

Uso de Programação Linear Inteira para geração e análise de agrupamentos de políticos da Câmara dos Deputados

Natália Lídia L. Coelho¹, Tatiane Fernandes Figueiredo¹, Rosa Figueiredo²

¹Universidade Federal do Ceará (UFC)
Russas – CE – Brasil

²Laboratoire Informatique d'Avignon – Université d'Avignon
Avignon – France

natalialidia@alu.ufc.br, tatianefernandes@ufc.br
rosa.figueiredo@univ-avignon.fr

Abstract. *Understanding the behavior of politicians and their votes has been the focus of several studies in the literature. Recent work, using heuristic techniques for clustering, presented some analyzes on groups of politicians, brazilians and europeans, in voting for proposals that occurred in a certain period of time. In this context, this work proposes the use of Integer Linear Programming techniques for clustering politicians from the Chamber of Deputies. The results obtained allowed an analysis of the factors that possibly influenced, or not, the votes of each group of politicians created.*

Resumo. *Compreender o comportamento dos políticos e seus votos tem sido foco de diversos estudos na literatura. Trabalhos recentes, utilizando técnicas heurísticas para clusterização, apresentaram algumas análises sobre grupos de políticos, brasileiros e europeus, em votações de propostas que ocorreram em um determinado período de tempo. Nesse contexto, este trabalho propõe o uso de técnicas de Programação Linear Inteira para agrupamentos de políticos da Câmara dos Deputados. Os resultados obtidos permitiram uma análise dos fatores que possivelmente influenciaram, ou não, os votos de cada grupo de políticos criado.*

1. Introdução

Estudos científicos sobre interações e comportamentos em parlamentos políticos tem sido foco de pesquisadores das áreas de humanas e exatas. Esses pesquisadores utilizam algoritmos que buscam solucionar problemas de agrupamentos que medem o grau de "equilíbrio" em uma rede social. De acordo com a *Teoria do Balanço Estrutural*, uma rede social é considerada equilibrada com base na consistência das relações de hostilidade e amizade. Essa teoria afirma que as interações entre os indivíduos de uma rede tendem a formar grupos sociais estáveis [Heider 1946].

Como exemplo, pode-se citar o trabalho de Levorato e Frota [Levorato and Frota 2017], no qual eles geraram várias redes sociais a partir da extração e análise de votações realizadas na Câmara dos Deputados (CD). Os autores utilizaram métricas de similaridade/dissimilaridade de votos entre pares de deputados para determinar o tipo de relacionamento existente entre eles, que pode ser positivo ou negativo. A partir das redes sociais obtidas, com objetivo de compreender a dinâmica

da CD, os autores apresentaram agrupamentos de parlamentares através da aplicação direta de um algoritmo utilizado para resolução do *Correlation Clustering*, um problema que consiste em particionar uma rede social, representada por um grafo de sinal, buscando maximizar as relações positivas dentro de cada subgrafo gerado, assim como as relações negativas entre subgrafos distintos [Bansal et al. 2002]. Os resultados obtidos por Levorato e Frota [Levorato and Frota 2017] serviram como base para realizar uma avaliação de aspectos grupais como lealdade e liderança de partidos, coligações, crises e polarização, sendo identificando aspectos como casos de infidelidade de partidos às suas coligações e a polarização existente entre dois grandes grupos políticos.

De forma semelhante, Arinik et al. [Arinik et al. 2017] também realizaram estudos sobre agrupamentos de políticos, porém com o foco no Parlamento Europeu, mais precisamente considerando votos da 7ª legislatura relacionados aos temas: *Agricultura e Desenvolvimento Rural e Economia e Assuntos Monetários*. Os autores também consideraram o problema estudado como uma aplicação direta do *Correlation Clustering*. A análise qualitativa dos agrupamentos formados revelou a identificação de diversos padrões de grupos. Alguns desses grupos seguiam divisões ideológicas tradicionais, enquanto outros mostraram comportamentos mais específicos. Por exemplo, foi observado que alguns políticos foram alocados de forma excepcional em grupos onde a maioria dos membros possuía ideologias opostas às deles. Apesar dos *insights* relevantes dos autores mencionados sobre os parlamentos estudados, é importante notar que os agrupamentos não levaram em conta características pessoais dos políticos envolvidos. Desta forma, surge a questão de se informações como residência, idade ou formação acadêmica podem auxiliar na compreensão de votos e na auto-organização dos políticos em um parlamento.

