

Parrot Talk: Retweeting Among Twitter Users During the 2018 Brazilian Presidential Election

Jordan K. Kobellarz
jordan@alunos.utfpr.edu.br
Universidade Tecnológica Federal do Paraná
Curitiba, Brazil

Michelle Reddy
mireddy@stanford.edu
Stanford University
Palo Alto, USA

Alexandre R. Graeml
graeml@utfpr.edu.br
Universidade Tecnológica Federal do Paraná
Curitiba, Brazil

Thiago H. Silva
thiagoh@utfpr.edu.br
Universidade Tecnológica Federal do Paraná
Curitiba, Brazil

ABSTRACT

This study investigates users engagement on Twitter around presidential elections. While existing studies using Twitter investigate citizen engagement in the United States, there is limited work on online citizen engagement during elections in other contexts. As a result, we draw on the 2018 election in Brazil, one of the most contentious and polarizing elections in recent history. Overall, we look at the extent to which citizens come into contact with diverse ideological perspectives, and the extent of polarization among citizens. We investigate polarization within the network of retweets, as well as the role of media on serving as a broker between polarized users. We found that, in Brazil, the right-wing political network, although more homogeneous and polarized, proved to be less cohesive than the left wing political network, which in turn was more heterogeneous, by containing the largest proportion of neutral users. In addition, we found that traditional media still plays a neutral and bridging role among polarized networks, contradicting the common sense belief that the traditional media is also politically oriented. This maybe should already be expected, after all, when the traditional media is criticized by people from both sides, in a polarized matter, it is probably because it is not serving the interest of any of them.

CCS CONCEPTS

• **Information systems** → **World Wide Web; Web mining;** • **Human-centered computing** → *Collaborative and social computing;* • **Applied computing** → *Law, social and behavioral sciences.*

KEYWORDS

elections, political polarization, social networks, Twitter

ACM Reference Format:

Jordan K. Kobellarz, Alexandre R. Graeml, Michelle Reddy, and Thiago H. Silva. 2019. Parrot Talk: Retweeting Among Twitter Users During the 2018

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

WebMedia '19, October 29–November 1, 2019, Rio de Janeiro, Brazil

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6763-9/19/10...\$15.00

<https://doi.org/10.1145/3323503.3349559>

Brazilian Presidential Election. In *Brazilian Symposium on Multimedia and the Web (WebMedia '19), October 29–November 1, 2019, Rio de Janeiro, Brazil*. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3323503.3349559>

1 INTRODUÇÃO

O fenômeno da polarização política é representado pela presença simultânea de princípios, tendências e pontos de vista conflitantes em um mesmo contexto [9]. Estudos mostram uma preocupação em relação aos efeitos que esse fenômeno pode provocar na consolidação da opinião de indivíduos em redes sociais *online*, causados pela formação de grupos polarizados e altamente coesos, que podem reforçar o posicionamento político individual [5].

Nesse sentido, as eleições presidenciais do Brasil de 2018 foram um exemplo de cenário polarizado, onde as pesquisas [7, 14] evidenciavam a dicotomia política entre os eleitores. Na disputa, em segundo turno, estavam Jair Messias Bolsonaro, representando a possibilidade de ruptura com 15 anos do Partido dos Trabalhadores (PT) no poder, e Fernando Haddad, representando a continuidade dos governos petistas, depois de um breve período em que uma presidente eleita pelo PT fora substituída pelo seu vice-presidente, depois de sofrer um processo de *impeachment*. Nesse período, as redes sociais foram usadas amplamente para expressar e disseminar informações políticas entre os usuários.

Partindo da premissa de que usuários da *web* tendem a formar grupos coesos em relação a posicionamentos político-partidários [5] e levando em consideração o cenário político polarizado do segundo turno para a presidência do Brasil de 2018 [7, 14], essa pesquisa teve como objetivo analisar o fenômeno da polarização política, a partir do engajamento de usuários em mensagens de cunho político postadas na rede social do Twitter. Primeiramente, foram coletados 35.655.906 *tweets* referentes a esse período, sobre os quais foram aplicados métodos de aprendizado semi-supervisionado para classificação do posicionamento político de *hashtags*. Isso possibilitou posicionar cada usuário em um espectro de tendência política (variando de “direita” a “esquerda”). Em seguida, foi criada uma rede de *retweets* (ação de propagar aos seguidores no Twitter um *tweet*), que permitiu estudar até que ponto os usuários entram em contato com perspectivas ideológicas diversas e o grau de polarização entre os usuários.

Alguns dos principais resultados foram:

- A rede de *retweets* foi composta principalmente por usuários com posicionamento político tendendo para a direita,

representando 59,04% do total, seguido por usuários com tendência para a esquerda, representando 37,36% do total e 3,60% de usuários com tendência neutra.

- Ao identificar os usuários mais centrais na rede de *retweets*, observou-se que a mídia tradicional desempenhou um papel de ponte entre usuários com posicionamentos políticos opostos, diferentemente das mídias livres, que, em geral, serviram como pontes entre usuários com mesmo posicionamento político.
- Ao utilizar um algoritmo de detecção de comunidades de usuários na rede de *retweets*, observou-se a existência de duas grandes comunidades de usuários, que juntas representam 99,76% de todos os usuários. Uma delas, denominada “comunidade azul”, foi composta majoritariamente de usuários com posicionamento político de direita (97,15% dos usuários) e, a outra, denominada “comunidade vermelha”, foi composta majoritariamente de usuários com posicionamento político de esquerda (93,82% dos usuários).
- A comunidade vermelha incorporou a maioria dos usuários neutros. Isso mostra um maior grau de alinhamento de usuários neutros com outros que possuem posicionamento tendendo para a esquerda. Além disso, foi possível observar que usuários da comunidade vermelha criaram grupos mais coesos (fortemente conectados) em relação à comunidade azul, mesmo a comunidade vermelha sendo ligeiramente menos homogênea.

Acredita-se que essas descobertas têm implicações mais amplas para o estudo dos crescentes movimentos políticos conservadores globalmente [10] e podem ser aplicadas no estudo de mecanismos para minimizar os efeitos da polarização em grupos [19].

2 POLARIZAÇÃO EM REDES SOCIAIS

Não há consenso sobre o quê caracteriza a polarização política, mas muitos consideram a bimodalidade da distribuição de posicionamentos políticos de uma população uma condição necessária, mas não suficiente, para a caracterização da polarização política [9]. Essa distribuição foi explorada em estudos sobre cenários políticos usando dados de redes sociais *online*, principalmente com dados do Twitter [5, 12, 17]. Redes sociais *online* são extensivamente usadas para discutir assuntos políticos [11], mas trazem consigo uma preocupação sobre a polarização política, uma vez que seus usuários podem controlar os conteúdos que gostariam de consumir ou receber automaticamente conteúdos alinhados com suas preferências e, em consequência, aumentar sua exposição a argumentos que reforçam seu posicionamento político [5], levando a posicionamentos extremos ou enviesados.

A presença da polarização política em redes sociais pode ser explicada pelo conceito de homofilia, proveniente da sociologia, que é a tendência natural de indivíduos se associarem com outros que possuem características similares [15]. Contudo, um problema emerge a partir do momento em que os indivíduos passam a fazer parte de uma espécie de “câmara de eco” [2, 12], dado que, quanto mais polarizada é uma sociedade, mais cresce a chance dos eleitores serem expostos à notícias de fontes enviesadas [4], que acabam reforçando seu posicionamento e retroalimentando o fenômeno da polarização.

Em relação à detecção de polaridade em redes sociais, Conover *et al.* (2011), identificaram diferentes topologias em redes de usuários do Twitter, evidenciando uma baixa comunicação entre comunidades com posicionamentos políticos opostos na rede de *retweets* (ato de um usuário propagar uma mensagem de outro para sua rede), mas uma alta comunicação entre comunidades com posicionamentos políticos opostos na rede de menções (ato de mencionar outro usuário em uma mensagem). Outro estudo [12], por meio de uma análise multipartidária, identificou que apoiadores de um partido tendem a trocar mais mensagens com outros usuários do mesmo partido e menos com os de outros partidos, possivelmente com o intuito de evitar conflito, em vez de discutir abertamente seu posicionamento com opositores.

3 DADOS PRELIMINARES E PROCEDIMENTOS METODOLÓGICOS

Esta seção descreve como os *tweets* foram coletados (Seção 3.1), que procedimentos de classificação de *hashtags* foram adotados (Seção 3.2), como elas foram usadas para estimar o posicionamento político dos usuários (Seção 3.3) e, por fim, como a rede de *retweets* foi construída para permitir a análise de usuários de acordo com seu posicionamento político (Seção 3.4).

3.1 Coleta de *tweets*

Os dados usados neste trabalho foram obtidos por meio da API de *streaming* do Twitter, uma rede social e *microblog*, onde os usuários postam mensagens de até 280 caracteres chamadas de *tweets*, que podem ser retransmitidas por meio de *retweets* [5]. Cada *tweet* pode conter um ou mais rótulos, chamados *hashtags*, cuja principal função é descrever um tópico de discussão relacionado ao *tweet* [5].

Para obter *tweets* relacionados com as eleições presidenciais do Brasil de 2018, foram usadas as principais *hashtags* de cunho político precedentes ao segundo turno das eleições, sendo elas: #bolsonaro, #haddad, #elenao, #elesim, #lula, #TodosComJair, #Bolsonaro17, #Haddad_Fernando, #jairbolsonaro, #ManuelaDavila, #ManuNoJaburu, #HaddadPresidente, #GeneralMourao e #SomosBolsonaro17. A coleta dos *tweets* foi realizada durante os vinte dias que precederam as votações do segundo turno, no período entre 08/10/2018 e 28/10/2018, obtendo-se um total de 35.655.906 *tweets*.

3.2 Classificação de *hashtags*

O próximo passo do estudo foi realizar a classificação das 114.512 *hashtags* únicas usadas nos *tweets* coletados. Dado o grande volume de *hashtags*, não seria possível realizar a classificação de todas de forma manual em tempo hábil. Por isso, inicialmente, foram selecionadas 100 *hashtags*-semente, as quais apareceram com maior frequência. Cada uma das *hashtags*-semente foi classificada manualmente com relação ao seu posicionamento político, sendo “L” usado para representar um posicionamento de esquerda ou de continuidade, “R” de direita ou de ruptura e “?” um posicionamento incerto. Para isso, foram convidados três avaliadores não autores e não especialistas, para classificar individualmente e sem interferência com os outros avaliadores as 100 *hashtags*-semente. As características dos avaliadores são apresentadas na Tabela 1, na qual a coluna “Posicionamento” refere-se ao posicionamento político relatado explicitamente pelo avaliador. Depois que os avaliadores classificaram

Tabela 1: Características dos avaliadores

Ocupação	Escolaridade	Sexo	Idade	Posicionamento
Professor de geografia	Ensino superior	Fem.	45	Esquerda
Auxiliar administrativo	Ensino superior	Fem.	31	Esquerda
Nutricionista	Ensino superior	Fem.	24	Centro direita

as *hashtags*, os resultados foram comparados e mantidas apenas as *hashtags* cuja classificação foi unânime pelos três avaliadores. Isso aconteceu com 76 das 100 *hashtags*-semente, apresentadas na Tabela 2.

Partindo da premissa de que *hashtags* descrevendo um tópico em comum costumam ocorrer juntas em um mesmo *tweet* [5], para classificar as 114.412 *hashtags* ainda não classificadas, foi construída uma rede de coocorrências de *hashtags* em todos os *tweets*, incluindo as *hashtags*-semente, onde cada nó representava uma *hashtag* e cada aresta a ocorrência de uma *hashtag* juntamente com outra. Para isso foi usada a biblioteca NetworkX [13]. Sobre essa rede foi executado um algoritmo de aprendizado semi-supervisionado de funções harmônicas [20], cujo funcionamento se baseia em campos aleatórios gaussianos sobre uma matriz de adjacências da rede, em que o peso das arestas é usado para computar e indicar a similaridade entre os nós, possibilitando classificar nós não classificados a partir de nós já classificados. Dessa forma, foram classificadas 78.649 *hashtags*, representando 68,7% do total de *hashtags* ainda não classificadas. As *hashtags* que ficaram sem classificação por não coocorrerem com nenhuma outra *hashtag*, representando 31,3% do total, foram eliminadas da análise. A Tabela 3 apresenta a quantidade de *hashtags* obtidas em cada classe.

Como a pesquisa teve como objetivo a análise apenas de *tweets* com conteúdo político, foram removidos todos os *tweets* que não usaram pelo menos uma das 78.649 *hashtags* classificadas, restando 4.789.627 *tweets* para análise.

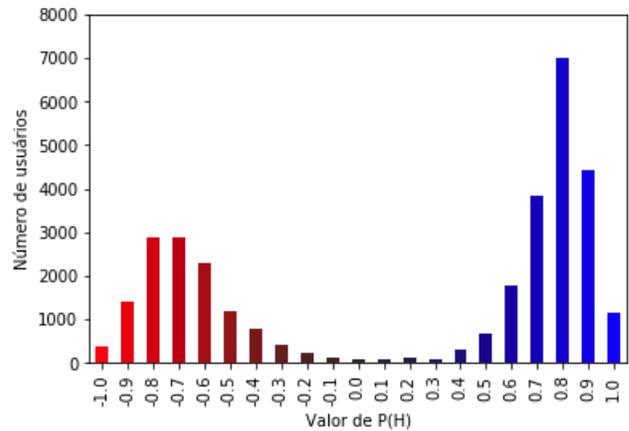
3.3 Classificação de usuários

Após a classificação das *hashtags*, o próximo passo foi classificar os usuários de acordo com seu posicionamento político. No total, foram encontrados 495.662 usuários únicos no conjunto de dados, dos quais foram removidos aqueles cujo número de *retweets* foi menor que 10 ou que não possuíam perfil em português, ou seja, usuários pouco ativos ou, provavelmente, estrangeiros discutindo sobre o cenário político brasileiro. Com a remoção, restaram 32.496 usuários, representando 6,56% do total. