HAL open science

Brazilian Congress structural balance analysis

Mario Levorato, Yuri Y. Frota

To cite this version:

Mario Levorato, Yuri Y. Frota. Brazilian Congress structural balance analysis. 2016. hal-01359672v1

HAL Id: hal-01359672

https://hal.science/hal-01359672v1

Preprint submitted on 2 Sep 2016 (v1), last revised 13 Apr 2017 (v5)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Brazilian Congress structural balance analysis

Mario Levorato[*1] **and Yuri Frota**[1]

[1]Department of Computer Science, Fluminense Federal University, Brazil

*Correspondance : mlevorato@ic.uff.br

Abstract

In this work, we study the behavior of Brazilian politicians and political parties with the help of clustering algorithms for signed social networks. For this purpose, we extract and analyze a collection of signed networks representing voting sessions of the lower house of Brazilian National Congress. We process all available voting data for the period between 2011 and 2016, by considering voting similarities between members of the Congress to define weighted signed links. The solutions obtained by solving Correlation Clustering (CC) problems are the basis for investigating deputies voting networks as well as questions about loyalty, leadership, coalitions, political crisis, and social phenomena such as mediation and polarization.

Keywords

social network; signed graph; structural balance; Correlation Clustering; metaheuristic; politics

I INTRODUCTION

Structural balance theory is based on the notion of cognitive consistency between friendship and hostility. For example, an enemy of a friend is probably my enemy as well, while a friend of a friend is probably my friend or can become one (Heider, 1946). In simple terms, the interaction of individuals follows the tendency to create stable (albeit not certainly conflict-free) social groups. This can be specially interesting to study similarity and correlation networks, like those originated from common voting patterns, or alliances and disputes among parties or nations (Traag and Bruggeman, 2009; Macon et al., 2012).

One appropriate criterium to measure the degree of balance in signed social networks is by solving the Correlation Clustering (CC) problem (Bansal et al., 2002; Demaine et al., 2006), which consists of partitioning a set of elements into clusters by analyzing the level of similarity between them. It aims to maximize the affinity inside each cluster (i.e. positive relationships) while, at the same time, minimizing the similarities between elements of different clusters (i.e. maximizing negative relationships).

The CC problem, which has been proved to be NP-hard (Bansal et al., 2002), can be applied in several areas, such as efficient document classification (Bansal et al., 2002), natural language processing (Elsner and Schudy, 2009), image segmentation (Kim et al., 2014) and, of course, signed social network analysis (Doreian and Mrvar, 1996). With this objective, the level of balance in a social group can be used by social network researchers to study how (and if) a group evolves to a possible balanced state.

A relaxed version of the CC problem called Symmetric Relaxed Correlation Clustering (SRCC) problem can also be used to evaluate balance in signed social networks. This variant, although

computationally harder to solve, allows the identification of special types of social relationship, such as polatization, mediation and differential popularity (Doreian and Mrvar, 2009), originally viewed as violations of structural balance.

We implemented an algorithm known as $ILS - CC$, which can efficiently solve the aforementioned problems, providing useful information for social network analysis. Using the House of Cunha website (Andrade, 2016) as inspiration, we provide a novel analysis of Brazilian politics inside the Chamber of Deputies (CD). In Brazil, the Chamber of Deputies (Câmara dos Deputados) is the lower house of the National Congress, comprised of 513 federal deputies (from 25 political parties), elected by a proportional representation of votes to serve a four-year term. Based on the CD voting records, we generate several instances of signed social networks, according to certain grouping criteria. The clustering results obtained when invoking the $ILS - CC$ procedure over these instances is the starting point of our study.

The analysis presented in this work can be applied to any network originated from voting patterns, where alliances and interest groups have strong inluence.

This paper is organized as follows. Section II presents a literature review regarding Correlation Clustering problems and signed social network analysis. Section III describes the method applied to extract signed networks from the Chamber of Deputies voting data. Section IV presents an analysis of structural balance on the Chamber of Deputies voting networks, based on the solutions obtained by using our methodology. Finally, we show our conclusions in Section V.

II RELATED WORKS

Heider (1946) was the first to state Structural Balance (SB) theory in order to define sentiment relations among people belonging to the same social group (such as like/dislike and love/hate). Signed graphs were later applied by Cartwright and Harary (1956), formalizing SB theory which affirmed that a stabilized social group could be divided two mutually hostile subgroups (or clusters), each having internal solidarity. Davis (1967) then proposed the more general notion of "weak balance" or clusterable signed graph, when a balanced social group can be divided into two or more mutually antagonistic subgroups, each having internal solidarity.

When solving a clustering problem, one wants to find the most balanced partition[1] of a signed graph. Using structural balance as a measure is equivalent to solving the optimization problem called Correlation Clustering (CC). To our knowledge, this problem was first addressed by Doreian and Mrvar (1996) (although not under this name), who provided a heuristic solution method for analyzing structural balance on real-world social networks. Their method was implemented in software Pajek (Batagelj and Mrvar, 2008). Having a document clustering problem in mind, Bansal et al. (2002) formalized the unweighted version of the CC problem and also discussed its NP-completeness proof. Later, Demaine et al. (2006) addressed the weighted version of the problem. Integer linear programming (ILP) can be used to solve the CC problem optimally, but only if the number of elements is small. Since it consists of a NP-hard minimization problem, the only available solutions for larger instances are either heuristic or approximate. The solution of the CC problem and of some of its variants has already been applied in several areas, such as portfolio analysis in risk management (Harary et al., 2003; Huffner et al., 2010), biological systems (DasGupta et al., 2007; Huffner et al., 2010), grouping of genes (Bhattacharya and De, 2008), efficient document classification (Bansal et al., 2002), detection of embedded

[1] A partition is here defined as the division of the set of vertices V into non-overlapping and non-empty subsets.

matrix structures (Gülpinar et al., 2004), image segmentation (Kim et al., 2014) and community structure (Traag and Bruggeman, 2009; Macon et al., 2012).

In Yang et al. (2007), the CC problem is known as *community mining* and an agent-based heuristic called FEC is proposed to its solution. Genetic algorithms have also been applied to document clustering, using the CC problem as objective function (Zhang et al., 2008). Lately, we presented a Greedy Randomized Adaptive Search Procedure (GRASP) (Feo and Resende, 1995) implementation that provides an efficient solution to the CC problem in networks of up to 8000 vertices (Drummond et al., 2013). Then, based on this method, we introduced sequential and parallel ILS (Iterated Local Search) (Lourenço et al., 2003) procedures for the CC problem(Levorato et al., 2015), which outperformed other solution methods from the literature on three huge real-world signed social networks. In this work, we will use this algorithm to evaluate the imbalance of voting networks.

Apart from the CC problem, alternative measures to structural balance and the associated clustering problems have also been discussed in the literature. In Doreian and Mrvar (2009), the definition of a k-balanced signed graph was informally extended in order to include relevant processes (polarization, mediation, differential popularity and subgroup internal hostility) that were originally viewed as violations of structural balance. For example, the existence of a group of individuals who share only positive relationships with everyone in the network counts as imbalance in the CC Problem. Nonetheless, the individuals in this group could be identified as mediators (i.e. their relations probably won't change over time) and, as pointed in Esmailian et al. (2014), their relations should not be considered as a contribution to the imbalance of the network.

Using this new definition, structural balance was generalized to a version labeled as *relaxed structural balance* (Doreian and Mrvar, 2009). Similarly to the CC problem, measuring the relaxed structural balance can be accomplished through the solution to the Relaxed Correlation Clustering (RCC) problem. It is originally defined on directed graphs (Figueiredo and Moura, 2013); however, a redefinition of relaxed imbalance of a partition P that takes into account only symmetric relationships (i.e. undirected graph) is possible. This gives rise to a new graph clustering problem, the Symmetric Relaxed Correlation Clustering (SRCC) Problem (Figueiredo and Moura, 2013), which will be used in this work. The SRCC problem allows us to analyse mediation processes (positive and negative). That is not the case of the RCC problem, where mediation and differential popularity cannot be pointed out.

It is worth noting that the SRCC problem is closely related with the CC problem but it is not a particular case nor is it a generalization. Actually, each feasible solution (a graph partition) of the SRCC problem is also feasible in the CC problem but the problems have different cost functions, i.e., there are different ways of evaluating the imbalance of a partition. The SRCC problem is intuitively as difficult as the CC problem and is indeed a NP-hard problem (Figueiredo and Moura, 2013).

Two solution methods were initially presented in the literature for RCC problems: a greedy heuristic approach (Doreian and Mrvar, 2009) and a branch-and-bound procedure (Brusco et al., 2011). Computational experiments with both procedures were reported over literature instances with up to 29 vertices and for random instances with up to 40 vertices (Doreian and Mrvar, 2009; Brusco et al., 2011). As far as we know, the $ILS - CC$ algorithm is the only metaheuristic approach that has been applied to solve RCC problems.

