adapting DynamicWord2Vec to the network domain, combining it with node2vec. That is, we modify the node2vec implementation so that it uses DynamicWord2Vec (instead of word2vec) to generate an embedding from the sampled walks. Next, we briefly review how node2vec works and how we combine it with DynamicWord2Vec. We refer the reader to [27, 67] for further details on each technique.

Node2vec [27] uses a strategy of neighborhood sampling through a biased random walk which behaves, at each step, either as breadth-first sampling (BFS) or as depth-first sampling (DFS). In BFS, the neighborhood of a given source vertex v_s is restricted to vertices that are immediate neighbors of the source, while DFS consists of vertices sequentially sampled at increasing distances from v_s. We here want the walk to enforce BFS more often than DFS to better capture the similarities in the ideological space, rather than structural equivalences in the network [27]. To control this behavior, node2vec has two parameters, p and q. Parameter p determines the likelihood of immediately going back to an already visited vertex. Parameter q allows us to control whether the walk stays close to the source vertex, exploring the same neighborhood (i.e., corresponding to BFS), or whether it should walk further away, exploring other vertices (i.e., corresponding to DFS). We here are focused on the former, i.e., sampling immediate neighbors of the source more often. Thus, we set the parameter values according to the authors' recommendations for such case [27], i.e., $p = 1$ and $q = 0.5$. By doing so, we skew the random walks towards the immediate neighborhood of each source vertex.

In addition to p and q, node2vec allows us to define the number of walks per vertex and the length of each walk (i.e., number of vertices visited in each walk). These parameters directly determine the sampling process, which tends to saturate at a certain point as they increase [27]. In our experiments, we found that 16 walks per vertex, each with length 40, are sufficient to perform the sampling process in our case studies. Increasing either the number of walks or the length of each walk further caused a proportional increase in the co-occurrence of vertices in the walks, without bringing further information.

After computing the probabilities of the possible paths according to p and q and sampling the walks, node2vec builds a walk matrix S of size $k \times l$, where k is the product of number of walks and number of vertices and l is the length of each walk. S contains all vertices visited in all walks performed, starting from all vertices in the graph as sources. In the original node2vec algorithm, given a matrix S, the representations of the vertices are optimized using stochastic gradient descent so that vertices in the same neighborhood appear more closely in the generated latent space. Instead, we here use the DynamicWord2Vec technique as follows.

We want to learn a single latent space covering ΔT successive time windows, given a set of graphs $\mathcal{G} = \{G^1, G^2, \ldots, G^{\Delta T}\}$ representing the networks produced for windows $w = 1, \ldots, \Delta T$. Let $\mathcal{S} = \{S_1, S_2, \ldots, S_{\Delta T}\}$ be the set of matrices generated by node2vec for each graph in $\mathcal{G}$, and $V = \{v_1, v_2, \ldots, v_n\}$ be the set of all vertices that appear at least once in any graph in $\mathcal{G}$.

Then, we use DynamicWord2Vec to observe the association of vertices over time according to the sampled walks, mapping them to a temporal ideological latent space. To do this, for each matrix S_w and each pair of vertices $v_i, v_j \in V$ (representing two party members in one of the case studies), we count: (1) the number $\#(v_i)$ of individual occurrences of v_i in the walks represented by rows of S_w ; (2) the number $\#(v_j)$ of individual occurrences of v_j; and (3) the number of co-occurrences of vertices v_i and v_j, restricted within a window of size L from v_i (either before or after v_i), denoted as $\#(v_i, v_j)$. Typically, L is set between 5 to 10 as proposed in [43]. Here, we use L=5, resulting in a window containing 10 vertices in addition to the middle vertex. The degree of association between v_i and v_j is captured by the pointwise mutual information (PMI) [35], defined as a function of the empirical probabilities of occurrences of v_i, occurrences of v_j and co-ocurrences of v_i and v_j in matrix S_w. Specifically, given $|S_w| = k \times l$, the PMI matrix entry corresponding to (v_i, v_j) is given by:

$$\mathrm{PMI}(S_w, L)_{v_i,v_j} = \log_2\left(\frac{\#(v_i, v_j)\cdot|S_w|}{\#(v_i)\cdot\#(v_j)}\right),\ \forall v_i, v_j \in V. \quad (3)$$

When v_i and v_j co-occur very frequently in the sampled walks, the corresponding PMI is high, indicating high proximity between them. On the other hand, when the argument inside $\log_2(.)$ is very small, PMI tends to take on negative values. According to [35, 67], the pairs (v_i, v_j) with more representative association have PMI values greater than 1, that is, they co-occur more than twice in the walks sampled. Thus, considering only the positive values of PMI does not significantly affect the solution while providing better numerical stability to matrix factorization. Thus, given a walk matrix S_w, we define a positive PMI matrix, referred to as PPMI(S_w, L), whose entry for given two vertices v_i and v_j is defined as:

$$\mathrm{PPMI}(S_w, L)_{v_i,v_j} = \max(\mathrm{PMI}(S_w, L)_{v_i,v_j}, 0) \coloneqq Y(w). \quad (4)$$

Given the PPMI matrix $Y(w)$, DynamicWord2Vec learns the embedding vectors $\mathbf{u}_{v_i}$ and $\mathbf{u}_{v_j}$ for vertices v_i and v_j, respectively, by applying a low-rank factorization such that, for any pair v_i and v_j, $\mathbf{u}_{v_i}^\top \mathbf{u}_{v_j} \approx \mathrm{PPMI}(w, L)_{v_i,v_j}$. Each $\mathbf{u}_{v_i}$ has length $d \ll |V|$. Thus, for each time window w, a temporal embedding $U(w)$=$\{\mathbf{u}_{v_1}, \ldots, \mathbf{u}_{v_n}\}$ must satisfy $U(w)U(w)^\top \approx Y(w)$.

This low-rank factorization is obtained by solving an optimization problem. Two regularization terms are added to the objective function in order to address, respectively, overfitting and alignment issues. To avoid overfitting, a typical penalty term based on the Frobenius norm[6] of each low-rank matrix $U(w)$ $(w = 1, \ldots, \Delta T)$ is added [19]. To enforce alignment, a penalty term that assumes some smoothness between subsequent time windows is added. Also, this term is based on the Frobenius norm of the differences between matrices $U(w-1)$ and $U(w)$ for $w = 2, \ldots, \Delta T$. The function to be minimized is

$$\min_{U(1),\ldots,U(T)} \frac{1}{2}\sum_{w=1}^{\Delta T} \|(Y(w) - U(w)U(w)^\top\|_F^2 + \frac{\lambda}{2}\sum_{w=1}^{\Delta T}\|U(w)\|_F^2 + \frac{\tau}{2}\sum_{w=2}^{\Delta T}\|U(w-1) - U(w)\|_F^2, \quad (5)$$

[6] The Frobenius Norm of a given matrix $M_{m\times n}$ is defined by: $\|M\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}$.

where λ, $\tau > 0$. Observe that each embedding $U(w)$ depends, indirectly, on all other $\Delta T - 1$ embeddings. The smoothing term $\|U(w-1) - U(w)\|_F^2$ enforces alignment across embeddings. Parameters λ and τ control the degree of the regularization and smoothness, respectively. Specifically, parameter τ controls the alignment of the embeddings for successive windows w: τ=0 implies no alignment, whereas $\tau \to \infty$ produces a static embedding with $U(1) = U(2) = \ldots = U(\Delta T)$. We discuss how to set parameters λ and τ in Section 7. In order to solve Equation (5), DynamicWord2Vec uses the block coordinate descent [68] obtaining a representation vector $\mathbf{u}_{v_i}(w)$ for each vertex $v_i \in V$ and for each time window w.

