Termômetro da Polarização: monitoramento contínuo da polarização parlamentar a partir de redes de votação

Frederico Dolher Souza Baker
Departamento de Ciência da Computação
Universidade Federal de Minas Gerais
Belo Horizonte, Brasil
fredericobaker@dcc.ufmg.br

Orientador: Fabrício Benevenuto
Departamento de Ciência da Computação
Universidade Federal de Minas Gerais
Belo Horizonte, Brasil
fabricio@dcc.ufmg.br

***Resumo*—A polarização política é um tema recorrente no debate público, mas a sua medição objetiva e o seu acompanhamento ao longo do tempo permanecem, em geral, restritos ao ambiente acadêmico. Este trabalho apresenta um sistema para o monitoramento contínuo da polarização na Câmara dos Deputados, construído sobre uma metodologia de análise de redes de votação com sinal. Diariamente, o sistema coleta as votações nominais a partir da API oficial da Câmara, reconstrói grafos com sinal em três granularidades temporais (legislatura, ano e mês) e recalcula um conjunto de métricas. A principal contribuição metodológica é um índice de polarização normalizado, expresso em graus, cuja escala se ancora em um limite teórico de 75°. Esse limite é derivado analiticamente do comportamento das tríades equilibradas em um grafo completo bipartido simétrico e corresponde ao valor máximo que a proporção de tríades que representam antagonismo podem atingir nesse cenário. A normalização torna o índice interpretável mesmo sem conhecimento técnico prévio, uma vez que ancora a leitura em um valor de referência bem definido. Por fim, os resultados são disponibilizados por meio de uma interface pública organizada em camadas de profundidade crescente, acessível tanto a especialistas quanto ao público geral.**

***Index Terms*—polarização política, grafos com sinal, equilíbrio estrutural, índice de polarização, Câmara dos Deputados, análise de redes**

I. Introdução

A forma como os representantes eleitos se alinham e se opõem no interior de um parlamento é um dos indicadores mais diretos da saúde de uma democracia. Há polarização quando os parlamentares se organizam em blocos coesos internamente mas antagônicos entre si, e quando esses blocos se tornam progressivamente menos permeáveis, com menos movimentação entre eles. O fenômeno da polarização não é recente nem exclusivo do Brasil: estudos sobre o Congresso americano documentam, ao longo de décadas, um afastamento ideológico crescente entre os dois principais partidos [10], [11]. Ainda assim, a intensidade e a rapidez com que essa dinâmica se consolidou no Brasil nas duas últimas legislaturas fazem do caso brasileiro particularmente interessante.

A análise da polarização a partir de dados legislativos apresenta uma vantagem em relação a outras abordagens. Boa parte da pesquisa recente recorre a dados de redes sociais como termômetro do clima político [8], [9]. Esse tipo de dado registra aquilo que os atores declaram e compartilham, mas não aquilo que efetivamente decidem. As votações nominais, por outro lado, documentam decisões com consequências legislativas concretas: cada voto é um posicionamento público, com data, hora e resultado registrados no sistema oficial da Câmara. Tal registro possui vieses próprios, uma vez que as votações nominais representam apenas uma fração do processo decisório e a seleção do que é submetido a esse tipo de votação não é aleatória. Ainda assim, ele não está sujeito a distorções de plataforma nem a comportamentos estratégicos de visibilidade, o que o torna um sinal mais objetivo.

Em trabalho anterior, esse registro foi explorado por meio de análise de redes [1]. A partir de todas as votações nominais disponíveis na API da Câmara desde 2003, foram construídas redes ponderadas com sinal: para cada par de deputados que participou de uma mesma votação, o peso da aresta entre eles é incrementado em $+1$ quando os votos coincidem e em -1 quando divergem. Sobre essa estrutura, foram calculadas três dimensões de análise: a polarização, por meio da teoria do equilíbrio estrutural; a coesão partidária, por meio da detecção de comunidades com o algoritmo de Louvain; e a permeabilidade ideológica, por meio de uma simulação do tipo SIR. Os resultados evidenciaram uma inflexão a partir da 55ª legislatura, quando a lógica de alinhamento na Câmara deixou de se organizar predominantemente por afinidade direta e passou a se organizar por antagonismo compartilhado, com os deputados se agrupando em função de um adversário comum.

Esse tipo de análise, no entanto, costuma permanecer restrito ao ambiente acadêmico. Seus resultados ficam restritos aos artigos, sem qualquer mecanismo de consulta contínua. Não é possível consultar o índice de polarização da semana atual, comparar a estrutura da rede em dois meses distintos ou identificar com quais parlamentares um determinado deputado mais concorda ou discorda. Para quem não é pesquisador, seja um jornalista que cobre o Congresso, um estudante de ciência política ou um cidadão interessado, o dado está teoricamente disponível, já que a API da Câmara é pública, mas o acesso à sua interpretação não está.

O presente trabalho parte precisamente dessa limitação. A proposta consiste em levar uma metodologia de análise de polarização legislativa [1] a um sistema de funcionamento contínuo, que coleta as votações diariamente, reconstrói os grafos, calcula os índices e os apresenta em uma interface navegável. Além de operacionalizar a metodologia, este trabalho propõe o índice de polarização adotando uma escala normalizada, com um teto de referência derivado de um resultado teórico, o que o torna consideravelmente mais comunicável do que a proporção bruta de tríades utilizada anteriormente.[1]

Essa normalização é centrar para atingir o objetivo de criar um sistema acessível ao público geral. Uma proporção de, por exemplo, 65% de tríades de um determinado tipo não oferece, isoladamente, qualquer âncora de interpretação, pois não se sabe se o valor é alto ou baixo, nem qual seria o máximo possível. O índice em graus resolve esse problema ao fixar uma referência: o valor de 100° corresponde ao teto teórico de um grafo completo perfeitamente bipartido, e qualquer medição passa a ser lida nessa escala. Em uma interface destinada ao público geral, no qual o visitante pode não ter formação técnica, essa diferença é relevante.