Buscando contribuir com essas lacunas científicas ainda existentes na literatura, este trabalho propõe o uso de um modelo de Programação Linear Inteira (PLI) como ferramenta para geração de agrupamentos considerando não só as redes sociais criadas por Levorato e Frota [Levorato and Frota 2017], mas também características pessoais de cada político participante da CD. O modelo de PLI aplicado, de forma semelhante aos algoritmos para resolução do *Correlation Clustering* apresentados na literatura, também faz uso de conceito de similaridade/dissimilaridade, onde parlamentares que possuem relações consideradas positivas são agrupados em um mesmo conjunto, enquanto parlamentares que possuem relações consideradas negativas são agrupados em conjuntos distintos, sendo seu grande diferencial a possibilidade de requisitar que cada grupo a ser gerado possua uma quantidade específica de parlamentares com características pessoais específicas. Neste trabalho, utilizou-se como informação adicional o partido político e estado de residência de cada parlamentar, e as redes sociais foram geradas a partir de dados de votações em temas gerais ocorridos em 2010 na Câmara dos Deputados.

Este trabalho está organizado da seguinte forma: a Seção 2 define formalmente o problema. A Seção 3 apresenta a metodologia adotada. A Seção 4 apresenta os resultados obtidos com as abordagens propostas. Por fim, a Seção 5 contém a conclusão do trabalho realizado e sugestões para trabalhos futuros.

2. O problema de agrupamento na Câmara dos Deputados

Devido a natureza aplicada do estudo realizado, o agrupamento de parlamentares considerando suas características pessoais pode ser visto como um Problema de Formação de

Equipes Competitivas, do inglês *Competitive Team Formation Problem* (CTFP), conforme definido por Figueiredo [Figueiredo 2021].

Utilizando conceitos da *Teoria do Balanço Estrutural* [Heider 1946], onde uma rede social pode ser representada por um grafo de sinal, tal que, os vértices deste grafo representem indivíduos e as arestas seus relacionamentos, positivo (+) ou negativo (-), indicando uma relação de afinidade ou competitividade entre dois indivíduos, respectivamente, o CTFP tem como objetivo encontrar múltiplas equipes (subgrafos), que possuam um determinado conjunto de indivíduos com características pessoais específicas (*skills*), ao mesmo tempo que busca maximizar a afinidade dos indivíduos participantes de uma mesma equipe (relações positivas) e a competitividade (relações negativas) entre participantes de equipes distintas. Mais formalmente, dado uma tupla $\vartheta = (G, S, s, T, t)$, onde:

- O grafo de sinais $G = (V, E, sign)$, em que o conjunto V de vértices representa os indivíduos na rede social, o conjunto E de arestas os relacionamentos entre estes indivíduos e a função $sign : E \rightarrow \{+, -\}$, que vincula a cada aresta um sinal positivo ou negativo;
- O conjunto de características pessoais (*skills*) S ;
- A função $s : V \rightarrow 2^S$ que retorna o conjunto $s(u) \subseteq S$ de *skills* de cada $u \in V$;
- O conjunto de equipes T ;
- A função de demanda $t : T \times S \rightarrow \mathbb{N}$ que retorna o número $t(j, s)$ de indivíduos necessários para cada *skill* $s \in S$ na equipe $j \in T$.

O CTFP busca determinar $|T|$ subgrafos induzidos disjuntos, onde:

1. Cada $u \in V$ pertença à no máximo um subgrafo, com uma *skill* $s \in s(u)$ única;
2. Cada subgrafo j tem no mínimo $t(j, s)$ vértices com a *skill* s , $\forall j \in T, \forall s \in S$;
3. O somatório de vértices alocados aos subgrafos seja maximizado ou a verificação de que as equipes não podem ser formadas.