Given the embedding vectors, we can compute the change of a given member v_i in the defined ideological space between two time windows w_1 and w_2 using a metric of distance between vectors. We here use the widely adopted cosine distance:

$$\cos(v_{i(w_1)}, v_{i(w_2)}) = 1 - \frac{\mathbf{u}_{v_i}(w_1)\cdot\mathbf{u}_{v_i}(w_2)}{\|\mathbf{u}_{v_i}(w_1)\|\|\mathbf{u}_{v_i}(w_2)\|}. \quad (6)$$

Cosine distance ranges from 0 to 1. Values close to 0 indicate that the two vertices $\mathbf{u}_{v_i}(w_1)$ and $\mathbf{u}_{v_i}(w_2)$ coincide, i.e., the corresponding party member did not change ideologically between windows w_1 and w_2. Values close to 1 indicate that the member drastically shifted his ideology within the period.

In the next four sections, we discuss the results of our analyses when tackling the research questions posed in Section 1.

4 Identifying Ideological Communities

We start by tackling our first research question (RQ1) and characterizing the ideological communities discovered in both Brazilian and US networks. Table 2 shows an overview of all networks for both countries, presenting some topological properties [20], i.e., numbers of vertices (*# of nodes*) and edges (*# of edges*), number of connected components (*# of CC*), average shortest path length (*SPL*), average degree, clustering coefficient and density[7]. The difference between the number of nodes in this table and the number of members in Table 1 corresponds to nodes that were removed after the edge filtering.

Table 2 also summarizes the characteristics of the ideological communities identified using the Louvain algorithm. In the four rightmost columns, it presents the number of communities identified, their *modularity* (*Mod.*) as well as average and standard deviation of the party discipline (*Avg PD* and *SD PD*), computed with respect to the ideological communities.

Starting with the Brazilian networks (top part of Table 2), we can observe great fluctuation in most topological metrics over the years, but, overall, the networks are sparse: the average shortest path length is long, the average clustering coefficient is moderate and the network density is low. Also, the number of communities identified is much smaller than the

[7] The *density* of a network is the ratio of the total number of existing edges to the maximum possible number of edges in the graph. The *clustering coefficient* measures the degree at which nodes tend to group together to form triangles, and is defined as the ratio of the number of existing closed triplets to the total number of open and closed triplets. A triplet is three nodes that are connected by two (open triplet) or three (closed triplet) undirected ties.

Table 2: Statistics of Networks and Ideological Communities (CC: connected components, SPL: shortest path length, Mod: modularity)

Year	Brazil											US										
	# of Nodes	# of Edges	# of CC	Avg. SPL	Avg. Degree	Avg. Clust.	Density	# of Comm.	Mod.	Avg. PD(%)	SD PD	# of Nodes	# of Edges	# of CC	Avg. SPL	Avg. Degree	Avg. Clust.	Density	# of Comm.	Mod.	Avg. PD(%)	SD PD
2003	342	9329	5	1.83	55.01	0.65	0.16	8	0.11	95.48	2.22	431	41892	2	1.11	194.39	0.95	0.45	2	0.48	93.60	1.03
2004	326	7079	2	1.90	43.43	0.62	0.13	4	0.14	92.68	3.36	426	40928	2	1.10	192.15	0.95	0.45	2	0.48	92.97	0.55
2005	359	7211	1	3.18	40.17	0.59	0.11	5	0.21	88.32	3.64	431	41892	2	1.10	194.39	0.95	0.45	2	0.48	92.60	0.79
2006	419	8613	1	2.47	41.11	0.61	0.09	4	0.36	90.50	2.36	426	41112	2	1.10	193.01	0.95	0.45	2	0.49	91.45	0.33
2007	427	11394	3	1.77	53.37	0.67	0.12	6	0.14	95.97	1.26	414	38471	2	1.12	185.85	0.94	0.45	2	0.44	91.55	3.78
2008	400	10180	2	1.62	50.90	0.70	0.12	5	0.08	95.78	1.94	424	40729	2	1.11	192.12	0.94	0.45	2	0.46	95.45	1.97
2009	434	10784	2	1.92	49.70	0.66	0.11	4	0.18	91.45	3.49	429	41698	2	1.15	194.40	0.94	0.45	2	0.40	93.86	2.42
2010	446	10151	1	2.42	45.52	0.64	0.10	4	0.19	92.01	1.29	420	39969	1	3.06	190.33	0.95	0.45	3	0.43	94.92	1.86
2011	408	11519	2	1.89	56.47	0.60	0.13	6	0.12	93.69	3.76	426	41119	2	1.18	193.05	0.96	0.45	3	0.44	90.31	1.91
2012	345	6527	3	2.47	46.11	0.48	0.11	4	0.33	87.00	4.25	417	40545	3	1.17	194.46	0.96	0.46	3	0.44	91.63	1.86
2013	449	10094	1	2.21	44.96	0.61	0.10	4	0.38	86.51	4.18	423	40921	2	1.11	193.48	0.96	0.45	2	0.47	93.23	1.03
2014	450	10036	1	2.18	44.60	0.58	0.09	3	0.43	91.14	1.79	418	40735	2	1.08	194.90	0.96	0.46	2	0.48	94.37	0.34
2015	490	12563	1	2.90	51.28	0.69	0.10	5	0.60	85.90	3.11	427	41890	2	1.09	196.21	0.95	0.46	2	0.47	94.40	1.36
2016	425	10159	2	1.44	47.81	0.66	0.11	4	0.38	92.62	1.83	423	40927	2	1.11	193.51	0.95	0.45	2	0.48	94.70	1.36
2017	396	9434	4	1.64	47.65	0.72	0.12	6	0.24	90.25	3.16	423	40928	2	1.09	193.51	0.95	0.45	2	0.46	96.02	0.44

total number of parties (see Table 1) confirming the fragmentation and ideological overlap of multiple parties. Yet, the *party discipline* of these communities is, on average, very close to, and, in some cases, slightly larger than the values computed for the individual parties, despite a somewhat greater standard deviation observed across communities. Thus, these communities are indeed very cohesive in their voting patterns.