A interface foi projetada tendo a interpretabilidade do usuário leigo como um dos focos principais. Grafos de redes são familiares a pesquisadores, mas não ao público em geral, razão pela qual a interface foi organizada em camadas de profundidade crescente. O visitante que acessa o sistema pela primeira vez encontra um número e um gráfico de linha histórico, de leitura imediata. Aquele que deseja se aprofundar encontra o grafo interativo e, em um nível seguinte, o perfil individual de cada deputado. Nenhuma camada depende da anterior para fazer sentido.

O restante deste artigo está organizado da seguinte forma. A Seção II apresenta o referencial teórico. A Seção III detalha a fonte de dados e a construção da rede. A Seção IV define o índice de polarização e deriva seu limite teórico. A Seção V descreve a arquitetura do sistema e a Seção VI, o pipeline de processamento. A Seção VII discute as decisões de visualização do grafo, a Seção VIII apresenta a interface e a Seção IX conclui o trabalho.

II. Referencial Teórico

A. Redes de votação parlamentar

A modelagem do comportamento legislativo como rede complexa constitui uma linha de pesquisa já razoavelmente consolidada na literatura. No contexto brasileiro, Brito et al. [5] construíram redes de concordância a partir de votações nominais da Câmara e aplicaram extração de backbone e detecção de comunidades para identificar agrupamentos ideológicos. Levorato e Frota [6] avançaram nessa direção ao representar as relações como grafos com sinal, isto é, incluindo explicitamente a discordância como aresta negativa, e ao empregar técnicas de *correlation clustering* na análise de coalizões. Vaz de Melo [7], por sua vez, propôs um algoritmo para reorganizar os partidos com base em padrões empíricos de votação, demonstrando ser possível reduzir substancialmente o número de siglas sem perda de coerência ideológica.

A opção por redes com sinal, adotada neste trabalho, justifica-se pelo que se desejava medir. Redes baseadas apenas em similaridade capturam quem concorda com quem, mas descartam a informação a respeito de quem discorda. Na análise de polarização, contudo, a discordância é justamente o aspecto central: blocos antagônicos se definem não apenas pela coesão interna, mas pela oposição a um bloco rival. Uma rede que ignora as arestas negativas não é capaz de distinguir um cenário de ampla coalizão de um cenário de polarização acentuada.

B. Teoria do Equilíbrio Estrutural

A teoria do equilíbrio estrutural foi originalmente proposta por Heider, no campo da psicologia social [2], a partir da observação de que determinadas configurações de relacionamento geram tensão cognitiva e tendem à instabilidade. A ideia central é a seguinte: se A se relaciona positivamente com B e B com C, há uma pressão natural para que A também se relacione positivamente com C. Configurações que violam essa lógica, em que A se relaciona bem com B e B com C, mas A se opõe a C, são instáveis e tendem a se resolver com o tempo.

Cartwright e Harary formalizaram a teoria do equilíbrio estrutural em termos de grafos [3]. Em um grafo com sinal, uma tríade (triângulo) é dita equilibrada quando o produto dos sinais de suas três arestas é positivo, o que ocorre em duas situações: três arestas positivas $(+\,+\,+)$ ou exatamente uma positiva e duas negativas $(+\,-\,-)$. As duas configurações restantes, duas positivas e uma negativa $(+\,+\,-)$ e três negativas $(-\,-\,-)$, são desequilibradas. A proposição central do arcabouço é que grafos com sinal tendem ao equilíbrio global: à medida que os indivíduos resolvem suas tensões relacionais, a rede se aproxima de uma configuração na qual todos os triângulos estão equilibrados.

Redes reais raramente atingem o equilíbrio perfeito, porém a proporção de triângulos equilibrados em redes legislativas costuma ser bastante elevada; análises anteriores da Câmara registraram valores superiores a 99% em todas as legislaturas examinadas [1]. O que varia de forma significativa é a distribuição entre os tipos equilibrados, e é precisamente nesse ponto que o sinal de polarização fica evidente.

C. Extração de Backbone

Redes construídas a partir de votações nominais são, por natureza, densas. Qualquer par de deputados que tenha participado de uma mesma votação dá origem a uma aresta e, ao final de uma legislatura com centenas de votações, a rede fica praticamente completa. Grande parte dessas conexões, entretanto, corresponde a alinhamentos pontuais, sem relevância estrutural: dois deputados que coincidiram em algumas votações de rotina não mantêm, por esse motivo, uma relação política significativa.

[1] Sistema disponível em https://termometrodapolarizacao.com.br. Código em https://github.com/FredericoBaker/termometro-da-polarizacao.

