

Rede de co-autores de projetos de lei na câmara federal brasileira

Guilherme Aguilar de Oliveira

¹Departamento Acadêmico de Informática (DAINF)
Universidade Tecnológica Federal do Paraná (UTFPR)
Curitiba, PR – Brasil

guilhermeoliveira.2019@alunos.edu.br

1. Introdução

Os dados abertos da câmara de deputados federais permite várias análises da política brasileira¹. Uma possível análise, explorada nesse trabalho, é a de proposições de lei como uma rede. A rede foi construída da forma que os deputados são os nós e as proposições são as arestas. Dessa forma, foi analisado proposições desde 1998 a 2022. O código-fonte utilizado para construir essa pesquisa está em <https://github.com/guilherme9718/rede-proposicoes-camara-federal>.

2. Objetivos

Os objetivos dessa pesquisa são esses:

1. Identificar diálogo de autores de projetos e partidos
2. Identificar comunidades
3. Identificar autores e partidos centrais

3. A construção da rede

Para construir a rede, primeiramente foi criado um *script* automatizado para baixar os dados. três tabelas de dados em *csv* para cada ano dentre 1998 a 2022 (2022 incluso) foram baixadas². As tabelas são proposições por ano de apresentação, classificação temática das proposições e autores das proposições por ano de apresentação. O *script* só foi possível visto que a API dos dados está bem acessível³.

A partir disso, preprocessei os dados para transformá-los em redes. A principal tabela utilizada para formar a rede e transformar-se em análise é a de autores das proposições por ano de apresentação. O preprocessamento da tabela de autores se deu dessa forma:

1. Filtrar colunas relevantes para a tabela: id da proposição, id do autor, nome do autor, sigla do partido e sigla da UF do autor;
2. Adicionar uma coluna de ano para identificar o ano da proposição;
3. Retirar autores com id nulo;
4. Concatenar as tabelas de todos os anos;

¹Disponível em <http://dadosabertos.camara.leg.br/>. Acesso em: 03 dez 2022

²Os dados foram baixados em 14 de novembro de 2022

³Disponível em <https://dadosabertos.camara.leg.br/swagger/api.html#staticfile>. Acesso em 03 de dez. de 2022

5. Retirar projetos que contém somente um autor;

Dessa forma, com a tabela preprocessada, foi possível agrupar de dois a dois autores que participaram dos mesmos projetos, identificando quantos e quais projetos cada autor compartilha com outro. Esse agrupamento caracteriza as arestas do grafo. Já para os atributos dos nós, foi agrupado para cada autor todos os partidos e todas as UF desse autor no decorrer dos anos (muitos autores mudaram de partido e alguns poucos mudaram de unidade federativa ao longo dos mandatos). Porém, uma limitação desse trabalho é que não foi considerado os partidos que mudaram de nome sendo iguais, para corrigir isso poderia ser levado em conta o id do partido em vez da sigla.

4. A rede de deputados e proposições

De 1998 a 2022 houveram 2139 deputados diferentes que participaram como autores de projetos. Desses 2139, 1884 participaram de projetos com múltiplos autores. Notou-se que dois partidos nesse período não possuíam projetos com múltiplos autores: o PPR e o PSTU. Já sobre os projetos, nesses anos existiram por 505 mil projetos e somente 30 mil deles contém mais de um autor. Porém, o mais curioso disso é o fenômeno de 30 mil proposições que 1884 autores participaram produzir mais de 640 mil conexões entre os autores.