Neste trabalho, cada equipe representa um grupo de políticos com opiniões semelhantes, enquanto a competitividade reflete as opiniões divergentes entre equipes distintas. Para tal, o grafo de sinal estudado foi gerado a partir de dados de votações na CD em temas gerais no ano de 2010, considerando o partido político e o estado de residência como *skills* de cada parlamentar.

3. Metodologia

Com o objetivo de analisar a formação de agrupamentos dentro da Câmara dos Deputados, o seguinte procedimento metodológico composto de 4 etapas foi seguido:

1. Coleta de dados da Câmara dos Deputados.
2. Geração de instâncias padronizadas conforme a definição do CTFP.
3. Implementação do CTFP para resolver as instâncias (i.e. gerar os agrupamentos).
4. Análise dos agrupamentos obtidos.

3.1. Coleta de dados

Os dados dos deputados da CD foram obtidos do trabalho de Levorato e Frota [Levorato and Frota 2017], contendo informações como nome, código de cadastro na Câmara dos Deputados, partido e estado de residência de cada deputado em 2010.

Também foram obtidos os grafos de sinal desenvolvidos pelos mesmos autores, em que cada vértice corresponde a um deputado e as arestas representam a concordância média entre os pares de deputados nas propostas votadas neste ano. A abordagem para rotulagem das arestas consistiu na comparação dos votos de cada par de deputados em uma proposta. O grau de concordância foi atribuído um valor de -1 quando um votou "a favor" e o outro "contra", e +1 quando ambos compartilharam o mesmo voto. O voto de "obstrução" foi considerado semelhante ao voto "contra", e para o voto de "abstenção", foram adotadas duas abordagens, sendo a considerada neste trabalho aquela que atribui "meia concordância" quando combinado com outros votos, ou seja, +0.5 quando combinado com votos "a favor", "contra" ou outra "abstenção". Para obter a concordância média entre cada par de deputados em um determinado ano, os autores realizaram o cálculo da média dos valores de concordância entre eles dentro do período.

3.2. Geração e padronização das instâncias

Como definido no início da Seção 2, uma instância do problema CTFP é constituída por um grafo de sinal, um conjunto de *skills* e uma demanda de parlamentares com determinadas *skills* para cada equipe. Como mencionado na Subseção 3.1, os grafos de sinal e o conjunto de *skills* foram obtidos através do trabalho de Levorato e Frota [Levorato and Frota 2017], sendo considerado o estado de residência de cada deputado como sua *skill*. Desta forma, foi criado para o ano de 2010, uma matriz $|V| \times |S|$, onde cada elemento é um valor pertencente a $\{1, 0\}$, ou seja, o valor 1 indica que um deputado $u \in V$ possui uma determinada *skill* $s \in S$, caso contrário o elemento recebe valor 0.

O número de grupos a ser gerado foi definido com base nas características específicas do grafo de sinal deste ano, sendo definida a criação de 2 grupos para a instância. A demanda de habilidades foi gerada levando em conta informações específicas dos políticos estudados. Por exemplo, há maior demanda de políticos de estados como São Paulo e Minas Gerais, devido à sua representação mais significativa no parlamento. A Tabela 1 mostra a demanda mínima de cada estado para os 2 grupos requisitados em 2010, representando cerca de $\frac{1}{5}$ dos deputados de cada estado em cada grupo.

Tabela 1. Total de deputados e a demanda de cada estado do ano de 2010.