In contrast, the topological structure of the identified communities, as expressed by the *modularity* metric, is very weak, especially in the former years. That is, there is still a lot of similarity across members of different ideological communities. We note that in the former years the government had greater support from most parties, as their members voted similarly in most sessions. Such approval dropped during a period of political turmoil that started in 2012, when the distinction of ideologies and opinions become more clear[8,9]. This may explain why the *modularity* starts low and increases in the most recent years, when there is greater distinction between different communities. This occurs despite the large average party discipline maintained by the communities. Thus, these two metrics offer complementary interpretations of the political scenario.

Turning our attention to the US (bottom part of Table 2), we note that, unlike in Brazil, most metrics remain roughly stable throughout the years. The networks are much more dense, with higher average clustering (Avg. Clust.) coefficient and density and shortest path length. The number of identified communities coincides with the number of connected components as well as with the number of political parties (see Table 1) in most years. These communities are more strongly structured, despite some ideological overlap, as expressed by moderate-to-large *modularity* value. They are also consistent in their ideologies, as expressed by large party disciplines, comparable to the original (party-level) ones. These metrics reflect the political behavior of a non-fragmented and stronger two-party system, quite unlike the Brazilian scenario.

In sum, in Brazil, the several parties can be grouped into just a few ideological communities, with strong disciplined members, although the separation between communities is not very clear. In the US, on the other hand, ideological communities are more clearly defined, both structurally and ideologically, though some inter-community similarity still remains.

5 Identifying Polarized Communities

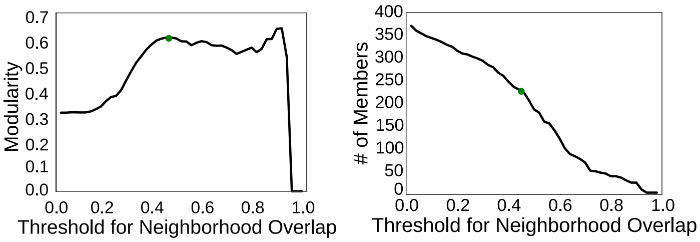
As mentioned, the ideological communities identified in the previous section still share some similarity, particularly for the Brazilian case. In this section, we address our second research question (RQ2), with the aim of identifying polarized communities, i.e., communities that have a more clear distinction from the others in terms of voting behavior. To that end, we take a step further and consider that members of the same polarized community should not only be neighbors (i.e., similar to each other) but should also share most of their neighbors. Thus two members that, despite voting similar to each other, have mostly distinct sets of neighbors should *not* be in the same group.

To identify polarized communities, we start with the networks used to identify the ideological communities and compute the *neighborhood overlap* for each edge. The neighborhood overlap of an edge (v_i, v_j) is the ratio of the number of nodes that are neighbors of both v_i and v_j to the number of neighbors of at least one of v_i or v_j [20]. The neighborhood overlap of v_i and v_j is taken as an estimate of the strength of the tie between the two nodes. Edges with tie strength below a given threshold are considered as *weak* ties, whereas edges with tie strength above that threshold are classified as *strong* ties. We consider that weak ties come from overlapping communities, and strong ties are edges within a polarized community. Thus, edges representing weak ties are removed. As before, nodes that become isolated after this new filtering are also removed.

Once again the selection of the best neighborhood overlap threshold is not straightforward as it involves a complex trade-off: larger thresholds lead to more closely connected communities and higher *modularity*, which is the goal, but also produce sparser graphs, resulting in a larger number of isolated nodes, which are disregarded. Thus, for each network, we selected a threshold that produced a good trade-off between the two metrics (i.e., the lowest threshold yielding modularity close to maximum). Figure 2 shows an example of this trade-off for one specific year (2017) in Brazil, with the selected threshold value shown in green. For Brazil, the selected threshold fell between 0.40 and 0.55, while for the US it was from 0.1 to 0.28. We

[8] https://www.vox.com/2016/4/21/11451210/dilma-rousseff-impeachment

[9] http://www.bbc.com/news/world-latin-america-19359111



(a) *Modularity* (b) Number of Members

Figure 2: Modularity values for different thresholds choices on Brazil's 2017 data. Green dot indicates selected threshold, 0.42.

then re-executed the Louvain algorithm to detect (polarized) communities in the new networks.

Table 3 presents the topological properties of the networks as well as the structural and ideological properties of the identified polarized communities, for both Brazil and US. Focusing first on the Brazilian networks (top part of the table), we see that the number of nodes with strong ties decreases drastically (by up to 66%) as compared to the networks analyzed in Section 4. This indicates the large presence of House members that, despite great similarity with other members, are not strongly tied (as defined above) to them, and thus do not belong to any polarized community. The number of connected components dropped for some years and increased for others, suggesting that some components in the first set of networks were composed of structurally weaker communities or of multiple smaller communities. Network density, average shortest path length, and clustering coefficient also dropped, indicating sparser networks, as expected.

The number of polarized communities somewhat differs from the number of communities obtained when all (strong and weak) ties are considered, increasing in most years. This suggests that some ideological communities identified in Section 5 may be indeed formed by multiple more closely connected subgroups. Yet, those numbers are still smaller than the number of parties in each year (Table 1). Moreover, compared to the ideological communities first analyzed, the polarized communities are stronger both structurally and ideologically, as expressed by larger values of *modularity* and average party discipline.

For the US case, the numbers in Table 3 are very similar to those in Table 2. Less than 2% of the nodes have only weak ties and were removed from the networks in all years. Thus, almost all members have strong ties to each other, building ideological communities that are, in general, very polarized.

In sum, despite the fragmented party system, polarization can be observed in Brazil, to some degree, in a number of smaller strongly tied communities. In the US, on the other hand, almost all members and communities are very polarized.

6 Temporal Analysis of Polarized Communities

We now turn to RQ3 and investigate how the polarized communities evolve over time. To that end, we compute two complementary metrics, namely *persistence* and normalized mutual information [63, 65], for each pair of consecutive years.

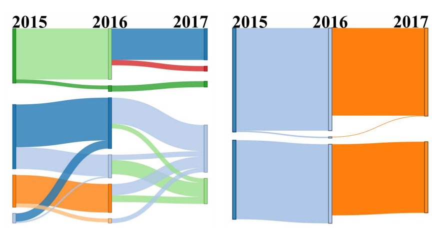
We define the **persistence** from year x to x+1 as the fraction of all members of polarized communities in x who remained in some polarized community in x+1. A *persistence* equal to 100% implies that all members of polarized communities in x remained in some polarized community in x+1. Yet, the membership of individual communities may have changed as members switched communities. To assess the extent of change in community membership over consecutive years, we compute the normalized mutual information (NMI) over the communities, taking only members who persisted over the two years.