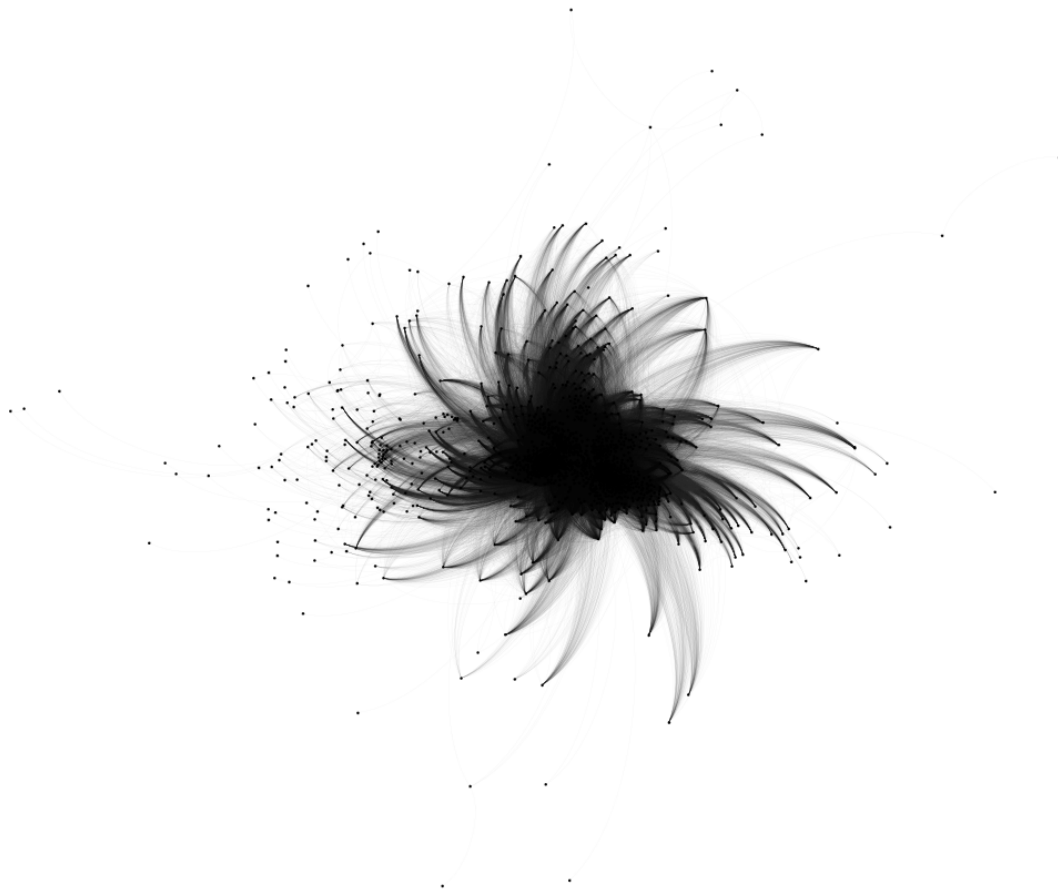


Figura 1. Rede com visualização Force Atlas 2

4.1. Métricas de grafo

São 1884 nós e 640958 links, uma rede extremamente conectada. A quantidade de links está somente a uma ordem de grandeza da quantidade máxima de links sendo por volta de 1,773 milhão. Além disso, o grau médio é de 680, muito maior que o limiar para a rede ser conectada que é $\ln(N) = 7,54$. Considerei os pesos das aresta como a quantidade de projetos compartilhado por dois nós e dessa forma o grau médio por peso é de 58816, significando que cada deputado participou, em média, de 58 mil projetos. Tenho a suposição de que deputados, mesmo com posições completamente distintas, participam de projetos conjuntos para aplicar sua pauta ou reduzir a influência da pauta do outro no projeto.

Existe uma diferença significativa de peso entre as arestas, então decidi analisar o grafo removendo arestas com um peso menor que um limite configurável. Percebeu-se que independente do limite, não havia comunidades e visualizações que surgiam a partir

disso, as arestas somente se concentravam em um bloco sólido central no grafo.

Os nós estão muito próximos uns dos outros da forma que a distancia média é 1,68, a distância máxima é 5 e o índice de clusterização é 0,841.

É possível observar na distribuição de graus (figura 2 e 3) que não há uma tendência em formar hubs. Classifico essa rede como *Small World*.