O filtro de disparidade proposto por Serrano et al. [4] trata desse problema de maneira adequada ao contexto da rede construída neste trabalho. Em vez de aplicar um limiar global de peso, o que eliminaria conexões localmente importantes em nós de força total baixa, o filtro avalia cada aresta no contexto local do nó. Para um nó i de grau k_i e força total $S_i = \sum_j |w_{ij}|$, a fração de força direcionada ao vizinho j é dada por:

$$p_{ij} = \frac{|w_{ij}|}{S_i} \tag{1}$$

Sob a hipótese nula de que as arestas de i distribuem sua força de modo uniforme, a probabilidade de se observar, por acaso, uma fração maior ou igual a p_{ij} é $(1 - p_{ij})^{k_i - 1}$. A aresta é preservada quando essa probabilidade é inferior a α:

$$p_{ij} > 1 - \alpha^{1/(k_i - 1)} \tag{2}$$

Uma aresta é mantida no backbone quando satisfaz o critério em relação a pelo menos um dos dois nós que conecta. O parâmetro α regula a intensidade da filtragem, sendo que valores baixos retêm apenas as conexões mais relevantes, ao passo que valores altos são mais permissivos.

D. Ferramentas existentes

O Voteview [12] é a principal referência de plataforma pública de análise legislativa. Apoiado no método NOMINATE [10], estima a posição ideológica de cada parlamentar americano em um espaço bidimensional contínuo, a partir do histórico de votações. Trata-se de uma abordagem relevante para comparar posições latentes ao longo do tempo. No entanto, o modelo ideológico é abstrato, e as coordenadas resultantes não possuem interpretação direta em termos das votações que as originaram.

A abordagem adotada neste trabalho é mais direta. Em vez de estimar posições em um espaço latente, mapeiam-se explicitamente as relações de concordância e discordância a partir dos registros de votação, sem pressupor qualquer dimensão ideológica subjacente. O índice de polarização não indica a posição de cada ator em um espectro; descreve, antes, como a estrutura de alianças e antagonismos se organiza na rede como um todo.

III. Dados e Construção da Rede

A. Fonte de dados

O sistema consome a API de Dados Abertos da Câmara dos Deputados [13], que disponibiliza, o histórico completo de votações nominais desde o início dos registros eletrônicos. As votações em plenário podem ser de três tipos: simbólica, nominal e secreta. Na votação simbólica, a mais comum, o presidente da Câmara proclama o resultado com base em manifestação coletiva, sem registro individual dos votos. Na votação secreta, o sigilo é garantido em casos específicos, como eleições internas. Apenas a votação nominal registra o voto de cada deputado individualmente, sendo, portanto, o único tipo que permite reconstruir as relações de concordância e discordância entre pares de parlamentares.

Cada voto nominal pode assumir seis categorias: Sim, Não, Abstenção, Obstrução, Art. 17 (ausência justificada) e Ausente. Para a construção da rede, consideram-se apenas os votos Sim e Não. Abstenção e Obstrução constituem posições neutras ou estratégicas, que não revelam alinhamento com nenhum dos lados; as ausências, justificadas ou não, indicam apenas que o deputado não participou da decisão. A inclusão dessas categorias geraria pares de arestas sem informação real de alinhamento e distorceria os pesos acumulados.

B. Construção das arestas

A rede inicia vazia e é atualizada de forma iterativa a cada votação nominal processada. Para cada par de deputados (i, j) que participou de uma mesma votação, a contribuição ao peso da aresta é $+1$ quando ambos votam no mesmo sentido (Sim-Sim ou Não-Não) e -1 quando votam em sentidos opostos (Sim-Não ou Não-Sim). Após o processamento de múltiplas votações, o peso acumulado reflete o saldo entre concordâncias e discordâncias ao longo do período.

Pesos positivos elevados indicam um alinhamento consistente, ao passo que pesos negativos indicam oposição recorrente. Um peso próximo de zero entre dois deputados que participaram de muitas votações em comum é também informativo: significa que eles concordam e discordam com frequência aproximadamente igual, sem um padrão claro de alinhamento.

IV. O Índice de Polarização

A. Por que usar tríades

Métricas diretas do grafo, como a proporção de arestas negativas ou a modularidade da rede, capturam aspectos da estrutura, mas não distinguem com precisão o tipo de polarização que se pretende medir. Uma rede com muitas arestas negativas pode resultar tanto de um plenário genuinamente dividido quanto, simplesmente, de um grande número de votações decididas por margens estreitas. O índice aqui proposto mede algo mais específico no caso a prevalência do alinhamento por um antagonismo compartilhado, em contraposição ao alinhamento por afinidade direta.

Esse padrão é capturado pela estrutura de tríades. Um triângulo no grafo com sinal sintetiza, de forma compacta, a relação entre três atores: o modo como cada par se alinha e a geometria de alianças e oposições que daí resulta. A classificação proposta pela teoria do equilíbrio estrutural agrupa essas geometrias em dois tipos equilibrados e dois desequilibrados, e é na distribuição entre os tipos equilibrados que reside a informação de polarização.

B. Os quatro tipos de tríade

Os tipos possíveis de tríade em um grafo com sinal, classificados pelo número de arestas positivas, são apresentados na Tabela I. A tríade T_{+++} representa cooperação mútua entre três atores. A tríade T_{+--} codifica a lógica do "inimigo do meu inimigo é meu amigo": dois deputados se alinham porque ambos se opõem a um terceiro, e não necessariamente porque concordem de forma consistente entre si. As tríades T_{++-}

Tabela I
CLASSIFICAÇÃO DAS TRÍADES EM GRAFOS COM SINAL

Tipo	Config.	Equil.?	Interpretação
T_{+++}	$+++$	Sim	Cooperação mútua
T_{++-}	$++-$	Não	Instabilidade estrutural
T_{+--}	$+--$	Sim	Aliança por antagonismo
T_{---}	$---$	Não	Instabilidade estrutural

e T_{---} são desequilibradas e indicam tensão estrutural. O índice considera exclusivamente as tríades equilibradas, T_{+++} e T_{+--}. As tríades desequilibradas correspondem a configurações instáveis que, segundo a teoria, tendem a desaparecer com o tempo. Ao trabalhar apenas com as equilibradas, mede-se o estado atual da estrutura de alianças já consolidadas, sem contaminação por relações possivelmente em transição. A proporção de tríades do tipo T_{+--} entre as equilibradas é definida como:

$$\rho = \frac{T_{+--}}{T_{+++} + T_{+--}} \tag{3}$$

Quando ρ é elevado, a maior parte dos triângulos estáveis da rede corresponde a alianças por antagonismo, e não a cooperação direta. Esse é o sinal estrutural da polarização.

C. O limite teórico de 75%

Para que ρ possa ser interpretado como índice, é necessário conhecer o valor máximo que assumiria em um cenário de polarização extrema. Esse limite pode ser obtido analiticamente a partir do grafo completo bipartido simétrico, que representa a divisão mais nítida possível.

Considere-se um grafo completo com n nós, divididos igualmente em dois blocos de tamanho $m = n/2$. No interior de cada bloco, todas as arestas são positivas; entre os blocos, todas são negativas. As tríades equilibradas que surgem são de dois tipos: quando os três nós pertencem ao mesmo bloco, as três arestas são positivas (T_{+++}); quando dois nós pertencem a um bloco e um ao outro, há uma aresta positiva, entre os dois do mesmo bloco, e duas negativas (T_{+--}). Não há outras configurações possíveis.

Contando cada tipo:

$$T_{+++} = 2\binom{m}{3} = \frac{m(m-1)(m-2)}{3} \tag{4}$$

$$T_{+--} = 2\binom{m}{2} \cdot m = m^2(m-1) \tag{5}$$

Substituindo em (3):

$$\rho = \frac{m^2(m-1)}{\dfrac{m(m-1)(m-2)}{3} + m^2(m-1)} = \frac{3m}{4m-2} \tag{6}$$

Quando $m \to \infty$, $\rho \to 3/4 = 0{,}75$. Conclui-se que, no cenário de polarização mais extremo que um grafo completo é capaz de exibir, a proporção de tríades do tipo $+--$ entre as equilibradas converge para 75%, sem nunca ultrapassar esse valor. Para valores finitos de m, ρ é sempre estritamente menor que 75%.

A razão pela qual o limite não é 100% torna-se evidente a partir da geometria. Mesmo em um grafo bipartido perfeito, há um grande número de tríades T_{+++} formadas por três nós de um mesmo bloco. Tais tríades "diluem" as tríades T_{+--}, e quanto mais densos os blocos internos, maior essa diluição. O teto de 75% reflete precisamente esse equilíbrio no limite.

D. O índice em graus e sua normalização

Estabelecido o limite teórico, define-se o índice de polarização como:

$$P = 100 \cdot \frac{\rho}{0{,}75} \tag{7}$$

A escala atribui 100 graus ao comportamento do grafo completo bipartido simétrico. Um índice de 80° corresponde a $\rho = 60\%$, ou seja, a 80% do teto teórico. Um índice de 40° indica que metade dos triângulos equilibrados corresponde a cooperação mútua e a outra metade, a alianças por antagonismo.

A unidade "graus" foi escolhida em analogia ao termômetro. Da mesma forma que a temperatura mede a intensidade de um fenômeno físico em uma escala de referência bem definida, o índice mede a intensidade da polarização estrutural em uma escala ancorada em um marco teórico.

E. Valores acima de 100°

Em grafos reais esparsos, sobretudo depois da filtragem de backbone, ρ pode ultrapassar 75%. Em redes esparsas, os blocos internos apresentam menos conexões, o que reduz a contagem de tríades T_{+++} em relação às T_{+--}. Podem ocorrer configurações locais em que quase todos os triângulos são do tipo $+--$, o que eleva ρ para além do limite do grafo completo. Valores de $P > 100$ indicam que a estrutura local de triângulos da rede concentra mais antagonismo do que seria possível em um grafo completamente conectado entre os mesmos indivíduos.

V. ARQUITETURA DO SISTEMA

O sistema é composto por quatro camadas: o pipeline de processamento, o banco de dados PostgreSQL, a API REST e o frontend web (Fig. 1).

A. Banco de dados em três camadas

O banco PostgreSQL está organizado em três níveis de abstração, cada um com um propósito distinto. O nível *raw* armazena os dados exatamente como retornados pela API da Câmara, em colunas JSONB que preservam todas as informações originais. A principal vantagem dessa organização é a possibilidade de reprocessar qualquer etapa posterior sem novas chamadas à API externa. Caso a lógica de transformação seja alterada, basta reprocessar a partir do dado bruto.

O nível *normalizado* contém as entidades estruturadas: partidos, deputados com seus respectivos mandatos por legislatura e votações. É nesse nível que os votos são normalizados, com Sim e Não convertidos em 1 e 0 e as demais categorias descartadas. Cada entidade normalizada mantém uma flag `transform_dirty`, que indica se ela foi alterada no dado bruto desde a última normalização.