NMI is based on Shannon entropy of information theory [60]. Given two sets of partitions X and Y, defining community assignments for nodes, the mutual information of X and Y can be thought as the informational "overlap" between X and Y, or how much we learn about Y from X (and about X from Y). Let $P(x)$ be the probability that a node picked at random is assigned to community x, and $P(x,y)$ the probability that a node picked at random is assigned to both x in X and y in Y. The NMI of X and Y is defined as:

$$NMI(X,Y) = \frac{\sum_x \sum_y P(x,y) \log \frac{P(x,y)}{P(x)P(y)}}{\sqrt{H(X)H(Y)}} \tag{7}$$

where $H(X) = -\sum_x P(x) \log P(x)$ is the Shannon entropy for X. NMI ranges from 0 (all members changed their communities) to 1 (all members remained in the same communities).

Table 4 shows *persistence* (*Pers*) and NMI results for all pairs of consecutive years and both countries. For Brazil (BR), the values of *persistence* varied over the years, ranging from 46% to 80%. Thus, a significant number of new nodes join polarized communities every year. Indeed, in most years, roughly half of the members of polarized communities are newcomers. The values of NMI are also small, especially in the earlier years, reflecting great change also in terms of nodes switching communities. This is consistent with a period of less clear distinction between the communities and weaker polarization, as discussed in the previous sections. Since 2012, the values of NMI fall around 0.6, reflecting greater stability in community membership. For the US, in contras, both *persistence* and NMI are very large, approaching the maximum of 1. Almost all members persist in their polarized communities over the years.



(a) *Brazil* (b) *United States*

Figure 3: Dynamics of Polarized Communities over 2015-2017.

A visualization of some of these results is shown in Figure 3 which presents the flow of nodes across polarized communities over the years of 2015 to 2017 in Brazil and in the US. Each vertical line represents a community, and its length represents the number of members belonging to that community who persisted in some polarized community in the following year. Thus, communities for which all members did not persist in any polarized community in the following year are not represented in the figure. Recall that, according to Table 3,

Table 3: Statistics of Strongly Tied Networks and Polarized Communities (CC: connected comp., SPL: shortest path length, Mod: modularity)

Year	Brazil											US										
	# of Nodes	# of Edges	# of CC	Avg. SPL	Avg. Degree	Avg. Clust.	Density	# of Comm.	Mod.	Avg. PD(%)	SD PD	# of Nodes	# of Edges	# of CC	Avg. SPL	Avg. Degree	Avg. Clust.	Density	# of Comm.	Mod.	Avg. PD(%)	SD PD
2003	186	1436	1	1.48	15.44	0.38	0.08	4	0.35	97.78	0.86	431	41872	2	1.11	194.30	0.95	0.45	2	0.47	93.60	1.03
2004	154	866	1	1.52	11.25	0.33	0.07	5	0.36	97.11	0.57	426	40741	2	1.12	191.27	0.95	0.45	2	0.48	92.97	0.55
2005	119	1210	2	1.19	20.34	0.59	0.17	4	0.37	95.40	0.93	431	41886	2	1.11	194.37	0.95	0.45	2	0.47	92.60	0.79
2006	136	590	10	1.37	8.68	0.52	0.06	12	0.57	96.62	2.16	426	41073	2	1.10	192.83	0.95	0.45	2	0.48	91.45	0.33
2007	175	977	3	1.68	11.17	0.32	0.06	6	0.44	97.31	1.36	414	38462	2	1.12	185.81	0.94	0.44	2	0.42	91.55	3.78
2008	216	1019	2	1.94	9.44	0.23	0.04	5	0.42	97.11	0.46	423	40708	2	1.11	192.47	0.95	0.45	2	0.43	95.49	1.93
2009	209	1217	1	1.30	11.65	0.41	0.05	5	0.56	94.57	1.67	428	41690	2	1.15	194.81	0.94	0.45	2	0.40	93.89	2.45
2010	225	726	6	1.45	6.45	0.22	0.02	11	0.51	94.31	1.80	418	39958	2	1.13	191.19	0.95	0.45	3	0.43	94.86	1.97
2011	250	1891	1	1.78	15.13	0.31	0.06	4	0.40	96.56	0.86	422	41112	2	1.15	194.84	0.97	0.46	3	0.45	90.01	3.16
2012	145	1151	3	1.84	29.82	0.48	0.11	6	0.37	94.42	1.98	413	40529	2	1.07	196.27	0.97	0.47	3	0.44	91.70	2.17
2013	318	4437	5	1.77	27.91	0.58	0.08	9	0.47	91.30	2.17	421	40910	2	1.10	194.35	0.96	0.46	2	0.46	93.32	0.94
2014	287	1672	3	1.37	11.65	0.41	0.04	5	0.63	94.04	1.28	417	40717	2	1.08	195.29	0.96	0.46	2	0.48	94.40	0.38
2015	372	6290	6	1.41	33.82	0.64	0.09	9	0.64	93.93	1.70	424	41759	2	1.08	196.98	0.95	0.46	2	0.47	94.53	1.41
2016	269	1726	3	1.43	12.83	0.44	0.04	8	0.63	95.08	1.21	418	40890	2	1.08	195.65	0.96	0.46	3	0.46	95.67	0.80
2017	227	1631	5	1.58	14.37	0.44	0.06	6	0.60	95.25	2.01	421	40923	2	1.08	194.41	0.95	0.46	2	0.48	95.37	0.11

Table 4: Temporal Analysis of Polarized Ideological Communities (NMI: normalized mututal information)

Consecutive Years	Brazil		United States	
	Persistence	NMI	Persistence	NMI
2003 - 2004	58.24%	0.14	98.13%	0.97
2004 - 2005	46.30%	0.16	90.80%	0.97
2005 - 2006	53.04%	0.20	98.36%	1.00
2007 - 2008	68.26%	0.22	97.57%	1.00
2008 - 2009	63.80%	0.18	86.74%	1.00
2009 - 2010	61.38%	0.26	96.24%	0.94
2011 - 2012	80.08%	0.14	96.18%	0.96
2012 - 2013	67.87%	0.59	96.76%	0.80
2013 - 2014	61.23%	0.56	97.85%	1.00
2015 - 2016	57.85%	0.65	97.63%	0.97
2016 - 2017	57.47%	0.58	86.26%	0.98

the number of polarized communities in Brazil in 2015, 2016 and 2017 was 9, 8 and 6, respectively. A cross-analysis of these results with Figure 3a indicates that members of only 4 out of 9 polarized communities in 2015 persisted polarized in the following year. Moreover, two polarized communities in 2016 were composed of only newcomers and both communities disappeared in 2017 (as they do not appear in the figure). Similarly, one polarized community in 2017 was composed of only newcomers. The figure also shows a great amount of switching, merging and splitting across communities over the years. Figure 3b, on the other hand, illustrates the greater stability of community membership in the US.