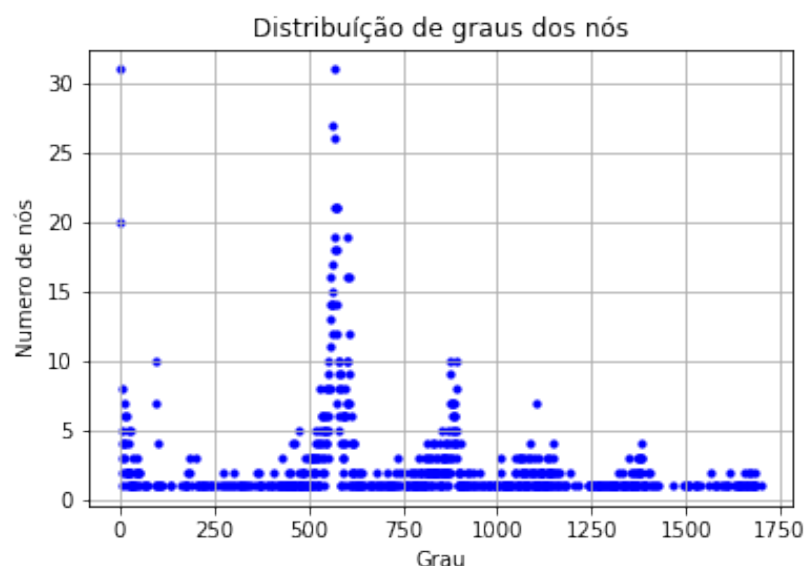


Figura 2. Distribuição de grau da rede

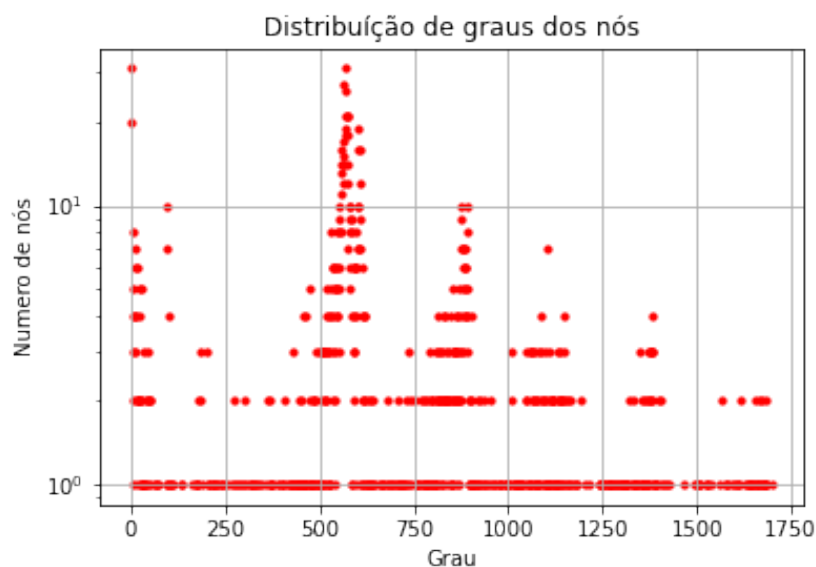


Figura 3. Distribuição de grau da rede em escala log

4.2. Centralidades

Aplicando o algoritmo de centralidades *page rank* e *betweenness* percebeu-se alguns padrões nos partidos e UF dos autores. Somando a centralidade de deputados para seus

respectivos partidos e unidades federativas. Nos partidos dos deputados mais centrais predominam partidos de direita e predominam estados com maior PIB com exceção de Pará e Bahia.

Tabela 1. Estados mais centrais por Page Rank e Betweenness

	Page Rank	Betweenness
SP	0.136	163267
PA	0.099	153972
BA	0.099	153972
MG	0.096	135130
RJ	0.094	97054
RS	0.059	93516
PR	0.056	86387
PE	0.048	64160

Tabela 2. Partidos mais centrais por Page Rank e Betweenness

	Page Rank	Betweenness
PFL	0.254	333279
PR	0.254	333279
PL	0.254	333279
PMDB	0.179	291632
PT	0.148	261861
PSDB	0.137	170609
PSD	0.117	139074
PAN	0.117	139074

4.3. Comunidades

Foi possível classificar seis comunidades na componente gigante pelo algoritmo de Louvain. Dentre as seis comunidades três representam mais de 90% dos nós.

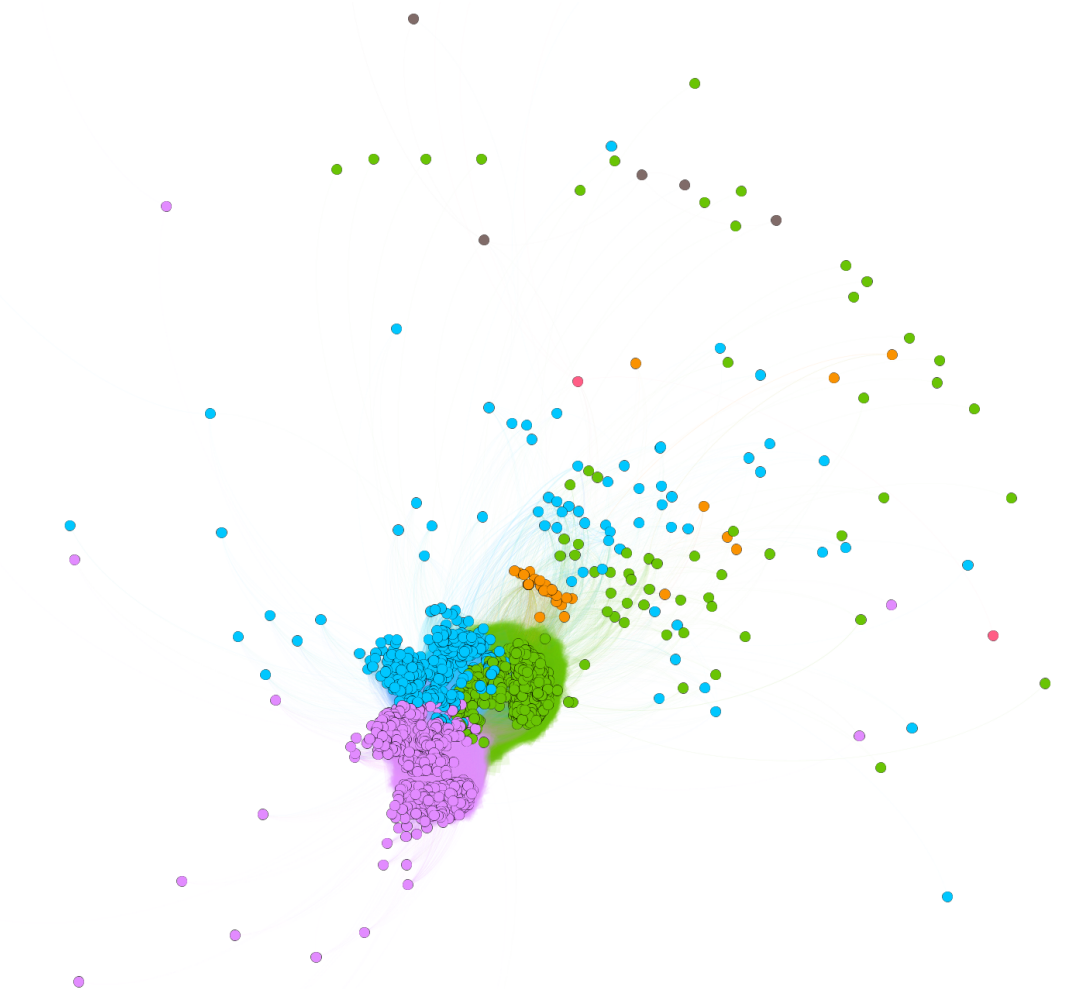


Figura 4. Rede com as comunidades classificadas pelo algoritmo Louvain

Em todas as comunidades, percebeu-se na tabela 3 forte presença de partidos de direita e uma forte presença do partido dos trabalhadores (PT).

Tabela 3. Comunidades com os maiores partidos representados nas comunidades

Presença	1 (rosa)	2 (verde)	4 (azul)
1º	PT	PMDB	PL
2º	PMDB	PSDB	PT
3º	PSDB	PFL	UNIÃO
4º	DEM	PP	PSD

5. Por tempo

Fazendo grafos por mandato (de quatro em quatro anos), pude observar as seguintes questões. De 1999 a 2002 existem pouquíssimas arestas comparado com outros anos sendo 421 nós e 7868 arestas. Pode haver alguma perda de dados ou poucas proposições

conjuntas (necessário investigar em pesquisas futuras). Já nos próximos 5 mandatos, os nós variam de 611 a 638 e arestas variam entre 151916 a 177859.