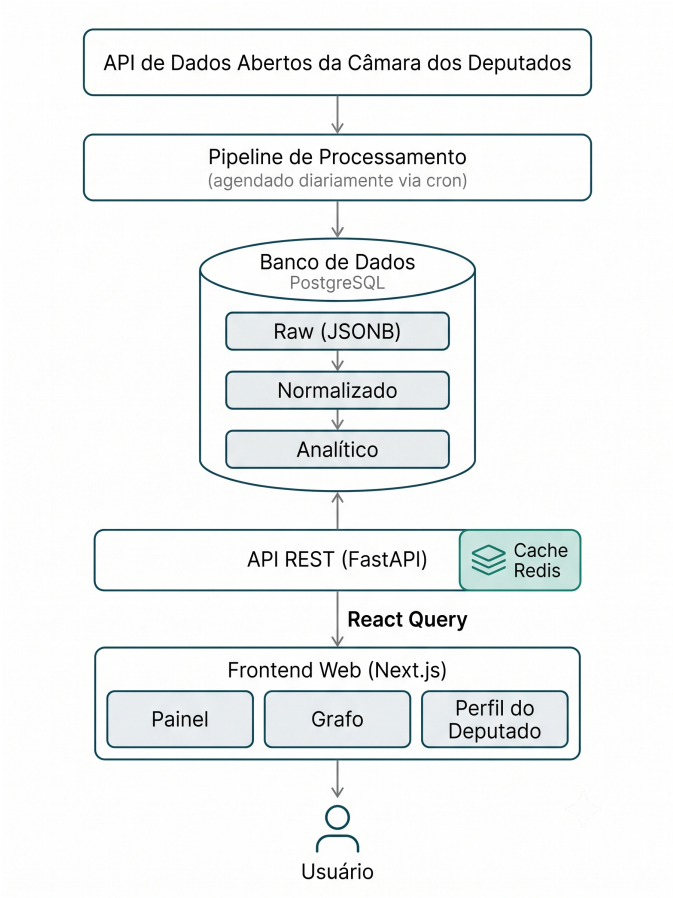


Figura 1. Arquitetura do sistema em quatro camadas. O pipeline atualiza o banco diariamente; a API serve os dados com cache Redis; o frontend consome a API via React Query.

O nível *analítico* armazena os produtos finais do pipeline: os grafos nas três granularidades temporais, as arestas com pesos acumulados, as coordenadas (x, y) dos nós no layout e as métricas de polarização de cada período. A API opera exclusivamente sobre esse nível; os níveis anteriores não são expostos ao frontend.

B. API e cache

A API foi construída com FastAPI e funciona como uma camada de leitura, não mantendo estado próprio. Os endpoints mais custosos, como o grafo completo de uma legislatura, a série temporal histórica e os rankings, são armazenados em cache no Redis. Esse cache é aquecido automaticamente ao final de cada execução bem sucedida do pipeline.

C. Frontend

O frontend, desenvolvido em Next.js, renderiza três telas principais. As páginas de dashboard e de perfil de deputado são renderizadas no servidor, o que garante que o HTML inicial já contenha os dados, favorecendo a indexação e o carregamento rápido. A tela do grafo, que exige interatividade e renderização em canvas, carrega os dados no cliente por meio do React Query, com cache local.

VI. Pipeline de Processamento

A. Estrutura e agendamento

O pipeline é executado diariamente, às 3h UTC, por um processo agendado com cron. Cada execução percorre quatro etapas em sequência: ingestão, transformação, construção do grafo e cálculo de métricas. Uma etapa só tem início quando a anterior é concluída com sucesso e o status é registrado na tabela `ingestion_log` para fins de auditoria.

O reprocessamento incremental é controlado por flags de "sujeira" (*dirty flags*). As votações novas são marcadas como `graph_dirty` e após o processamento na etapa de construção do grafo, os grafos afetados passam a `metrics_dirty`. Dessa forma, o pipeline evita recalcular o que já foi processado, já uma execução que não encontra dados novos não realiza nenhum recálculo. A Fig. 2 ilustra o fluxo completo.

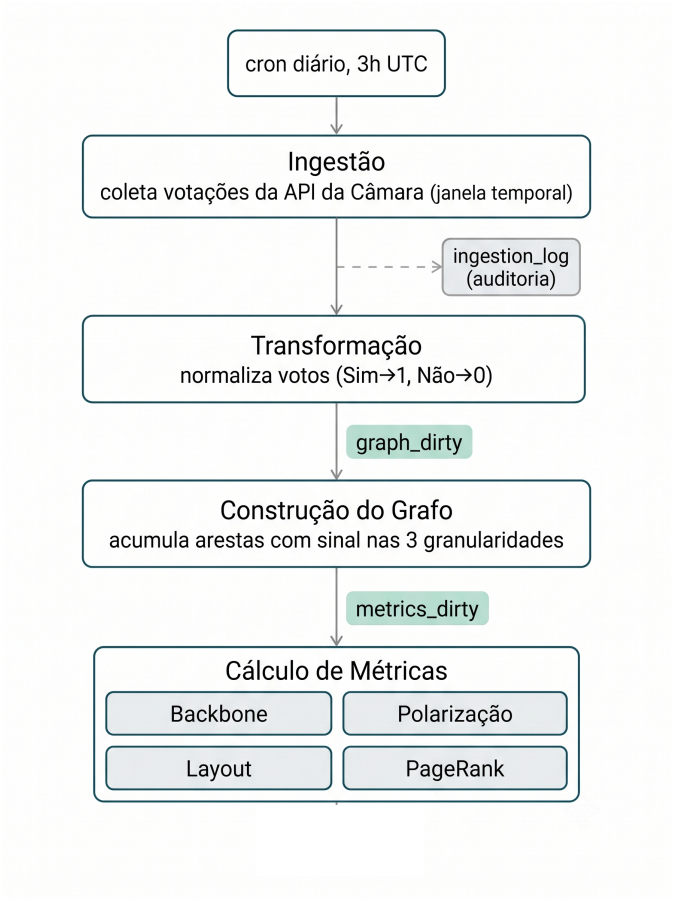


Figura 2. Fluxo do pipeline com as quatro etapas e o controle por dirty flags.