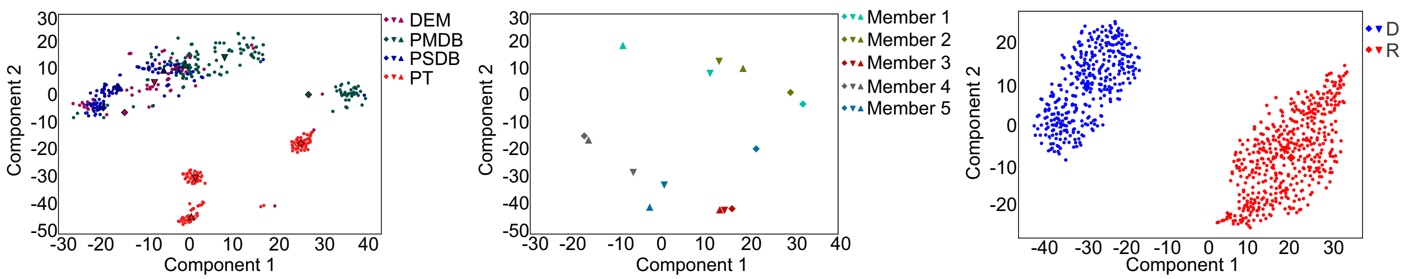
7 Evaluating Ideological Changes

We now turn to our final analyses of changes in ideological behavior. We employ the strategy described in Section 3.4 to model an ideological space and track individual members over time in this space. For analysis purposes, we focus on changes during a period ΔT equal to the duration of a term of office (time period during which elected members should serve), divided in yearly time windows. In Brazil, party members are elected for a 4-year term (named *legislature*), whereas in the US, they are elected for a 2-year period (called *Congress*,).

We start by defining, for each case study and each term of office specified in Table 1, a corresponding sequence of graphs $\mathcal{G} = \{G^1, G^2, \ldots, G^{\Delta T}\}$ representing the networks produced for windows $w = 1, \ldots, \Delta T$, as described in Section 3.3. For each such sequence $\mathcal{G}$ we then produce a single latent ideological space following the method in Section 3.4. Specifically, for each window w (year), we obtain a matrix of embedding vectors $U(w) = \{\mathbf{u}_{v_1}(w), \mathbf{u}_{v_2}(w), \ldots, \mathbf{u}_{v_n}(w)\}$ where $V = \{v_1, v_2, \ldots, v_n\}$ is the set of vertices in $\mathcal{G}$. Recall that our model is robust to missing values, allowing us to infer an ideological representation of a member v_i in w from $(w-1)$ and $(w+1)$. Nevertheless, we choose to include in V only members who appeared in $\mathcal{G}$ in at least two years. This choice is based on a conservative approach to improve robustness, particularly for the Brazilian case, which, as already discussed, has greater instability and a longer term of office (4 years).

We train our ideological space model for a given term of office by carrying out a grid search to determine the best values of parameters λ and τ, as proposed in [67]. We consider various combinations of parameter values, varying λ in [0;100] and τ in [0;100]. For each combination, we first generate our latent space model and the vertice representations (embeddings) for each window w. We then evaluate the goodness of these embeddings (and correspondingly of the generated latent space) as follows. We apply the spherical k-means algorithm [8], which uses cosine similarity as distance metric, to group the vertice embeddings produced for window w, $\mathbf{u}_v(w)$, into k clusters, where k is the number of ideological communities detected for the same window w (see Section 4). We then calculate the Normalized Mutual Information (NMI) (Equation 7) between the ideological communities and the clusters detected by the spherical k-means on the embeddings yielded by our model. The most representative latent space model (i.e., the best parameter values) is the one that best recovers the originally defined ideological groups, thus yielding a higher NMI result.

The results of the grid search were very consistent across most terms of office for each case study. For the US, the same values of $\lambda = 15$ and $\tau = 20$ were found to be the best in all cases. Also, the NMI values were very high, being at least 0.97 and very often reaching the maximum of 1, reflecting the clear structural and ideological separation of the networks. For Brazil, the best values are $\lambda = 5$ and $\tau = 10$ for all but the last term of office for which the best parameter values are $\lambda = 10$ and $\tau = 5$. The NMI values are lower than in the US, yet still reasonably high, especially in the most recent terms (the NMI reached 0.85 in the last 55^{th} legislature), reflecting the stronger community structure and more clear ideological separation of



(a) Brazil, 55th Leg. (top-4 parties) (b) Brazil, 55th Leg. (5 selected members) (c) United States 114th Congress

Figure 4: 2-D representation of latent ideological space. Symbols $\diamond$, $\triangledown$ and $\triangle$ represent party centroids (a,c) or members (b) respectively during 1st, 2nd and, in case of Brazil, 3rd year of legislature as well.

party members in the period (as discussed in Section 4).

For the sake of visualization, we select one example term of office from each case study and plot the generated latent ideological space in a low-dimensional 2-D view using the t-distributed Stochastic Neighbor Embedding (t-SNE) [37]. Figures 4a and 4c show the representations obtained for the 55^{th} legislature in Brazil[10] and the 114^{th} congress in the United States, respectively. Each color corresponds to one political party and each point corresponds to one party member in one of the covered years in the latent ideological space. For the sake of graph readability, Figure 4a shows only the four brazilian parties (the largest ones). Also, the figures do not distinguish between different members of the same party nor different locations of the same member over the years (in case the member changed position over time): all of them are represented by points of the same color. Yet, to illustrate changes in ideological behavior, we plot the centroids of each party, distinguishing its location in each year of the analyzed term by using different representations. Each centroid is represented by a *diamond* in the first year, by an *upside down triangle* in the second year, and, in case of Brazil, by a regular *triangle* in the third year.

We further illustrate changes in individual ideological behavior by focusing now on 5 selected Brazilian party members. Figure 4b shows the locations of their corresponding vertice embeddings in the years of the 55^{th} legislature in the same low-dimensional 2-D view. Each member is shown in a different color and, once again, we use a *diamond*, an *upside down triangle* and a regular *triangle* to represent their locations in the first, second and third years, respectively.