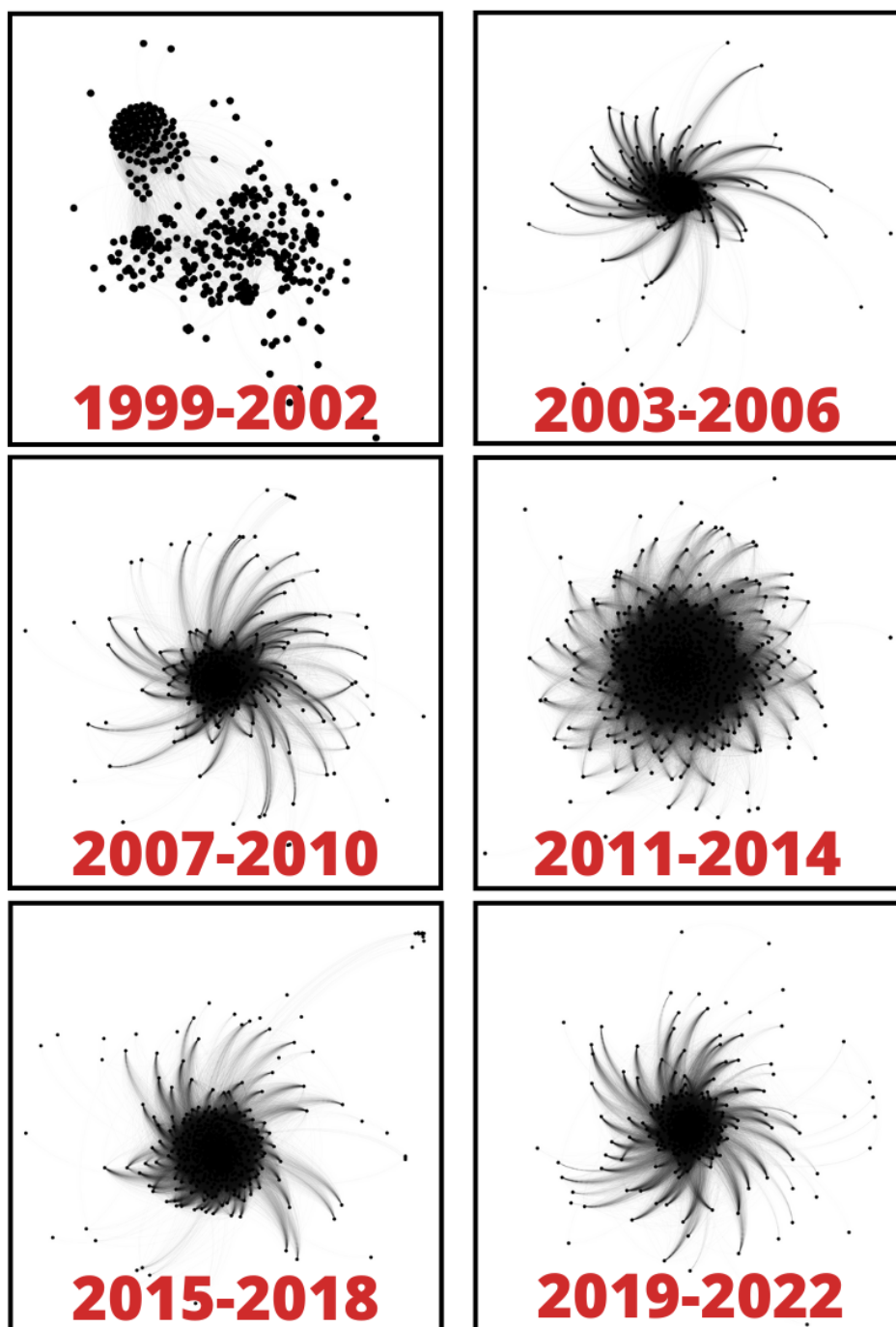


Figura 5. Grafos por mandato em visualização Force Atlas 2

6. Rede de partidos

Foi também construída uma rede de projetos compartilhados por partidos diretamente. Podemos considerar como a versão macro da rede de deputados. Dessa forma, essa rede contém 53 partidos e 939 arestas. Também é uma rede muito conectada relativo à rede de deputados. Os partidos são tão próximos entre eles que a maior distância é dois e o caminho médio é de 1,31.

Outro ponto interessante é que se observou que dentre os partidos mais centrais pelo *Betweenness* e *Page Rank*, aparecem muitos partidos de esquerda, ao contrário da nossa análise com os deputados pois nesse caso não depende totalmente do número de deputados de cada partido eleito na câmara.