B. Ingestão e janela temporal

A ingestão coleta partidos, deputados, votações e rollcalls da API da Câmara [13]. Cada entidade é armazenada no nível raw do banco, com o payload JSON original intacto. A coleta opera sobre uma janela temporal em que a cada execução, o pipeline determina o intervalo de datas a processar com base no log da execução anterior.

Para evitar a perda de votações em caso de falha ou de indisponibilidade temporária da API, a janela inicia-se sempre três dias antes do término da última execução bem-sucedida. Essa sobreposição é intencional e não gera duplicatas, pois a etapa de construção do grafo verifica, votação a votação, quais grafos já a processaram, ignorando o que já foi feito.

C. Transformação

A transformação converte os dados brutos nas entidades estruturadas. O processo mais relevante nessa etapa é a normalização das votações, em que votos Sim são mapeados para 1, votos Não para 0, e todas as demais categorias (Abstenção, Obstrução, Art. 17, Ausente) são descartadas. A decisão de descartar, em vez de manter, as categorias neutras simplifica a lógica de construção do grafo, que deixa de precisar tratar casos especiais.

A transformação também consolida os mandatos dos deputados. Como um mesmo deputado pode aparecer em múltiplas legislaturas, os registros da tabela `deputies_legislature_terms` vinculam cada deputado a cada legislatura em que exerceu mandato. Esse vínculo é o que permite ao frontend exibir a trajetória completa de legislaturas de cada parlamentar.

D. Construção do grafo

Para cada votação marcada como `graph_dirty = true`, o pipeline identifica a legislatura, o ano e o mês a que ela pertence e atualiza os três grafos correspondentes simultaneamente. Essa atribuição tripla é importante, pois garante que o mesmo dado alimente as três granularidades sem processamento redundante.

A fronteira temporal aqui é importante, dado que as legislaturas da Câmara iniciam-se em 1º de fevereiro e têm duração de quatro anos. Por essa razão, a lógica de atribuição trata janeiro como parte do ano anterior. Uma votação de janeiro de 2024 pertence ao "ano 2023", em consonância com a legislatura iniciada em fevereiro de 2023. O mesmo ajuste aplica-se aos limites de legislatura.

Para cada par de deputados (i, j) presente em uma votação, o pipeline calcula o delta de peso ($+1$ ou -1) e o acumula por meio de um upsert atômica, de forma que caso a aresta já exista, o peso é somado e, caso contrário, ela é criada.

E. Paralelização

A construção do grafo é a etapa mais custosa do ponto de vista computacional. Uma única votação com 513 deputados presentes gera até $\binom{513}{2} \approx 131.000$ pares. Considerando-se centenas de votações em uma legislatura, o volume total de operações atinge a ordem de dezenas de milhões de atualizações de arestas.

O pipeline paraleliza o processamento de votações distintas. Essa paralelização é segura porque as votações são independentes entre si, uma vez que o peso da aresta entre dois deputados é a soma das contribuições de cada votação, e tal soma é comutativa e associativa. O resultado final, portanto, independe da ordem de processamento. A única coordenação necessária ocorre no nível do banco de dados, onde os upserts são atômicos por construção do PostgreSQL.

F. Cálculo de métricas

Sobre cada grafo marcado como `metrics_dirty`, quatro métricas são calculadas em sequência.

Backbone. O filtro de Serrano et al. [4] é aplicado separadamente sobre o subgrafo positivo e o subgrafo negativo, e as arestas sobreviventes de ambos são reunidas para formar o backbone final. O parâmetro α é selecionado automaticamente a partir de uma lista crescente de candidatos (de 0,01 a 0,50). Adota-se o menor valor que mantém pelo menos 80% dos nós no maior componente conexo. Esse critério reduz ao máximo o ruído e, ao mesmo tempo, preserva a estrutura global da rede.

Polarização. Sobre as arestas do backbone, o algoritmo enumera todas as tríades pelo método de interseção de vizinhos e calcula ρ e P, conforme as equações (3) e (7).

Layout. As coordenadas (x, y) de cada nó são calculadas sobre as arestas positivas do backbone pelo algoritmo de Fruchterman-Reingold.

PageRank. A centralidade de cada nó na rede de concordâncias é utilizada para dimensionar os nós no grafo. O cálculo emprega apenas as arestas positivas do backbone, de modo que o PageRank reflita a influência exercida por meio de concordância.

VII. Visualização do Grafo

A. Algoritmo de posicionamento

A posição dos nós é calculada pelo algoritmo de Fruchterman-Reingold [14], um método de layout dirigido por força no qual cada aresta é modelada como uma mola, com força de atração proporcional ao peso, e os nós não conectados se repelem. A implementação utiliza a versão ponderada da biblioteca NetworkX, com 50 iterações.

B. Por que usar apenas arestas positivas

A decisão mais relevante no processo de layout é a de utilizar exclusivamente as arestas positivas do backbone, ignorando por completo as arestas negativas no cálculo de posição. Essa escolha justifica-se em dois níveis.

O primeiro é que o algoritmo de *spring layout* foi concebido para grafos sem sinal, nos quais cada aresta representa atração. Interpretar pesos negativos como repulsão exigiria modificar o funcionamento do algoritmo de maneira significativa, já que seria necessário definir como a força de repulsão se relaciona com a magnitude do peso negativo, como ela se combina com a repulsão geral entre nós não conectados e como o sistema converge nessas condições. Não há consenso na literatura sobre a melhor forma de fazer isso, e qualquer escolha introduziria hiperparâmetros adicionais, sem garantia de um resultado interpretável.

O segundo nível é mais fundamental. A polarização, enquanto estrutura, emerge naturalmente da dinâmica de atração. Deputados que concordam com frequência gravitam uns em direção aos outros, ao passo que aqueles que concordam com

blocos distintos são atraídos para regiões diferentes do espaço. Como resultado, os dois grupos antagônicos se separam no grafo sem que nenhuma força de repulsão explícita seja necessária. O grafo resultante é visualmente coerente, de maneira cada bloco aparece como uma região densa e compacta, e o espaço entre eles reflete a ausência de concordância.

C. Codificação visual

Sobre as posições calculadas pelo layout, três dimensões adicionais são codificadas visualmente nos nós:

- **Cor**: o partido político do deputado. As cores são atribuídas de forma consistente a cada partido em todos os períodos, de modo que o mesmo partido apareça sempre na mesma cor.
- **Tamanho**: o PageRank do deputado na rede de concordâncias. Quanto mais central o deputado, maior o nó, o que torna imediata a identificação dos atores mais influentes de cada bloco.
- **Arestas**: exibidas em verde, para indicar concordância, com espessura proporcional ao peso acumulado. Apenas as arestas do backbone são exibidas; a rede completa seria ilegível.

D. Renderização

A renderização utiliza o Sigma.js 3.0, tendo o Graphology como estrutura de dados subjacente. O Sigma opera sobre canvas WebGL, o que permite manipular redes com centenas de nós e milhares de arestas com interatividade fluida no navegador, sem degradação perceptível de desempenho nas granularidades mais densas. A opção por uma biblioteca dedicada a grafos, em vez de uma biblioteca de visualização genérica, deveu-se à necessidade de suportar operações de pan, zoom, busca por nó e clique com desempenho aceitável em dispositivos comuns.

VIII. A Interface

A. Princípios de design

A fim de criar uma interface amigável para o público geral, foi definida uma arquitetura em camadas de profundidade crescente. A primeira camada consiste em um número e em um gráfico de linha. Qualquer pessoa familiarizada com gráficos de evolução temporal compreende o que está sendo apresentado, sem necessidade de conhecimento prévio em análise de redes ou em equilíbrio estrutural. A segunda camada é o grafo interativo, que requer um pouco mais de familiaridade com a representação, mas é acompanhado de texto explicativo e legenda. A terceira é o perfil individual do deputado, voltada a quem se interessa por um parlamentar específico.

O ponto mais delicado é o grafo. Redes, como forma de visualização, são familiares a quem atua na área, mas não ao público geral. A página inicial aborda essa dificuldade ao explicar, em linguagem simples, o que o grafo representa. O objetivo não é que o visitante necessariamente precisa dominar análise de redes, mas que compreenda que os pontos representam deputados, que pontos próximos concordam com frequência e que dois agrupamentos distintos correspondem a dois blocos políticos. Essa leitura básica deve ser possível sem nenhuma leitura adicional.

B. Painel principal

A tela principal reúne o índice de polarização atual, em graus, com comparação ao mês anterior; uma série temporal histórica em granularidade mensal desde 2003; e rankings com os pares de maior concordância e de maior discordância na legislatura atual (Fig. 3). A granularidade mensal foi escolhida

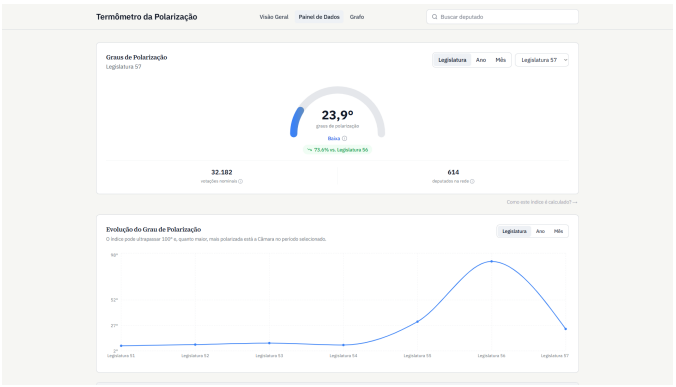


Figura 3. Painel principal com o índice atual em graus, comparação mensal, série temporal histórica e rankings de concordância e discordância.

para a série temporal por equilibrar detalhe e legibilidade. A granularidade anual ocultaria variações internas nas legislaturas que podem ser politicamente relevantes como, por exemplo, um mês marcado por votações em torno de uma emenda constitucional que pode apresentar um pico distinto do restante do ano. As granularidades semanal e diária, por sua vez, seriam excessivamente ruidosas, uma vez que o índice oscila conforme o volume de votações disponíveis em cada período.

C. Exploração do grafo

A tela do grafo (Fig. 4) permite alternar entre as três granularidades: legislatura, ano e mês. O clique sobre um

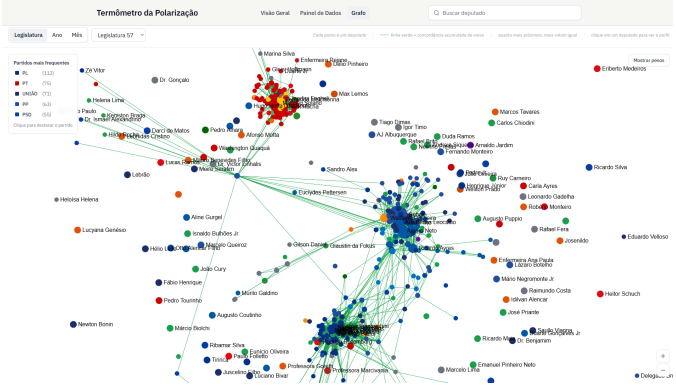


Figura 4. Tela de exploração do grafo com seletor de granularidade temporal e controles de zoom. Clicar num nó navega ao perfil do deputado correspondente.

nó conduz ao perfil do deputado correspondente. A busca

por nome permite localizar um parlamentar específico sem a necessidade de identificá-lo visualmente na rede.