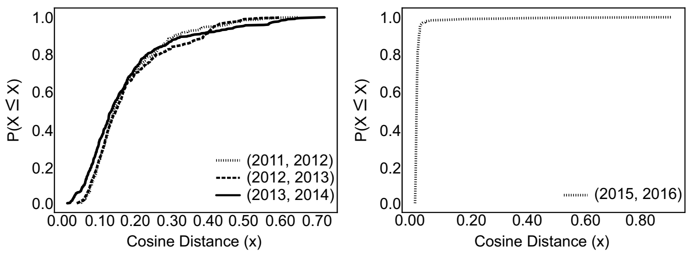
As Figure 4a shows, all four Brazilian parties have changed their ideological behavior over the years, as illustrated by the changes in the locations of their centroids in the ideological space. However, some of them remained quite cohesive throughout the period, that is, the changes were mostly in group. For example, despite individual changes, the distinction between the Work Party (PT) and the Brazilian Social Democracy Party (PSDB), represented in red and blue, respectively, is clear in all three years. These two parties have faced each other for over twenty years in the presidential elections in Brazil. In any of the years, the cosine distance of any given two members (one from each party) is close to 1, indicating great ideological distinction. On the other hand, the distance between any two members from the same party tends to be close to 0, indicating strong ideological alignment. Another interesting example is the Brazilian Democratic Movement Party (PMDB) which started the 55^{th} legislature ideologically aligned with PT, but approached the opposition, composed of PSDB and DEM (among others), as the years went by. The change in Figure 4a reflects what happened in reality as the second government of president Dilma Rousseff (PT) started with the support of PMDB. However, the party of the vice-president decreased its support to the left-wing president and shifted towards center, more aligned with PSDB and DEM. Such movement, which was replicated by other supporting parties, culminated with the presidential impeachment in 2016.

The changes in individual ideological behavior over the three analyzed years can be more clearly visualized in Figure 4b. Note that some members, such as member 3, exhibit very small changes in the ideological space, whereas others have a much more dynamic behavior, falling into different regions of the space over the years. Also note that while some seem to be converging to the same region of the space (e.g., members 4 and 5), others are drifting away (e.g., member 1). In contrast, Figure 4c shows that, in the US, party members have quite stable and distinct ideological behavior.

We delve deeper into the Brazilian scenario, by looking into the 54^{th} legislature, encompassing years 2011-2014, which, as already mentioned, consisted of a period of great political turmoil during which ideological distinctions became more clear. Using the temporal embedding obtained for this term of office, we compute, for each individual member, her *ideological shift*, i.e., the cosine distance between her embedding representations, in consecutive years. Figure 5a shows the cumulative distributions of the ideological shift of individual members for each pair of consecutive years. The three distributions are similar, but we can see a trend towards larger distances in the more recent years, in alignment with our discussion in the previous sections. Also, although most members exhibit some ideological shift over consecutive years, there is great variability across members. For comparison purposes, Figure 5b shows the distribution for the two years of the 114^{th} US congress, when practically all members remained unchanged, confirming the consistency of ideologies over time.

The greater variability in the Brazilian case can be explained, to some extent, by the heterogeneity in ideological behavior between polarized and non-polarized party members. To further analyze this issue, we separate, for each pair of consecutive years, the party members into *polarized*, i.e., members

[10] Recall that our dataset covers only 3 years of 55^{th} legislature.



(a) Brazil (b) United States

Figure 5: CDF of ideological shift of members over consecutive years (measured w.r.t. cosine distance).

who persisted in some polarized community in the two years (as in Table 4), and the other, *non-polarized members*. Figures 6a and 6b show the distributions of ideological shift for each group and for each pair of consecutive years of the 54^{th} legislature. Clearly, non-polarized party members exhibit much greater changes (larger cosine distances), while polarized members do exhibit a more consistent ideology over time.

Yet, even polarized members do experience changes over time, which indirectly affect the membership of the polarized communities. Indeed, as already discussed in Section 5, polarized members often switch between polarized communities, especially in the earlier years. To get a hint of the extent to which such polarized members shift in the ideological space but still remain polarized, we compare them with members who started polarized but left the polarized ideological communities (i.e., became non-polarized) in the following year. Figure 6c shows the distributions of ideological shift for the latter for the same period. Clearly, members who ceased being polarized tend to exhibit greater changes in ideological behavior. Thus, the changes in membership of polarized communities are mostly due to members changing to nearby (polarized) communities, as they slowly shift in the ideological space. Moreover, as shown in Figure 6c, the shift in the ideological space of members who ceased being polarized decreases over the years. This suggests that the polarized communities and, thus, their members, strengthen polarization over time.

8 Conclusions and Future Work

We have proposed a methodology to analyze the formation and evolution of ideological and polarized communities in party systems, applying it to two strikingly different political contexts, namely Brazil and the US. Our analyses showed that the large number of political parties in Brazil can be reduced to only a few ideological communities, maintaining their original ideological properties, that is well disciplined communities, with a certain degree of redundancy. These communities have distinguished themselves both structurally and ideologically in the recent years, a reflection of the transformation that Brazilian politics has been experiencing since 2012. For the US, the country's strong and non-fragmented party system leads to the identification of ideological communities in the two main parties throughout the analyzed period. However, there are still some highly similar links crossing the community boundaries. Moreover, for some years, a third community emerged, without however affecting the strong discipline, ideology and community structure of the American party system.

We then took a step further and focused on polarized communities by considering only tightly connected groups of nodes. We found that in Brazil, despite the party fragmentation and the existence of some degree of similarity even across the identified ideological communities, it is still possible to find a subset of members that organize themselves into strongly polarized ideological communities. However, these communities are highly dynamic, changing a large portion of their membership over consecutive years. In the US, on the other hand, most ideological communities identified are indeed highly polarized and their membership remain mostly unchanged over the years.

Finally, we delved deeper into the individual ideological behavior of party members by proposing a temporal ideological space model. Based on temporal vertice embeddings, our model allows to analyze the ideological shift of individual members over the period of a term of office. We observed that in Brazil, the vast majority of party members did exhibit some change in ideological behavior over time, though the extent of which varies greatly across members. Whereas members of polarized communities had a somewhat more consistent ideological behavior, members of non-polarized communities fluctuated much more ideologically, especially after 2012. In contrast, the representations of US party members in an ideological space confirm much greater stability over time.

As future work, we intend to extend our analyses to other case studies as well as characterize members in terms of their temporal centrality. We also plan to apply our temporal ideological space model in other domains that require temporal alignment (e.g., node classification, link prediction and graph reconstruction on online social networks).

Acknowledgments

This research is partially funded by the FAPEMIG-PRONEX-MASWeb project – Models, Algorithms and Systems for the Web (process number APQ-01400-14), and by the authors individual scholarships from CNPq, CAPES and FAPEMIG.

References

[1] P. Agathangelou et al. “Understanding Online Political Networks: The Case of the Far-Right and Far-Left in Greece”. In: *Proc. of the International Conference on Social Informatics*. 2017.

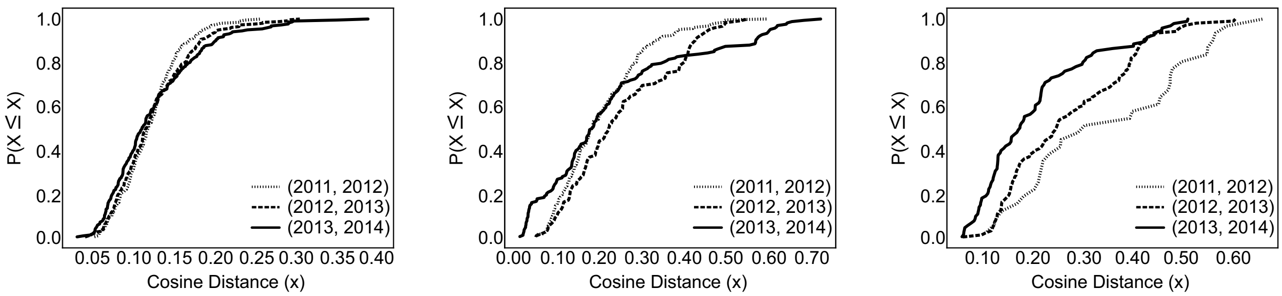
[2] B. Ames. *The Deadlock of Democracy in Brazil*. University of Michigan Press, 2001.

[3] C. Andris et al. “The Rise of Partisanship and Super-Cooperators in the U.S. House of Representatives”. In: *PLOS ONE* 10.4 2015, pp. 1–14.

[4] N. Arinik, R. Figueiredo, and V. Labatut. “Signed Graph Analysis for the Interpretation of Voting Behavior”. In: 2017.

[5] J. Bafumi et al. “Practical Issues in Implementing and Understanding Bayesian Ideal Point Estimation”. In: *Political Analysis* 13 2005, pp. 171–187.

[6] D. Baldassarri and A. Gelman. “Partisans Without Constraint: Political Polarization and Trends in American public opinion”. In: *American Journal of Sociology* 114.2 2008, pp. 408–446.



(a) Members of polarized communities

(b) Non-polarized members

(c) Members who left polarized communities

Figure 6: CDF of ideological shift of Brazilian members over consecutive years (w.r.t. cosine distance) grouped by polarization.

[7] R. Bamler and S. Mandt. "Dynamic Word Embeddings". In: *Proc. of the International Conference on Machine Learning.* 2017.

[8] A. Banerjee et al. "Clustering on the Unit Hypersphere Using Von Mises-Fisher Distributions". In: *Journal of Machine Learning Research* 6 2005, pp. 1345–1382.

[9] D. A. Bateman and L. J. "Ideal points and American Political Development: Beyond DW-NOMINATE". In: *Studies in American Political Development* 30 2016, pp. 147–171.

[10] V. D. Blondel et al. "Fast Unfolding of Communities in Large Cetworks". In: *Journal of Statistical Mechanics: Theory and Experiment* 10 2008.

[11] I. Budge and M. J. Laver. *Party Policy and Government Coalitions.* Springer, 2016.

[12] Q. Cai et al. "A Survey on Network Community Detection Based on Evolutionary Computation". In: *International Journal of Bio-Inspired Computation* 2016, pp. 84–98.

[13] D. Cherepnalkoski et al. "Cohesion and Coalition Formation in the European Parliament: Roll-Call Votes and Twitter Activities". In: *PLOS ONE* 11 2016, pp. 1–27.

[14] J. Clinton, S. Jackman, and D. Rivers. "The Statistical Analysis of Roll Call Data". In: *American Political Science Review* 98.2 2004, pp. 355–370.

[15] M. Coscia and F. M. Neffke. "Network Backboning with Noisy Data". In: *Proc. of the International Conference on Data Engineering.* 2017.

[16] P. Cui et al. "A Survey on Network Embedding". In: *IEEE Transactions on Knowledge and Data Engineering* 31.5 2019, pp. 833–852.

[17] C. Dal Maso et al. "Voting Behavior, Coalitions and Government Strength through a Complex Network Analysis". In: *PLOS ONE* 9 2015, pp. 1–13.

[18] K. Darwish, W. Magdy, and T. Zanouda. "Trump vs. Hillary: What Went Viral During the 2016 US Presidential Election". In: *International Conference on Social Informatics.* 2017.

[19] M. A. Davenport and J. Romberg. "An Overview of Low-Rank Matrix Recovery From Incomplete Observations". In: *Journal of Selected Topics in Signal Processing* 10.4 2016, pp. 608–622.

[20] D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World.* Cambridge University Press, 2010.

[21] V. Eidelman, A. Kornilova, and D. Argyle. "How Predictable is Your State? Leveraging Lexical and Contextual Information for Predicting Legislative Floor Action at the State Level". In: *Proc. of the International Conference on Computational Linguistics.* 2018.

[22] C. H. G. Ferreira, B. de Sousa Matos, and J. M. Almeira. "Analyzing Dynamic Ideological Communities in Congressional Voting Networks". In: *Proc. of the International Conference on Social Informatics.* Springer. 2018, pp. 257–273.

[23] M. P. Fiorina and S. J. Abrams. "Political Polarization in the American Public". In: *Annual Review of Political Science* 11.1 2008, pp. 563–588.

[24] S. Fortunato. "Community Detection in Graphs". In: *Physics Reports* 486 2010, pp. 75–174.

[25] P. Goyal, S. R. Chhetri, and A. Canedo. *dyngraph2vec: Capturing Network Dynamics using Dynamic Graph Representation Learning.* 2018. arXiv: `1809.02657 [cs.SI]`.

[26] P. Goyal et al. "DynGEM: Deep Embedding Method for Dynamic Graphs". In: *Proc. of the International Workshop on Representation Learning for Graphs.* 2018.

[27] A. Grover and J. Leskovec. "node2vec: Scalable Feature Learning for Networks". In: *Proc. of the International Conference on Knowledge Discovery and Data Mining.* 2016.

[28] S. Hix, A. Noury, and G. Roland. "Power to the Parties: Cohesion and Competition in the European Parliament, 1979–2001". In: *British Journal of Political Science* 35.2 2005, pp. 209–234.

[29] X. Huang, J. Li, and X. Hu. "Accelerated Attributed Network Embedding". In: *Proc. of the International Conference on Data Mining.* 2017.

[30] J. J. Jones et al. "Inferring Tie Strength from Online Directed Behavior". In: *PLOS ONE* 8 Jan. 2013, pp. 1–6.

[31] A. Kornilova, D. Argyle, and V. Eidelman. "Party Matters: Enhancing Legislative Embeddings with Author Attributes for Vote Prediction". In: *Annual Meeting of the Association for Computational Linguistics.* 2018.

[32] M. Kusner et al. "From Word Embeddings to Document Distances". In: *International Conference on Machine Learning.* 2015.