Previous works have employed signed graph clustering methods to analyze networks of international alliances and disputes (Traag and Bruggeman, 2009; Macon et al., 2012). In Levorato et al. (2015), by using the $ILS-CC$ algorithm, we presented a historical and geopolitical analysis of the results obtained from the voting on resolutions in the United Nations General Assembly (UNGA). Mendonça et al. (2015) have then applied a parallel version of the $ILS-CC$ algorithm to analyze a collection of signed networks representing voting sessions of the European Parliament. The obtained results were compared to a selection of community detection algorithms designed to process only positive links.

III NETWORK EXTRACTION

In this section, we explain the retrieval of raw voting data, and how we extracted signed networks from it.

3.1 Brazilian Chamber of Deputies

The Chamber of Deputies provides web services which supply information about each of its members, including the vote cast by a specific deputy for each proposition evaluated at the CD. A deputy is described by its name, state (one of 27 Brazilian Federative Units) and political party.

For a given proposition, a deputy can express his vote in either of four ways: Sim (For: the deputy wants the proposition to be accepted), Não (Against: s/he wants the proposition to be rejected), Abstenção (Abstain: s/he wants to express his neutrality) and Obstrução (Filibuster: a form of obstruction, where debate over a proposition is extended, in order to delay or entirely prevent a vote on the proposal).

Besides the previous votes, a deputy may not vote at all, which leads to a fifth vote type: Ausência (Absent: the deputy was not present during the voting session).

The Chamber of Deputies' web services provide raw voting data, which describe the behavior of deputies apart from the others. Nonetheless, since a network is naturally relational (relationships between individuals are the product of their opinion about topics of interest), voting data has to be processed to generate the networks we wish to analyse.

3.2 Extraction algorithm

The procedure starts with a comparison between all pairs of deputies, analyzing the similarity of their voting choices. The obtained measures make up what is known as the agreement matrix M. Each element m_{uv} of this matrix indicates the average agreement between two deputies u and v, in other words, their level of accordance taking into consideration all propositions voted in a given time period.

While filtering the results is a relatively simple task, processing agreement scores may seriously alter the resulting network, depending on the methodology applied. Given a certain pair of deputies u and v and a proposition p_i, the proposition-wise agreement score m_{uv} (p_i) is determined by comparing the votes of both deputies. It ranges from -1 if they fully disagree (one voted FOR and the other AGAINST), to +1 if they entirely agree (they share the same vote: FOR or AGAINST).

But, as previously stated, a voting record may contain, besides FOR and AGAINST, other values which should be equally taken into account. The first case refers to absence of one deputy or both of them (it is worth remembering that the analysis is based on pairs of deputies). The

	FOR	ABSTAIN	AGAINST
FOR	+1	+0.5	-1
ABSTAIN	+0.5	+0.5	+0.5
AGAINST	-1	+0.5	+1

Table 1: Vote weights representing abstention as half an agreement.

	FOR	ABSTAIN	AGAINST
FOR	+1	0	-1
ABSTAIN	0	+1	0
AGAINST	-1	0	+1

Table 2: Vote weights representing abstention as absence of opinion.

general approach is to leave out all propositions p_i that fall into this case (Porter et al., 2005; Dal Maso et al., 2014). But, since certain deputies have low attendance rates, this might lead to distorted agreement or disagreement average scores, due to the small number of common voting sessions. To prevent this, we assume a neutral score of zero if at least one deputy is absent when voting a given proposition.

Abstention is a little more complicated process to understand. For example, if the political party supports a completely different view from the deputy, such pressure may be enough to lead to him/her to take a step towards abstention, despite the fact that s/he is FOR or AGAINST the proposition under analysis. Similarly, abstention may simply represent the deputy's neutral position before a specific topic proposed (whether or not the subject is approved, it makes no difference at all as long as the deputy is concerned). Literature provides different views to deal with ABSTAIN-FOR, ABSTAIN-AGAINST and ABSTAIN-ABSTAIN situations (Macon et al., 2012; Porter et al., 2005; Dal Maso et al., 2014). In this work, we make use of two different ways of calculating the scores. The first one (Table 1) treats abstention as half an agreement whenever it is paired with FOR, AGAINST or other abstention, yielding a value of +0.5. In the second one (Table 2), whenever two deputies abstain at the same time, this is viewed as a full agreement (+1 value). As opposed to that, if only one abstains, a zero score is assigned, since there is not sufficient information to assert they are in agreement or disagreement. So to make things more clear, absence was not included in the tables.

The last case is filibuster, a practice used to create difficulties or hindrances in a systematic way to delay or impede the approval of a bill in parliament. It is normally used by minority groups which do not have the necessary number of representatives to effectively hold back a decision taken by the majority. Therefore, any vote marked as obstruction is here regarded as AGAINST.

The proposition-wise agreement score is fully specified by choosing one of the previous processing strategies. By averaging this score over all considered propositions, the average agreement can be calculated. In a formal way, consider two users u and v, as well as the propositions resulting from the filtering stage: p_1, ..., p_ℓ, for which both u and v voted. The average agreement m_{uv} between these two deputies is:

$$m_{uv} = \frac{1}{\ell} \sum_{i=1}^{\ell} m_{uv}(p_i) \tag{1}$$

We generated one signed graph for each year (from 2011 until June 2016), taking into account all the voting sessions in that year. Graph edges with weight smaller than 0.001 were removed from the graph. The set of vertices in each signed graph represents the list of deputies who voted at least one time in the corresponding year.

IV STRUCTURAL BALANCE ANALYSIS

In this section, based on the clustering results obtained with the ILS-CC algorithm, we investigate some aspects of Brazilian politics in the Chamber of Deputies, including loyalty, leadership, coalition, crisis, as well as social phenomena such as mediation and polarization.

As explained in the previous section, we followed two approaches when generating voting networks for each year in the period between January 2011 and June 2016. We will refer to each network as either v1 or v2, depending on the strategy while dealing with abstentions:

v1 : abstention is worth half an agreement (+0.5), whenever it is paired with any kind of vote (FOR, AGAINST or other abstention);

v2 : abstention is viewed as full agreement (+1 value) only if both deputies abstain. Otherwise, if only one abstains, a zero score is assigned.

All charts and tables used in this analysis are available on-line[2].

4.1 A brief introduction to Brazilian politics

From 1994 to 2002, Brazil was governed by president Fernando Henrique Cardoso, member of the PSDB (Brazilian Social Democracy Party). In 2002, PSDB was defeated in the presidential elections by PT (Brazilian Labor Party) and Lula was elected for a four-year term, being reelected in 2006 for one more period of four years. Then, in 2010, Dilma Rouseff (also a PT member and supported by Lula) won the elections, becoming the next president and, like Lula, was also reelected in 2014 for an additional four-year term.

Nevertheless, due to the intense political and economical crisis, agravated by successive scandals of corruption in the heart of the government, the president was turned away from her post and an impeachment process was started.

In order to help understanding the political groups and parties referenced in the analysis, we first provide a list of the party alliances during the presidential elections held in 2010 (Table 3) and 2014 (Table 4).

Coalition name	Parties
Dilma PT	PT, PRB, PDT, PMDB, PTN, PSC, PR, PTC, PSB, PCDOB
Serra PSDB	PSDB, PTB, PPS, DEM, PMN, PTDOB

Table 3: Coalitions in the 2010 presidential elections.

Coalition name	Parties
Dilma PT	PT, PMDB, PSD, PP, PR, PDT, PRB, PROS, PCDOB
Aécio PSDB	PSDB, SD, PMN, PEN, PTN, PTC, DEM, PTDOB, PTB

Table 4: Coalitions in the 2014 presidential elections.

Another useful piece of information is the list of parties according to their orientation:

[2]Please visit https://public.tableau.com/profile/mario.levorato#!/

Orientation	Parties
Left	PCDOB, PCB, PCO, PSOL, PSTU, PT
Center-left	PDT, PPL, PPS, PROS, PSB
Center	DEM, PHS, PMB, PSD, PSL, PTDOB, PEN, PMDB, PRP, PSDC, PTC, PV, PTN
Center-right	Novo, PRB, PSC
Right	PP

Table 5: List of Brazilian political parties according to their orientation.

Although some parties classify their orientation as center-left or center-right, a great portion of them can be regarded as center parties. As of 2016, the block known as ”super-center” includes PP, PR, PSD, PTB, PROS, PSC, SD, PRB, PEN, PTN, PHS and PSL.

4.2 Methodology

We attempt to identify groups of deputies (and their respective parties) in the Chamber of Deputies signed networks, generated based on voting session records publicly made available by the open data initiative of the Brazilian Government [3].

To do so, we apply an efficient solution method to Correlation Clustering problems, called $ILS - CC$ (Levorato et al., 2015), capable of processing network instances of up to 100 thousand vertices. It is based on the Iterated Local Search (ILS) (Lourenço et al., 2003) metaheuristic. This procedure was used to solve the two problems introduced in Section II: the Correlation Clustering (CC) problem and the Symmetric Relaxed Correlation Clustering (SRCC) problem.