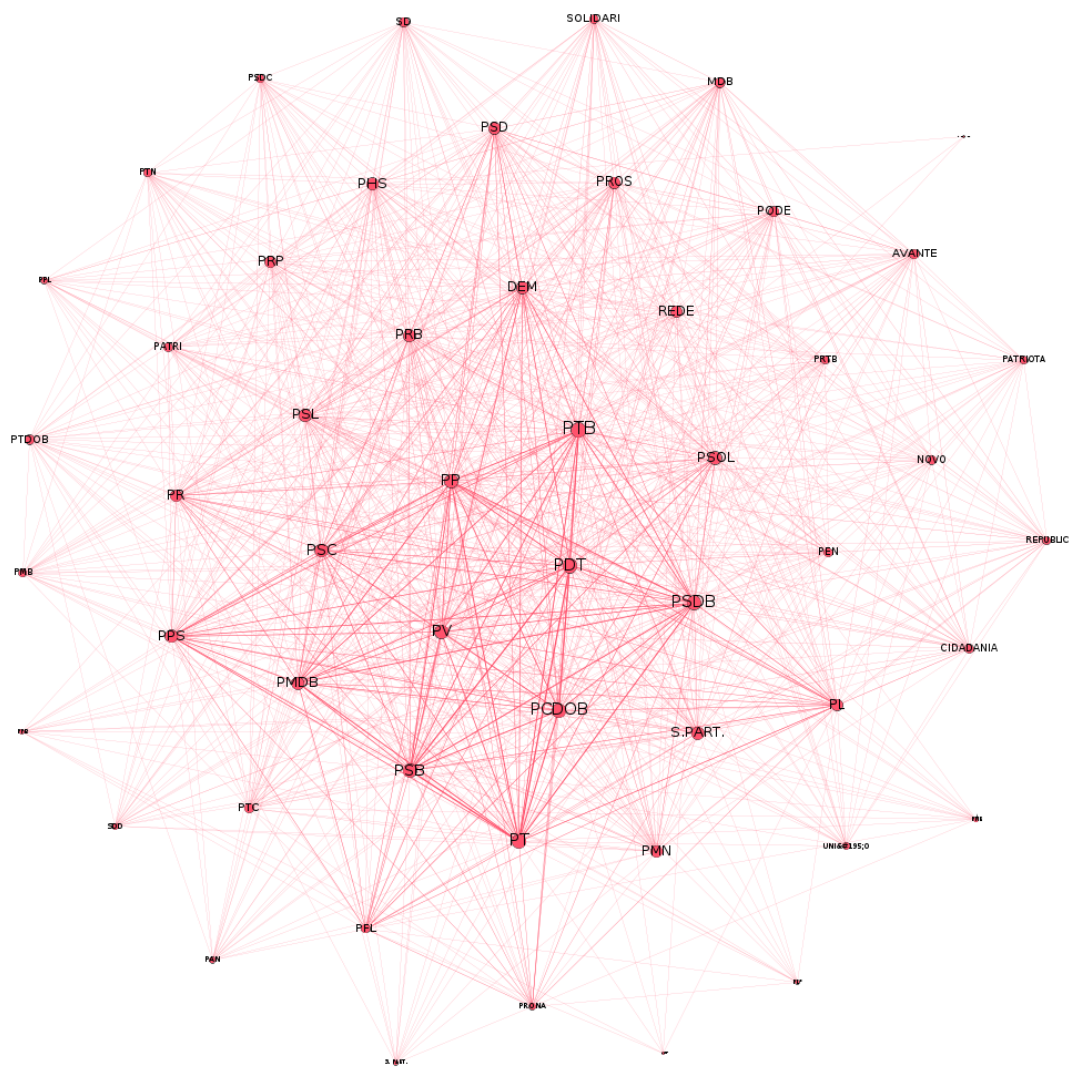


Figura 6. Partidos que compartilham projetos

Tabela 4. Partidos mais centrais na rede de partidos

Partido	Page Rank
PTB	0.027
PDT	0.026
PCDOB	0.026
PT	0.026
PSDB	0.026

7. Considerações finais

Observou-se que o diálogo entre partidos e deputados em proposições de lei é muito alto, principalmente entre 2003 a 2022. Porém se faz necessário estudar esse diálogo ao longo de janelas de tempo, por exemplo de 4 em 4 anos (mesmo que existem projetos que tem duração maior que 4 anos). É preciso estudar a falta de proposições entre 1998 a 2002, se é um fenômeno por falta de dados ou por algum motivo político.

Partidos de direita estão presentes em todas as comunidades e em conjunto com seus deputados são centrais. Porém, no macro partidos de esquerda também são centrais e participam de muitos projetos.

Sobre as limitações do projeto e sugestões para pesquisas futuras, levanto os seguintes apontamentos. Seria interessante analisar os temas das proposições e como circula esse tema na rede por partido, por UF entre outros. É preciso ter janelas de tempo melhor definidas. Também poderia aplicar técnicas de previsão de links para medir a probabilidade de um deputado se conectar com outro novamente, ou um partido se conectar com outro. Por fim, é preciso em futuros trabalhos uma análise embasada em ciência política e direito.