[33] B. E. Lauderdale and A. Herzog. "Measuring Political Positions from Legislative Speech". In: *Political Analysis* 24.3 2016, pp. 374–394.

[34] M. Levorato and Y. Frota. "Brazilian Congress structural balance analysis". In: *Journal of Interdisciplinary Methodologies and Issues in Science* 3 2017.

[35] O. Levy and Y. Goldberg. "Neural Word Embedding As Implicit Matrix Factorization". In: *Proc. of the International Conference on Neural Information Processing Systems*. Vol. 2. 2014, pp. 2177–2185.

[36] J. Li et al. "Attributed Network Embedding for Learning in a Dynamic Environment". In: *Conference on Information and Knowledge Management*. 2017.

[37] L. v. d. Maaten and G. Hinton. "Visualizing Data Using t-SNE". In: *Journal of Machine Learning Research* 9 2008, pp. 2579–2605.

[38] S. Mahdavi, S. Khoshraftar, and A. An. "dynnode2vec: Scalable Dynamic Network Embedding". In: *Proc. of the International Conference on Big Data*. 2018.

[39] S. Mainwaring and M. S. Shugart. *Presidentialism and Democracy in Latin America*. Cambridge University Press, 1997.

[40] T. E. Mann and N. J. Ornstein. *It's Even Worse Than it Looks: How the American Constitutional System Collided with the New Politics of Extremism*. Basic Books, 2016.

[41] J. McGee, J. Caverlee, and Z. Cheng. "Location Prediction in Social Media Based on Tie Strength". In: *Proc. of the International Conference on Information & Knowledge Management*. 2013.

[42] I. Mendonça, A. Trouve, and A. Fukuda. "Exploring the Importance of Negative Links Through the European Parliament Social Graph". In: *Proc. of the International Conference on E-Society, E-Education and E-Technology*. 2017.

[43] T. Mikolov et al. "Distributed Representations of Words and Phrases and their Compositionality". In: *Advances in Neural Information Processing Systems*. 2013, pp. 3111–3119.

[44] J. Moody and P. J. Mucha. "Portrait of Political Party Polarization". In: *Network Science* 1.1 2013, pp. 119–121.

[45] M. E. J. Newman. "Modularity and Community Structure in Networks". In: *National Academy of Sciences* 103.23 2006, pp. 8577–8582.

[46] G. H. Nguyen et al. "Continuous-time dynamic network embeddings". In: *Companion of the The Web Conference 2018 on The Web Conference 2018*. 2018.

[47] V.-A. Nguyen et al. "Tea Party in the House: A Hierarchical Ideal Point Topic Model and its Application to Republican Legislators in the 112th Congress". In: *Proc. of the International Joint Conference on Natural Language Processing*. Vol. 1. 2015.

[48] M. Plantié and M. Crampes. "Survey on Social Community Detection". In: *Social Media Retrieval*. Ed. by N. Ramzan et al. Springer London, 2013, pp. 65–85.

[49] K. T. Poole and H. Rosenthal. "A Spatial Model for Legislative Roll Call Analysis". In: *American Journal of Political Science* 29.2 1985, pp. 357–384.

[50] K. T. Poole and H. Rosenthal. *Congress: A Political-economic History of Roll Call Voting*. Vol. 58. 2. Oxford University Press on Demand, 2000.

[51] K. T. Poole and H. Rosenthal. "D-nominate After 10 years: A Comparative Update to Congress: A Political-economic History of Roll-call Voting". In: *Legislative Studies Quarterly* 2001, pp. 5–29.

[52] K. T. Poole and H. L. Rosenthal. *Ideology and congress*. Vol. 1. Transaction Publishers, 2011.

[53] M. A. Porter et al. "A Network Analysis of Committees in the U.S. House of Representatives". In: *National Academy of Sciences* 102.20 2005, pp. 7057–7062.

[54] S. A. Rice. "The Behavior of Legislative Groups: A Method of Measurement". In: *Political Science Quarterly* 40.1 1925, pp. 60–72.

[55] H. Rosenthal. "Introduction to the issue in honor of Keith T. Poole". In: *Public Choice* 176.1 2018, pp. 1–5.

[56] G. Rossetti and R. Cazabet. "Community Discovery in Dynamic Networks: A Survey". In: *ACM Computing Surveys* 51.2 2018.

[57] T. Sakamoto and H. Takikawa. "Cross-national Measurement of Polarization in Political Discourse: Analyzing Floor Debate in the US the Japanese Legislatures". In: *Proc. of the International Conference on Big Data*. 2017.

[58] G. Sartori. *Parties and Party Systems: A framework for analysis*. Vol. 1. ECPR press, 2005.

[59] M. Á. Serrano, M. Boguná, and A. Vespignani. "Extracting the Multiscale Backbone of Complex Weighted Networks". In: *National Academy of Sciences* 106.16 2009, pp. 6483–6488.

[60] C. E. Shannon. "A Mathematical Theory of Communication". In: *ACM SIGMOBILE Mobile Computing and Communications Review* 5 2001, pp. 3–55.

[61] "The Strength of Weak Ties". In: *The American Journal of Sociology* 78.6 1973, pp. 1360–1380.

[62] P. O. S. Vaz de Melo. "How Many Political Parties Should Brazil Have? A Data-Driven Method to Assess and Reduce Fragmentation in Multi-Party Political Systems". In: *PLOS ONE* 10 2015, pp. 1–24.

[63] N. X. Vinh, J. Epps, and J. Bailey. "Information Theoretic Measures for Clusterings Comparison: Variants, Properties, Normalization and Correction for Chance". In: *Journal of Machine Learning Research* 11 2010, pp. 2837–2854.

[64] A. S. Waugh et al. "Party polarization in congress: A network science approach". In: 1 2009.

[65] W. Wei and K. M. Carley. "Measuring Temporal Patterns in Dynamic Social Networks". In: *ACM Transactions on Knowledge Discovery from Data* 10.1 2015, p. 9.

[66] J. Wiese et al. ""You Never Call, You Never Write": Call and SMS Logs Do Not Always Indicate Tie Strength". In: *ACM Conference on Computer Supported Cooperative Work and Social Computing*. 2015.

[67] Z. Yao et al. "Dynamic Word Embeddings for Evolving Semantic Discovery". In: *Proc. of the International Conference on Web Search and Data Mining*. 2018.

[68] H. Yu et al. "Scalable Coordinate Descent Approaches to Parallel Matrix Factorization for Recommender Systems". In: *Proc. of the International Conference on Data Mining*. 2012.

[69] L. Zhu et al. "Scalable Temporal Latent Space Inference for Link Prediction in Dynamic Social Networks". In: *Transactions on Knowledge and Data Engineering* 28.10 2016, pp. 2765–2777.