D. Perfil do deputado

Cada deputado possui uma página própria, com foto, nome, partido e estado, além da lista de mandatos registrados na base de dados. O elemento central dessa página é o subgrafo local, que consiste na a vizinhança imediata do deputado, com os nós coloridos por partido. Esse subgrafo evidencia com quais parlamentares o deputado mais se alinha e em que blocos esses aliados se inserem. Abaixo do subgrafo, dois rankings listam os principais aliados e os principais antagonistas do deputado, ordenados pelo peso absoluto acumulado das arestas correspondentes. Os aliados são os deputados com quem ele mais concorda; os antagonistas, aqueles com quem mais discorda. Em conjunto, o subgrafo e os rankings respondem à pergunta mais natural de quem acessa o perfil de um parlamentar: com quem este deputado se alinha?

IX. Conclusão

Este trabalho converteu uma metodologia de análise de polarização legislativa [1] em um sistema funcional e com atualização contínua. O pipeline coleta votações diariamente, reconstrói os grafos, calcula as métricas e disponibiliza os resultados por meio de uma interface acessível a qualquer pessoa.

A contribuição metodológica mais significativa é o índice de polarização em graus. O limite de 75% em um grafo completo bipartido fornece uma âncora para a normalização e torna o índice mais facilmente interpretável sem conhecimento técnico prévio. Essa propriedade é mais importante do que pode parecer no contexto de uma ferramenta pública, já que um número sem referência tende a tornar-se especulativo, ao passo que um número expresso em uma escala com marco bem definido é mais provável de ser interpretado corretamente.

O sistema apresenta limitações que apontam caminhos futuros de extensão. A análise atual agrega todas as votações nominais sem distinção de tema, como um índice de polarização por área de política pública, como a econômica, a ambiental ou a de segurança, revelaria dimensões que o índice geral oculta. A comparação com assembleias legislativas estaduais, ou com câmaras de outros países, seguiria a mesma metodologia, com pouca alteração de arquitetura. Por fim, a coesão partidária e a permeabilidade ideológica, presentes na análise original [1] mas ausentes deste sistema, constituem candidatas naturais a uma versão futura.

Referências

[1] F. Baker e F. Benevenuto, "Quantificando Polarização, Coesão e Permeabilidade na Câmara dos Deputados do Brasil," *Anais do XXXI Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia)*, SBC, 2025, pp. 394–402.

[2] F. Heider, "Attitudes and cognitive organization," *The Journal of Psychology*, vol. 21, n. 1, pp. 107–112, 1946.

[3] D. Cartwright e F. Harary, "Structural balance: a generalization of Heider's theory," *Psychological Review*, vol. 63, n. 5, pp. 277–293, 1956.

[4] M. Á. Serrano, M. Boguñá e A. Vespignani, "Extracting the multiscale backbone of complex weighted networks," *Proceedings of the National Academy of Sciences*, vol. 106, n. 16, pp. 6483–6488, 2009.

[5] A. C. M. Brito, F. N. Silva e D. R. Amancio, "A complex network approach to political analysis: Application to the Brazilian Chamber of Deputies," *PLOS ONE*, vol. 15, n. 3, p. e0229928, 2020.

[6] M. Levorato e Y. Frota, "Brazilian Congress structural balance analysis," *Journal of Interdisciplinary Methodologies and Issues in Science*, vol. 2, 2017.

[7] 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," *PLOS ONE*, vol. 10, n. 10, p. e0140217, 2015.

[8] L. S. de Oliveira, W. Costa, P. O. S. Vaz de Melo e F. Benevenuto, "How politicians communicate in social media: A cross-platform study," *Anais do XXVIII Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia)*, SBC, 2022, pp. 79–87.

[9] G. F. Cunha e A. P. C. da Silva, "Caracterizando polarização em redes sociais: Um estudo de caso das discussões no Reddit sobre as eleições brasileiras de 2018 e 2022," *Anais do XXX Simpósio Brasileiro de Sistemas Multimídia e Web (WebMedia)*, SBC, 2024, pp. 365–369.

[10] K. T. Poole e H. Rosenthal, "Patterns of congressional voting," *American Journal of Political Science*, vol. 35, n. 1, pp. 228–278, 1991.

[11] B. Shor e N. McCarty, "The ideological mapping of American legislatures," *American Political Science Review*, vol. 105, n. 3, pp. 530–551, 2011.

[12] J. B. Lewis, K. DeVine, K. T. Poole e H. Rosenthal, "Voteview: Congressional Roll-Call Votes Database," Disponível em: https://voteview.com/. Acesso em: jun. 2026.

[13] Câmara dos Deputados, "Dados Abertos: API de Dados da Câmara dos Deputados," Disponível em: https://dadosabertos.camara.leg.br/. Acesso em: jun. 2026.

[14] T. M. J. Fruchterman e E. M. Reingold, "Graph drawing by force-directed placement," *Software: Practice and Experience*, vol. 21, n. 11, pp. 1129–1164, 1991.