	Total de deputados	Demanda por equipe		Total de deputados	Demanda por equipe		Total de deputados	Demanda por equipe		Total de deputados	Demanda por equipe		Total de deputados	Demanda por equipe	
AC	8	2	CE	23	5	MG	53	11	PE	25	5	RO	10	2	
AL	9	2	DF	9	2	MS	8	2	PI	11	2	RR	9	2	
AM	8	2	ES	10	2	MT	10	2	PR	32	6	RS	32	6	
AP	9	2	GO	17	3	PA	18	4	RJ	51	10	SC	18	4	
BA	44	9	MA	20	4	PB	12	2	RN	8	2	SE	9	2	
													SP	72	14
													TO	10	2

3.3. Implementação do modelo PLI

Como mencionado anteriormente, o modelo de PLI utilizado neste trabalho foi definido por Figueiredo [Figueiredo 2021], sendo apresentado a seguir:

$$x_{us}^j = \begin{cases} 1, & \text{se o indivíduo } u \text{ está na equipe } j \text{ com a habilidade } s, \forall u \in V, \forall j \in T, \forall s \in s(u). \\ 0, & \text{caso contrário.} \end{cases}$$

$$(CTFP) \max \sum_{j \in T} \sum_{u \in V} \sum_{s \in s(u)} x_{us}^j \quad (1)$$

$$\text{s.a: } \sum_{s \in s(u)} x_{us}^j + \sum_{s \in s(v)} x_{vs}^j \leq 1, \quad \forall uv \in E^-, \forall j \in T \quad (2)$$

políticos, representando 34% do total nesta equipe. Já na equipe 1, apenas 14 parlamentares são desses partidos, ou seja, 8% do total.

Apesar da influência dos partidos mencionados, ainda há uma divisão significativa de opiniões entre políticos do mesmo partido em 2010. Por exemplo, nos partidos PCDOB, PHS, PRB e PSOL, há uma divisão aproximada de 50% dos políticos em cada equipe. Outros partidos também apresentam uma distribuição, embora não tão equitativa, dos seus políticos em equipes distintas. É importante destacar o partido PMDB, que tinha um número considerável de políticos ativos na Câmara em 2010, com 41% alocados na equipe 1 e 59% na equipe 2, evidenciando uma clara divisão de opiniões.

5. Conclusões e trabalhos futuros

Este trabalho apresentou os resultados preliminares de um projeto de pesquisa cujo objetivo principal é gerar agrupamentos de parlamentares da Câmara dos Deputados. Os agrupamentos foram feitos levando em conta características pessoais e a similaridade/dissimilaridade de votos em temas gerais em 2010. A priori, considerou-se o estado de residência e o partido político de cada parlamentar como características pessoais.

A partir da análise dos resultados, concluiu-se que no ano de 2010 o estado de residência dos parlamentares não teve influência significativa em suas decisões de voto. Para a maioria dos estados, foram identificados subgrupos de parlamentares com votos divergentes. Além disso, constatou-se que alguns partidos políticos exerceram possíveis influências nos votos dos parlamentares, com a predominância de parlamentares de um mesmo partido em apenas um dos grupos, indicando opiniões semelhantes entre os membros desse grupo.

Após análises iniciais, observou-se que certos padrões de votos não estão associados nem ao estado de residência nem ao partido político do parlamentar. Em futuros trabalhos, pretende-se explorar outras características pessoais, como faixa etária e formação acadêmica, para identificar novos padrões. Além disso, planeja-se ampliar a análise para outros anos e parlamentos, como o Parlamento Europeu.

Referências

- Arinik, N., Figueiredo, R., and Labatut, V. (2017). Signed graph analysis for the interpretation of voting behavior. In *CEUR Workshop Proceedings*, volume 2025.
- Bansal, N., Blum, A., and Chawla, S. (2002). Correlation clustering. In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pages 238–247.
- Figueiredo, T. F. (2021). *Team formation problems: an integer linear optimization approach*. PhD thesis, Universidade Federal do Ceará.
- Heider, F. (1946). Attitudes and cognitive organization. *The Journal of psychology*, 21(1):107–112.
- Levorato, M. and Frota, Y. (2017). Brazilian Congress structural balance analysis. *Journal of Interdisciplinary Methodologies and Issues in Sciences*, Vol. 2 - Graphes et systèmes sociaux.