However, based on the obtained results, we chose to rely our analysis solely on SRCC clustering results. The reason is that all CC solutions presented only one or two clusters as output, which, to our knowledge, did not accurately represent the political groups in the Chamber of Deputies. One possible explanation is that, as stated in Section II, when compared to the SRCC problem, the CC problem tends to over-evaluate the imbalance of a network, for penalizing relationships associated, for instance, with mediation processes. As we shall see next, parliament mediation groups were indeed detected when solving the SRCC problem.

Next we present several clustering results that help answering interesting questions concerning political dynamics. Each question and its respective analysis is organized in a subsession.

4.3 To what extent did parties of the same coalition remain loyal to one another?

We have extracted a table which, for each year, coalition and party, gives details about the percentage of deputies in each cluster. This way it is possible to spot if the majority of the deputies of a specific party does not belong to the most populous coalition cluster, which constitutes a strong evidence that such a party is unfaithful to its coalition. By using this data, one can verify that, for example, in 2011 (Table 6), only 41% of PDT, 38% of PR and 42% of PRB deputies were classified inside the largest ruling coalition cluster, formed by 206 deputies. In 2012 (Table 7), only 16% (3 in 19) of PSC deputies accompanied the biggest government group, comprised of 237 deputies. Finally, in 2014 (on both network versions), just half of PT and PDT deputies followed the government coalition (see Tables 8 and 9).

[3] The data services of the Brazilian Chamber of Deputies website can be found at http://www2.camara.leg.br/transparencia/dados-abertos

4.4 Has the government coalition won or lost support?

We start by analyzing two tables that provide the amount of deputies according to their respective party alliance and the cluster to which they belong. The first table (Table 10) refers to the period from 2011 to 2014 (54th legislature of the Chamber of Deputies), while the second one (Table 11) gives information about the years of 2015 and 2016 (55th legislature, corresponding to Mrs. Rouseff's second presidential term).

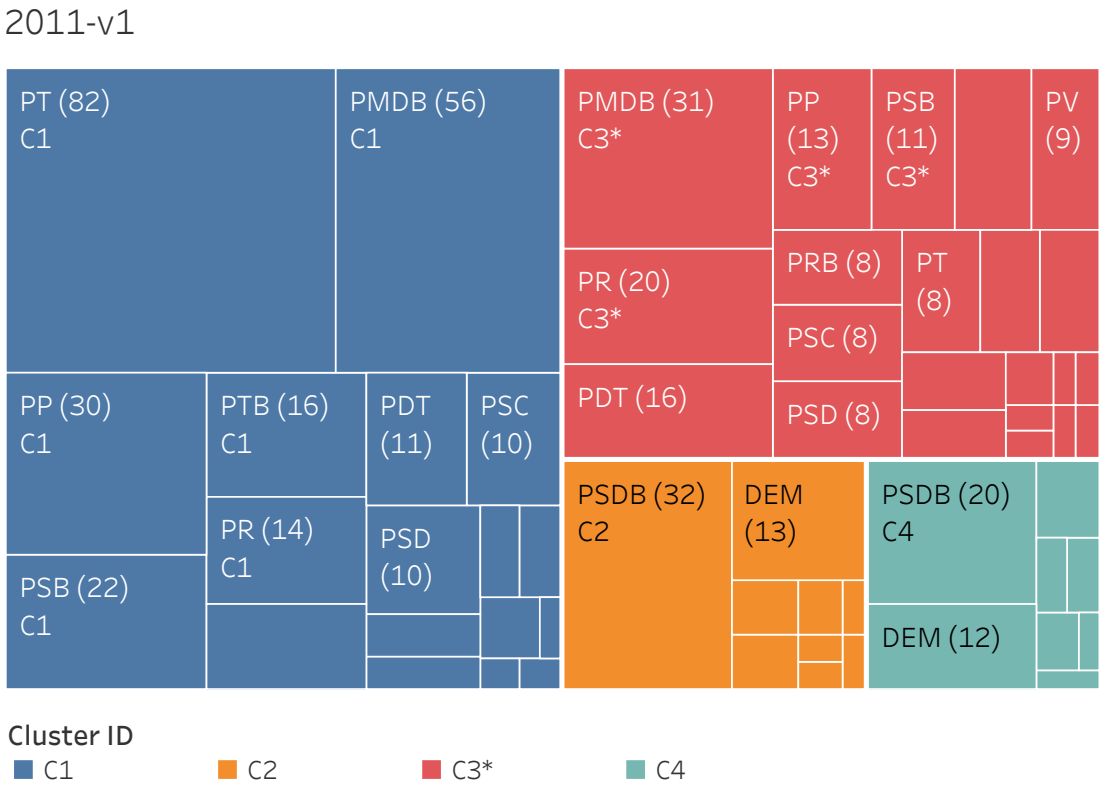

Figure 1: SRCC clustering results for the year of 2011, when using version v1 of the voting network. The treemap shows the parties inside each cluster, the number of deputies of each party in a specific cluster (in parathesis) and also the cluster identifier/label (begins with letter C). Cluster labels marked with an asterisk (*) consist of mediation groups.

We observe that, in the first year of Mrs. Rouseff's government (2011), the government coalition is divided, roughly speaking, in two or three great groups, depending on the network version on which the analysis is based. According to version v1 (Figure 1), the largest cluster (C1) has 64% of the allied deputies. Also, the great majority of the president's party (PT), 82 deputies, are to be found in this cluster.

From 2012 onwards, a clear basis consolidation can be observed, with 77% of the allied deputies in the same group (C1 in both network versions, see Figures 2 and 3). This cluster also holds more than 80 deputies of president's party (PT).

In 2013 (Figure 4), the percentage of allied deputies inside the largest cluster (C1) rises to 82% of the coalition (74 PT deputies). However, in 2014 (the last year of Mrs. Rouseff's first term), a change of course comes about. This measure falls to 66% (Figure 5) and, even worse, only about half of PT's deputies are inside the main coalition group (C3).

A close look at Mrs. Rouseff's second presidential term is surprising. In 2015, the biggest group of what should be the government's new coalition (cluster C1 in Figures 6 and 7) is formed by

70% of the total number of deputies of the coalition as a whole. Notwithstanding, this group houses at most 10 deputies of the president's party (PT). Consider as well that the greatest part of PT deputies is in fact isolated in a smaller cluster, together with a few deputies from less influential parties. Note that both network versions show almost identical results (Figures 6 and 7).

A similar picture takes place in 2016 (Figure 8), when about two thirds of the supposedly allied deputies belong to the same group, which contains only 11 PT deputies. Similarly, 50 PT deputies can be found in another cluster.

Briefly speaking, results point out that in the years of 2015 and 2016, even though there are still large groups in which most deputies are from the so-called governement coalition, such groups are no longer in accordance with the president's party, which is perfectly understandable because of the political crisis and the loss of parliamentary support, news widely broadcast.

4.5 Is party leadership as a whole strong?

This study was carried out as follows: for each year from 2011 to 2016 and for each party, we scanned data about the deputies and the clusters on which they make part. This information was then cross-referenced with the cluster where the leader of the respective party is found. This way it is possible to have a clear view of how strong the leadership of each party is. The full results with the information about the deputies classified in the same cluster as their respective party leader (percentual numbers) are available in Table 12 for 2011-2014 and in Table 13 for 2015-2016.

For each year, the following parties have been identified as having low percentage ($\rho < 50\%$) of deputies who vote after their party leaders, independently of the analyzed network version:

- 2011: PRB, PRP, PSC, PSD.
- 2012: PR, PSB, PSC, PSD, PTB, PV.
- 2013: PCDOB, PRP, PV.
- 2014: PCDOB, PDT.
- 2015: PSC.
- 2016: PCDOB, PMB, PSL, PTDOB.

Deep consideration into this list will reveal that, as we spot a considerably great number of deputies arranged in clusters where their party leaders are not present, there is strong evidence that, on average, voting recomendations from party leaders have not been followed by many deputies.

4.6 Was the algorithm capable of detecting mediation groups in the Chamber of Deputies?

When the ILS-CC algorithm analyzes the voting networks of the Chamber of Deputies, two versions for each year, it proves to be quite successful in identifying positive mediation groups, that means, clusters whose most internal and external relationships are positive. In the conducted study, a group was classified as showing mediation properties whenever its positive relationship percentage was above 90% and also its internal positive link ratio exceeded that same level.

Table 14 lists the mediation clusters, for each year and voting network version. In every figure that contains a treemap (see Figure 5) showing what is inside each cluster (parties and amount of deputies), the information about which groups have mediation properties is also present in the cluster labels marked with an asterisk (*).

4.7 Did the president's political party (PT) split when the Brazilian political crisis began?

The clustering results for 2014 strongly suggest that the president's party split, with 47 deputies in the first cluster, 39 in the second and 8 in the third. As seen on Figure 5, the treemap shows the fragmentation of PT in the last year of the president's first term.

4.8 After Mrs. Dilma Rouseff was reelected, did the government coalition loose support?

According to the results obtained by the ILS-CC algorithm, in 2015, after Mrs. Rouseff's reelection, three clusters cover 99% of the deputies (Figures 6 and 7). The parties inside each cluster reveal the main political groups at that time:

- the largest group includes mainly center parties, such as the majority of PMDB, PSD, PP and PR;
- the second biggest group is formed by opposition parties like PSDB and DEM;
- the last one represents the government core parties, such as PT (59 deputies) and PCDOB (12 members).

A comparison between 2015 and previous years (see Figures 1, 2 and 3) reveals that the government coalition has gone through a substantial loss of support, mostly from center parties.

4.9 When the government coalition lost power, did the center parties move towards opposition?

Looking at the data for the years of 2015 (Figures 6 and 7) and 2016 (Figures 8 and 9), one can observe that the majority of center party and opposition deputies started sharing the same group. There was a strong approximation between PMDB (center), PSDB and DEM (opposition), which have previously been in separate clusters. According to the charts, one can notice that center parties have moved towards opposition.

In 2015, there was a large movement of parties from the governement coalition, which went to a "super-centered" group. These parties include: PROS (12), PRB (12), PDT (22), PR (25), PP (28), PSD (33) and PMDB (71).

In 2016, the following coalition parties have effectively migrated to what can be interpreted as a huge opposition cluster: PDT(17), PRB (20), PSD (21), PP (30), PR (33) and PMDB (56).

4.10 At any time, do numbers show polarization between political groups?

In 2012 (on both network versions), the chamber of deputies is polarized in two large groups (see Figures 2 and 3). The first one with 238 members, led by the majority of PT and PMDB deputies (government base). The other cluster is mainly characterized by opposition parties, such as PSDB and DEM, but it also includes dissidents from center parties like PMDB and PSD.

V CONCLUSION

In this article, we have investigated some of the aspects inherent to signed voting networks and political relationships, by using data from the Brazilian Chamber of Deputies (CD). We have first extracted a collection of networks based on voting patterns of the CD members. We have also applied a clustering algorithm specifically designed for signed networks, called $ILS-CC$, which aims to reduce structural balance.

The analysis of the identified clusters has shown that mediation groups really exist in the Chamber of Deputies. They have been detected in different years. The applied algorithm has allowed us to gather evidence that certain parties are indeed unfaithful to their coalition. Besides, the obtained data perfectly confirms the news broadcast about the Brazilian political situation, such as the loss of support that government coalition experienced.

Equally, the algorithm has proved to be a useful tool to spot parties under weak leadership and the existance of polarization between two large political groups.

References

Andrade N. (2016). *House of Cunha: Who's who in the Brazilian House of Deputies*. http://houseofcunha.com.br/. Accessed: 2016-08-25.

Bansal N., Blum A., Chawla S. (2002). Correlation clustering. In *Proceedings of the 43rd annual IEEE symposium of foundations of computer science*, Vancouver, Canada, pp. 238–250.

Batagelj V., Mrvar A. (2008, 05). *Pajek Wiki*. URL http://pajek.imfm.si/. Accessed on May 2014.

Bhattacharya A., De R. K. (2008). Divisive correlation clustering algorithm (dcca) for grouping of genes: detecting varying patterns in expression profiles. *bioinformatics 24*(11), 1359–1366.

Brusco M., Doreian P., Mrvar A., Steinly D. (2011). Two algorithms for relaxed structural balance partitioning: linking theory, models and data to understand social network phenomena. *Sociological Methods & Research 40*, 57–87.

Cartwright D., Harary F. (1956). Structural balance: A generalization of heider's theory. *Psychological Review 63*, 277–293.

Dal Maso C., Pompa G., Puliga M., Riotta G., Chessa A. (2014). Voting behavior, coalitions and government strength through a complex network analysis. *PloS one 9*(12), e116046.

DasGupta B., Encisob G. A., Sontag E., Zhanga Y. (2007). Algorithmic and complexity results for decompositions of biological networks into monotone subsystems. *BioSystems 90*, 161–178.

Davis J. (1967). Clustering and structural balance in graphs. *Human Relations 20*, 181–187.

Demaine E. D., Emanuel D., Fiat A., Immorlica N. (2006). Correlation clustering in general weighted graphs. *Theoretical Computer Science 361*(2), 172–187.

Doreian P., Mrvar A. (1996). A partitioning approach to structural balance. *Social Networks 18*, 149–168.

Doreian P., Mrvar A. (2009). Partitioning signed social networks. *Social Networks 31*, 1–11.

Drummond L., Figueiredo R., Frota Y., Levorato M. (2013). Efficient solution of the correlation clustering problem: An application to structural balance. In Y. Demey and H. Panetto (Eds.), *OTM 2013 Workshops*, Volume 8186 of *LNCS*, pp. 674–683. Springer Berlin Heidelberg.

Elsner M., Schudy W. (2009). Bounding and comparing methods for correlation clustering beyond ilp. In *ILP'09 Proceedings of the Workshop on Integer Linear Programming for Natural Language Processing*, pp. 19–27.

Esmailian P., Abtahi S. E., Jalili M. (2014). Mesoscopic analysis of online social networks: The role of negative ties. *Physical Review E 90*(4), 042817.

Feo T. A., Resende M. G. (1995). Greedy randomized adaptive search procedures. *Journal of global optimization 6*(2), 109–133.

Figueiredo R., Moura G. (2013). Mixed integer programming formulations for clustering problems related to structural balance. *Social Networks 35*(4), 639–651.

Gülpinar N., Gutin G., Mitra G., Zverovitch A. (2004). Extracting pure network submatrices in linear programs using signed graphs. *Discrete Applied Mathematics 137*, 359–372.

Harary F., Lim M., Wunsch D. C. (2003). Signed graphs for portfolio analysis in risk management. *IMA Journal of Management Mathematics 13*, 1–10.

Heider F. (1946). Attitudes and cognitive organization. *Journal of Psychology 21*, 107–112.

Huffner F., Betzler N., Niedermeier R. (2010). Separator-based data reduction for signed graph balancing. *Journal of Combinatorial Optimization 20*, 335–360.

Kim S., Yoo C. D., Nowozin S., Kohli P. (2014). Image segmentation usinghigher-order correlation clustering. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 36*(9), 1761–1774.

Levorato M., Drummond L., Frota Y., Figueiredo R. (2015). An ils algorithm to evaluate structural balance in signed social networks. In *Symposium on Applied Computing, SAC 2015, Salamanca, Spain - April 13 - 17*, pp. 1117–1122.

Lourenço H. R., Martin O. C., Stützle T. (2003). *Iterated local search*. Springer.

Macon K., Mucha P., Porter M. (2012). Community structure in the united nations general assembly. *Physica A: Statistical Mechanics and its Applications 391*, 343–361.

Mendonça I., Figueiredo R., Labatut V., Michelon P. (2015). Relevance of negative links in graph partitioning: A case study using votes from the european parliament. In *Network Intelligence Conference (ENIC), 2015 Second European*, pp. 122–129. IEEE.

Porter M. A., Mucha P. J., Newman M. E., Warmbrand C. M. (2005). A network analysis of committees in the us house of representatives. *Proceedings of the National Academy of Sciences of the United States of America 102*(20), 7057–7062.

Traag V., Bruggeman J. (2009). Community detection in networks with positive and negative links. *Physical Review E 80*, 036115.

Yang B., Cheung W., Liu J. (2007). Community mining from signed social networks. *IEEE Transactions on Knowledge and Data Engineering 19*, 1333–1348.

Zhang Z., Cheng H., Chen W., Zhang S., Fang Q. (2008). Correlation clustering based on genetic algorithm for documents clustering. In *IEEE Congress on Evolutionary Computation*, pp. 3193–3198.

Year	Version	Party Alliance	Party	Cluster ID C1	C2	C3	C4	Total
2011	v2	Government	PDT	41.38% 12	6.90% 2	27.59% 8	24.14% 7	100.00% 29
			PMDB	67.42% 60	2.25% 2	12.36% 11	17.98% 16	100.00% 89
			PR	37.84% 14	8.11% 3	18.92% 7	35.14% 13	100.00% 37
			PRB	41.67% 5		33.33% 4	25.00% 3	100.00% 12
			PSB	64.71% 22	2.94% 1	14.71% 5	17.65% 6	100.00% 34
			PSC	61.11% 11		22.22% 4	16.67% 3	100.00% 18
			PT	91.11% 82		5.56% 5	3.33% 3	100.00% 90
			PTC				100.00% 1	100.00% 1
			Total	66.45% 206	2.58% 8	14.19% 44	16.77% 52	100.00% 310
		Opposition	DEM	7.89% 3	68.42% 26	7.89% 3	15.79% 6	100.00% 38
			PMN	75.00% 3		25.00% 1		100.00% 4
			PPS		58.33% 7	25.00% 3	16.67% 2	100.00% 12
			PSDB	1.72% 1	91.38% 53	3.45% 2	3.45% 2	100.00% 58
			PTB	72.73% 16		13.64% 3	13.64% 3	100.00% 22
			Total	17.16% 23	64.18% 86	8.96% 12	9.70% 13	100.00% 134
		Total		51.58% 229	21.17% 94	12.61% 56	14.64% 65	100.00% 444

Table 6: Party coalition and clustering details for the year of 2011, when using version v2 of the voting network.

Year	Version	Party Alliance	Party	Cluster ID C1	C2	C3	C4	Total
2012	v2	Government	PDT	60.71% 17	3.57% 1	10.71% 3	25.00% 7	100.00% 28
			PMDB	80.23% 69		10.47% 9	9.30% 8	100.00% 86
			PR	61.11% 22		19.44% 7	19.44% 7	100.00% 36
			PRB	90.00% 9		10.00% 1		100.00% 10
			PSB	81.08% 30		13.51% 5	5.41% 2	100.00% 37
			PSC	15.79% 3		36.84% 7	47.37% 9	100.00% 19
			PT	92.47% 86	1.08% 1	4.30% 4	2.15% 2	100.00% 93
			PTC	100.00% 1				100.00% 1
			Total	76.45% 237	0.65% 2	11.61% 36	11.29% 35	100.00% 310
		Opposition	DEM	16.67% 5		43.33% 13	40.00% 12	100.00% 30
			PMN	100.00% 2				100.00% 2
			PPS	18.18% 2		63.64% 7	18.18% 2	100.00% 11
			PSDB	5.36% 3	7.14% 4	48.21% 27	39.29% 22	100.00% 56
			PTB	57.14% 12		14.29% 3	28.57% 6	100.00% 21
			Total	20.00% 24	3.33% 4	41.67% 50	35.00% 42	100.00% 120
		Total		60.70% 261	1.40% 6	20.00% 86	17.91% 77	100.00% 430

Table 7: Party coalition and clustering details for the year of 2012, when using version v2 of the voting network.

				Cluster ID				
Year	Version	Party Alliance	Party	C1	C2	C3	C4	Total
2014	v1	Government	PDT	63.16% 12	5.26% 1	10.53% 2	21.05% 4	100.00% 19
			PMDB	81.25% 65	5.00% 4	1.25% 1	12.50% 10	100.00% 80
			PR	87.88% 29			12.12% 4	100.00% 33
			PRB	80.00% 8		10.00% 1	10.00% 1	100.00% 10
			PSB	88.46% 23		3.85% 1	7.69% 2	100.00% 26
			PSC	73.33% 11	6.67% 1		20.00% 3	100.00% 15
			PT	51.06% 48		38.30% 36	10.64% 10	100.00% 94
			Total	70.76% 196	2.17% 6	14.80% 41	12.27% 34	100.00% 277
		Opposition	DEM	89.29% 25			10.71% 3	100.00% 28
			PMN	100.00% 3				100.00% 3
			PPS	88.89% 8		11.11% 1		100.00% 9
			PSDB	91.84% 45		2.04% 1	6.12% 3	100.00% 49
			PTB	85.00% 17	5.00% 1		10.00% 2	100.00% 20
			Total	89.91% 98	0.92% 1	1.83% 2	7.34% 8	100.00% 109
		Total		76.17% 294	1.81% 7	11.14% 43	10.88% 42	100.00% 386

Table 8: Party coalition and clustering details for the year of 2014, when using version v1 of the voting network.

				Cluster ID				
Year	Version	Party Alliance	Party	C1	C2	C3	C4	Total
2014	v2	Government	PDT	10.53% 2	5.26% 1	47.37% 9	36.84% 7	100.00% 19
			PMDB	1.25% 1	5.00% 4	77.50% 62	16.25% 13	100.00% 80
			PR			81.82% 27	18.18% 6	100.00% 33
			PRB	10.00% 1		90.00% 9		100.00% 10
			PSB	3.85% 1		61.54% 16	34.62% 9	100.00% 26
			PSC		6.67% 1	80.00% 12	13.33% 2	100.00% 15
			PT	41.49% 39		50.00% 47	8.51% 8	100.00% 94
			Total	15.88% 44	2.17% 6	65.70% 182	16.25% 45	100.00% 277
		Opposition	DEM			75.00% 21	25.00% 7	100.00% 28
			PMN			66.67% 2	33.33% 1	100.00% 3
			PPS	11.11% 1		66.67% 6	22.22% 2	100.00% 9
			PSDB	2.04% 1		83.67% 41	14.29% 7	100.00% 49
			PTB		5.00% 1	95.00% 19		100.00% 20
			Total	1.83% 2	0.92% 1	81.65% 89	15.60% 17	100.00% 109
		Total		11.92% 46	1.81% 7	70.21% 271	16.06% 62	100.00% 386

Table 9: Party coalition and clustering details for the year of 2014, when using version v2 of the voting network.

Coalition Loyalty 2011-2014

			Cluster ID				
Year	Version	Party Alliance	C1	C2	C3	C4	Total
2011	v1	Government	199 64.19%	4 1.29%	103 33.23%	4 1.29%	310 100.00%
		Opposition	22 16.42%	48 35.82%	28 20.90%	36 26.87%	134 100.00%
	v2	Government	206 66.45%	8 2.58%	44 14.19%	52 16.77%	310 100.00%
		Opposition	23 17.16%	86 64.18%	12 8.96%	13 9.70%	134 100.00%
2012	v1	Government	238 76.77%	68 21.94%	2 0.65%	2 0.65%	310 100.00%
		Opposition	27 22.50%	89 74.17%	4 3.33%		120 100.00%
	v2	Government	237 76.45%	2 0.65%	36 11.61%	35 11.29%	310 100.00%
		Opposition	24 20.00%	4 3.33%	50 41.67%	42 35.00%	120 100.00%
2013	v1	Government	132 44.59%	2 0.68%	158 53.38%	4 1.35%	296 100.00%
		Opposition	64 57.66%		46 41.44%	1 0.90%	111 100.00%
	v2	Government	239 82.13%	6 2.06%	28 9.62%	18 6.19%	291 100.00%
		Opposition	70 65.42%	1 0.93%	22 20.56%	14 13.08%	107 100.00%
2014	v1	Government	196 70.76%	6 2.17%	41 14.80%	34 12.27%	277 100.00%
		Opposition	98 89.91%	1 0.92%	2 1.83%	8 7.34%	109 100.00%
	v2	Government	44 15.88%	6 2.17%	182 65.70%	45 16.25%	277 100.00%
		Opposition	2 1.83%	1 0.92%	89 81.65%	17 15.60%	109 100.00%

Table 10: Party coalition during the 2010 presidential elections, for the 2011-2014 term. For each year and network version (v1, v2), the table shows the amount of deputies in each party alliance (government, oposition) classified in each cluster.

Coalition Loyalty 2015-2016

			Cluster ID					
Year	Version	Party Alliance	C1	C2	C3	C4	C5	Total
2015	v1	Government	229 69.60%	3 0.91%	82 24.92%	15 4.56%		329 100.00%
		Opposition	49 41.88%		3 2.56%	65 55.56%		117 100.00%
	v2	Government	210 63.83%	3 0.91%	95 28.88%	21 6.38%		329 100.00%
		Opposition	28 23.93%		7 5.98%	82 70.09%		117 100.00%
2016	v1	Government	201 63.41%	89 28.08%	23 7.26%	4 1.26%		317 100.00%
		Opposition	104 86.67%	7 5.83%	9 7.50%			120 100.00%
	v2	Government	23 7.26%	92 29.02%	186 58.68%	14 4.42%	2 0.63%	317 100.00%
		Opposition	11 9.17%	8 6.67%	88 73.33%	11 9.17%	2 1.67%	120 100.00%

Table 11: Party coalition during the 2014 presidential elections, for the 2015-2018 term. For each year and network version (v1, v2), the table shows the amount of deputies in each party alliance (government, oposition) classified in each cluster.

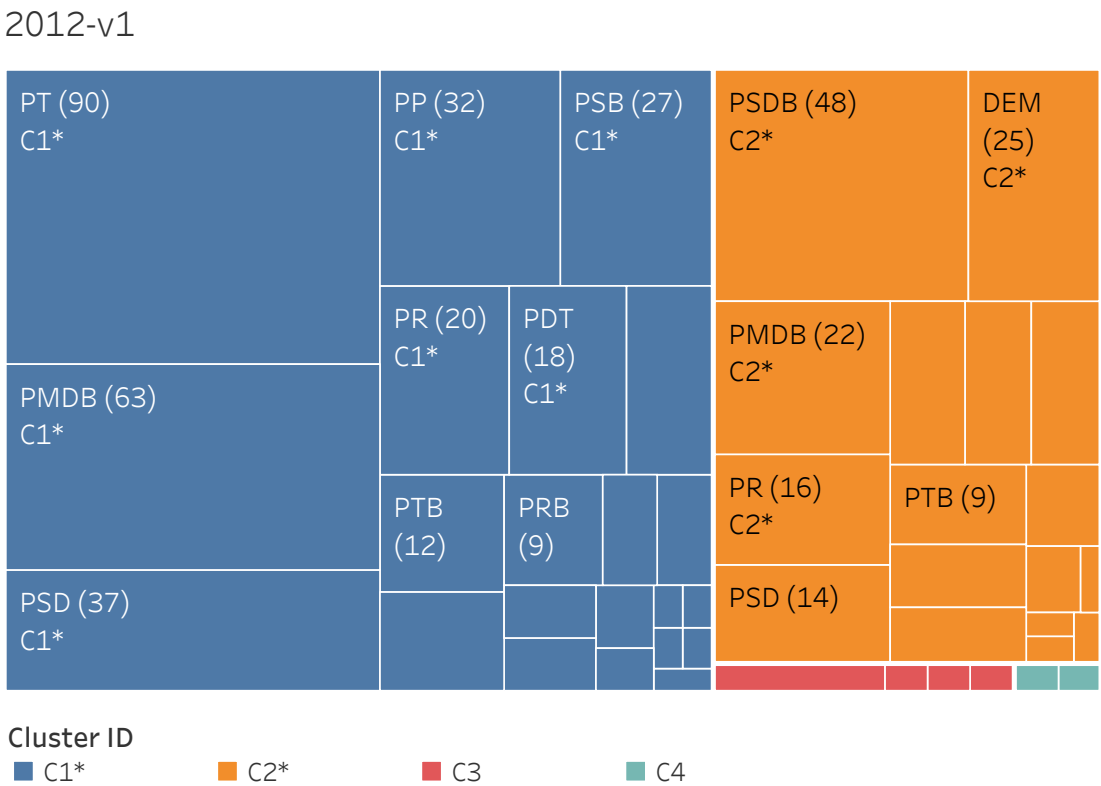

Figure 2: SRCC clustering results for the year of 2012, when using version v1 of the voting network. The treemap shows the parties inside each cluster, the number of deputies of each party in a specific cluster (in parathesis) and also the cluster identifier/label (begins with letter C). Cluster labels marked with an asterisk (*) consist of mediation groups.

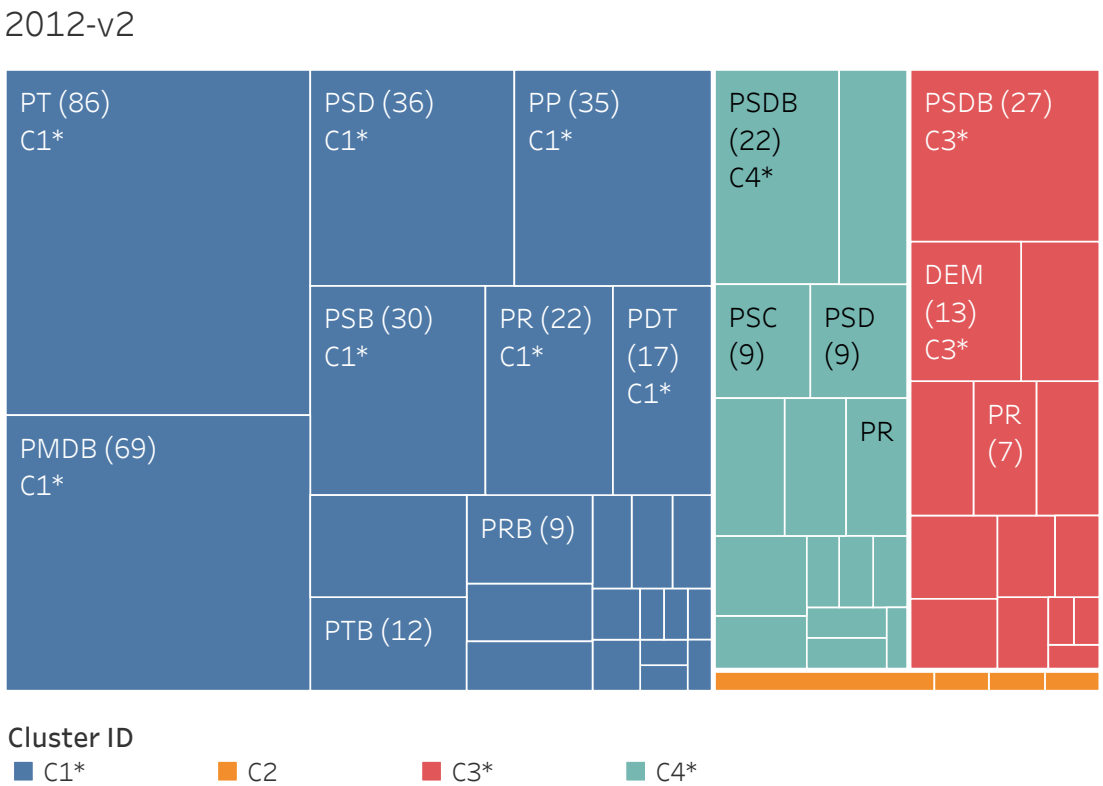

Figure 3: SRCC clustering results for the year of 2012, when using version v2 of the voting network. The treemap shows the parties inside each cluster, the number of deputies of each party in a specific cluster (in parathesis) and also the cluster identifier/label (begins with letter C). Cluster labels marked with an asterisk (*) consist of mediation groups.

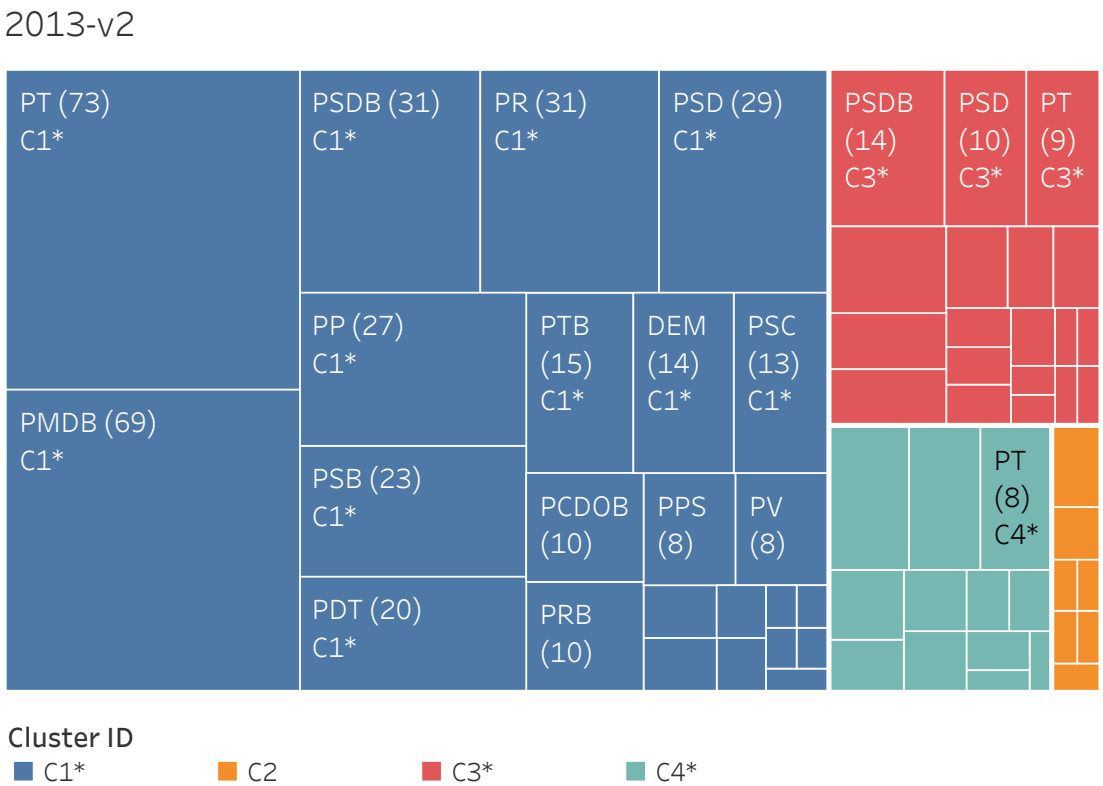

Figure 4: SRCC clustering results for the year of 2013, when using version v2 of the voting network. The treemap shows the parties inside each cluster, the number of deputies of each party in a specific cluster (in parathesis) and also the cluster identifier/label (begins with letter C). Cluster labels marked with an asterisk (*) consist of mediation groups.

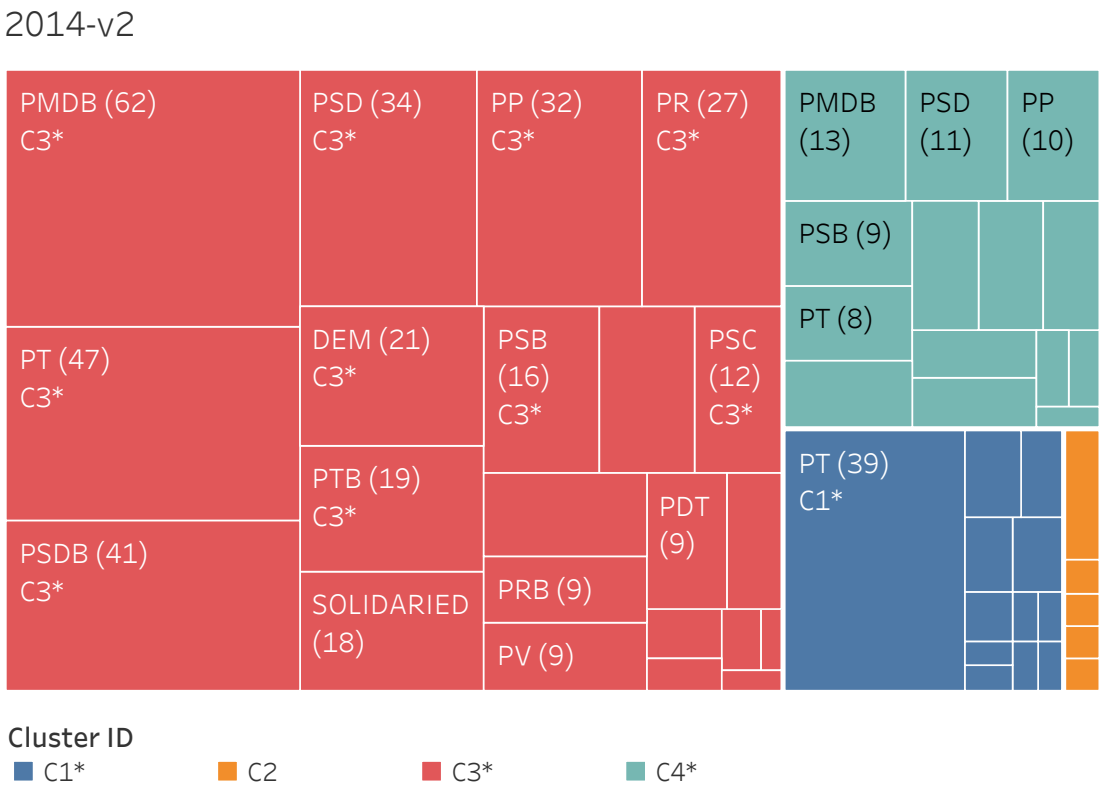

Figure 5: SRCC clustering results for the year of 2014, when using version v2 of the voting network. The treemap shows the parties inside each cluster, the number of deputies of each party in a specific cluster (in parathesis) and also the cluster identifier/label (begins with letter C). Cluster labels marked with an asterisk (*) consist of mediation groups.

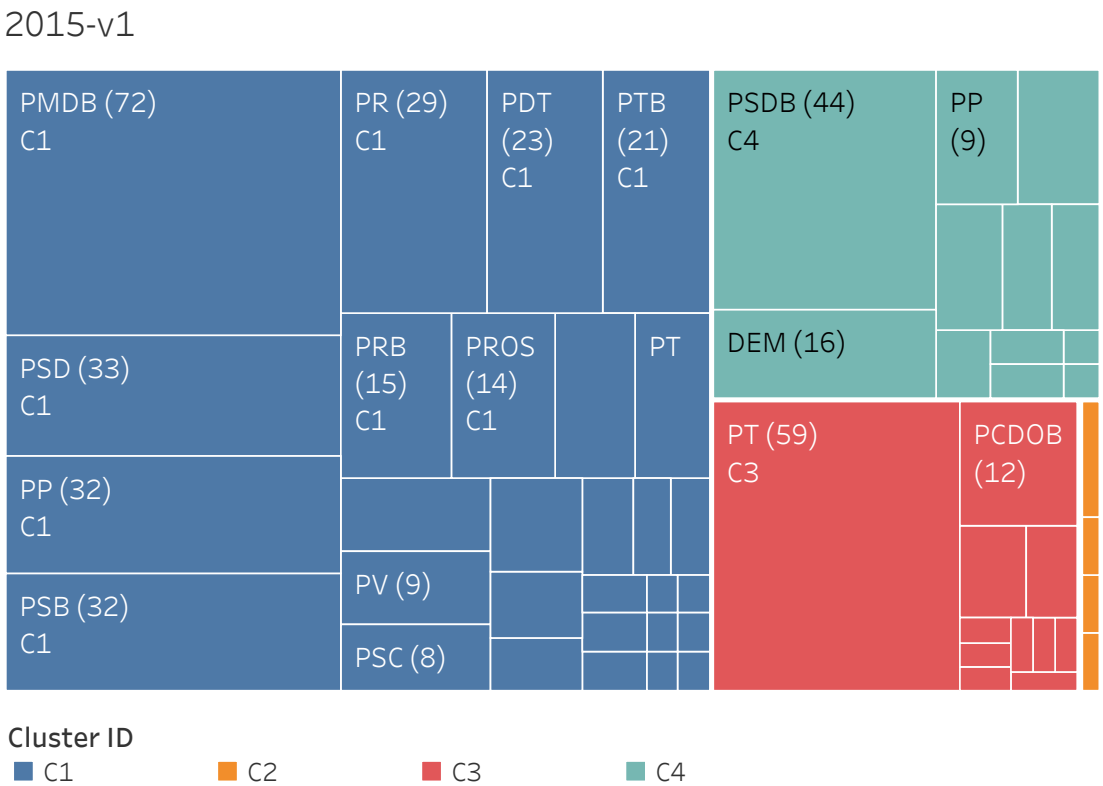

Figure 6: SRCC clustering results for the year of 2015, when using version v1 of the voting network. The treemap shows the parties inside each cluster, the number of deputies of each party in a specific cluster (in parathesis) and also the cluster identifier/label (begins with letter C). Cluster labels marked with an asterisk (*) consist of mediation groups.

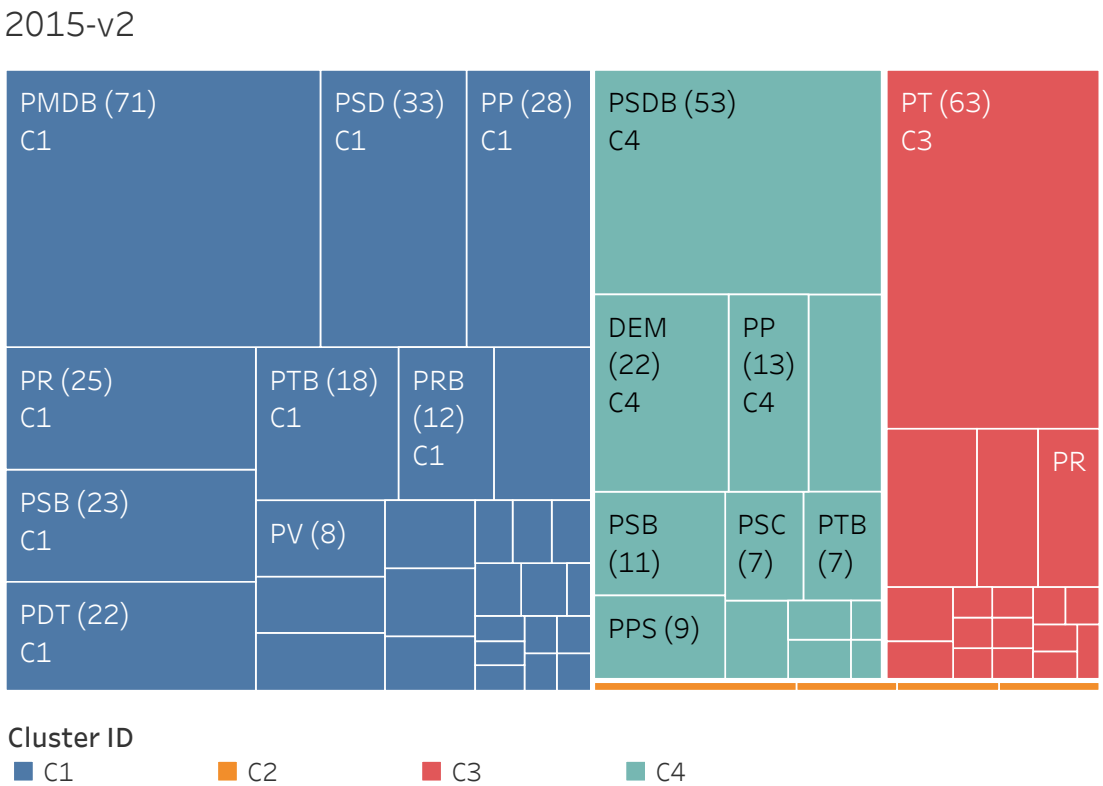

Figure 7: SRCC clustering results for the year of 2015, when using version v2 of the voting network. The treemap shows the parties inside each cluster, the number of deputies of each party in a specific cluster (in parathesis) and also the cluster identifier/label (begins with letter C). Cluster labels marked with an asterisk (*) consist of mediation groups.

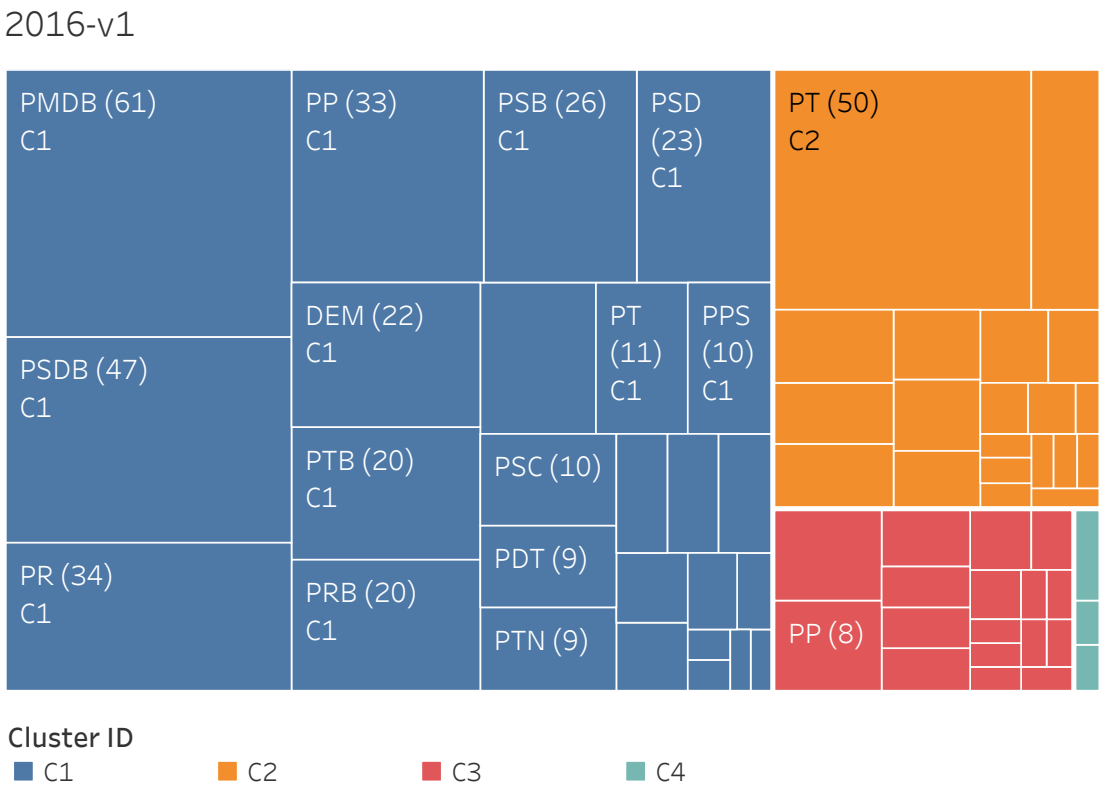

Figure 8: SRCC clustering results for the year of 2016, when using version v1 of the voting network. The treemap shows the parties inside each cluster, the number of deputies of each party in a specific cluster (in parathesis) and also the cluster identifier/label (begins with letter C). Cluster labels marked with an asterisk (*) consist of mediation groups.

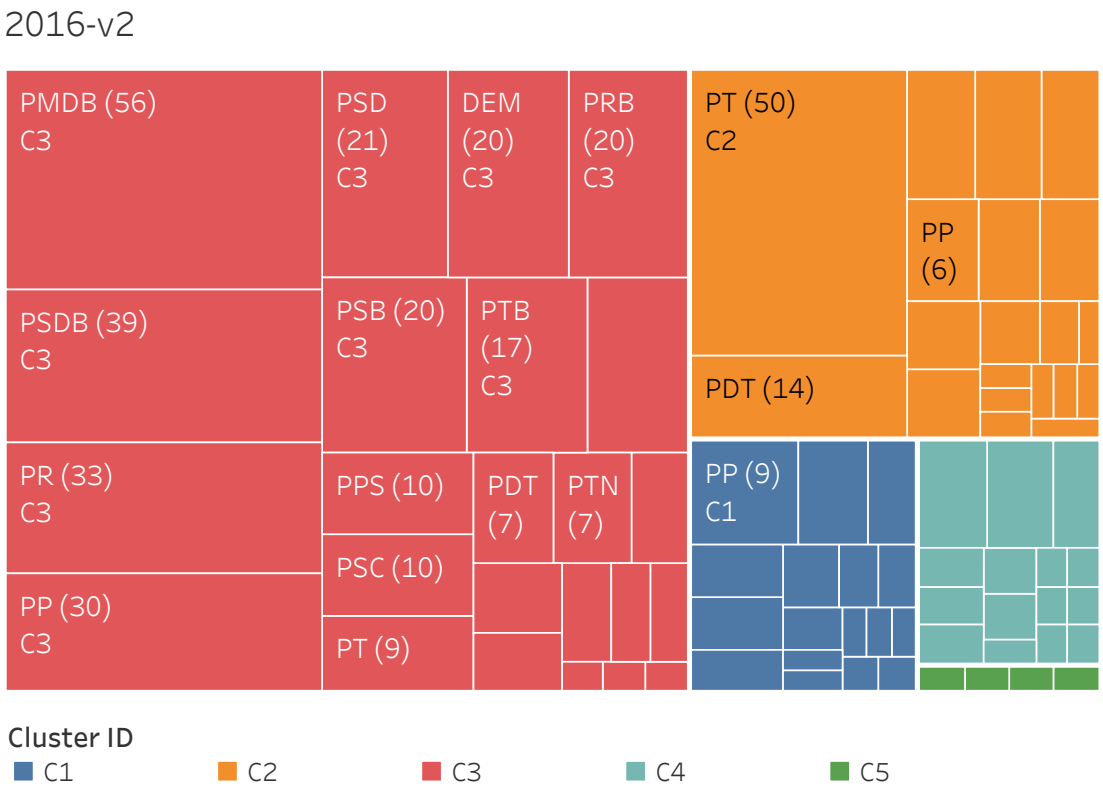

Figure 9: SRCC clustering results for the year of 2016, when using version v2 of the voting network. The treemap shows the parties inside each cluster, the number of deputies of each party in a specific cluster (in parathesis) and also the cluster identifier/label (begins with letter C). Cluster labels marked with an asterisk (*) consist of mediation groups.

Party Leadership (2011-2014)

	Year / Version							
	2011		2012		2013		2014	
Party	v1	v2	v1	v2	v1	v2	v1	v2
DEM	34%	68%	83%	40%	69%	32%	89%	75%
PCDOB	73%	80%	100%	100%	23%	23%	27%	27%
PDT	55%	28%	64%	25%	44%	80%	21%	47%
PMDB	63%	67%	73%	80%	42%	82%	81%	78%
PMN					67%	67%	100%	33%
PP	67%	67%	82%	90%	31%	75%	91%	74%
PPS	33%	58%	64%	18%	55%	73%	89%	22%
PR	54%	35%	44%	19%	62%	84%	88%	82%
PRB	33%	42%	90%	90%	90%	100%	10%	90%
PROS					100%	100%	67%	62%
PRP	50%	50%			50%	50%	100%	100%
PSB			27%	14%	70%	85%	88%	62%
PSC	44%	17%	47%	47%	53%	11%	73%	80%
PSD	35%	17%	27%	18%	35%	59%	79%	71%
PSDB	34%	91%	86%	48%	65%	6%	92%	84%
PSOL	33%	67%	100%	100%	67%	67%	100%	100%
PT	91%	91%	97%	92%	62%	81%	51%	50%
PTB	73%	73%	43%	29%	68%	88%	85%	95%
PTDOB	75%	75%	100%	67%	67%	100%	100%	100%
PV	75%	33%	50%	40%	50%	20%	89%	100%

Table 12: Percentage of deputies who vote after their party leaders (i.e. deputies classified in the same group of their party leader), for each year between 2011 and 2014. On certain periods, the numbers associated with a party may not have been shown. Either because the party still did not exist at that time or did not have any representation in parliament at all.

Party Leadership (2015-2016)

	Year / Version			
	2015		2016	
Party	v1	v2	v1	v2
DEM	30%	96%	85%	77%
PCDOB	92%	92%	50%	50%
PDT	100%	96%	59%	64%
PEN	100%	100%	75%	75%
PHS	100%	100%	100%	100%
PMB			14%	14%
PMDB	96%	95%	11%	77%
PP	78%	68%	69%	4%
PPS	58%	75%	100%	100%
PR	83%	71%	83%	80%
PRB	75%	40%	100%	100%
PROS	100%	86%	71%	71%
PRP			100%	100%
PSB	94%	68%	79%	61%
PSC	46%	43%	71%	71%
PSD	89%	89%	70%	64%
PSDB	80%	96%	87%	2%
PSL			25%	25%
PSOL	100%	100%	100%	100%
PT	83%	89%	82%	82%
PTB	81%	69%	95%	81%
PTDOB	100%	50%	50%	50%
PTN	75%	25%	82%	18%
PV	90%	20%	67%	67%
REDE	100%	100%	80%	80%

Table 13: Percentage of deputies who vote after their party leaders (i.e. deputies classified in the same group of their party leader), for the period between January 2015 and June 2016. On certain periods, the numbers associated with a party may not have been shown. Either because the party still did not exist at that time or did not have any representation in parliament at all.

Year	Network version	Mediation clusters
2011	v1	C3
	v2	C3, C4
2012	v1	C1, C2
	v2	C1, C3, C4
2013	v1	C1, C3
	v2	C1, C3, C4
2014	v1	C1, C2, C3, C4
	v2	C1, C3, C4

Table 14: List of mediation clusters, separated by year and voting network version.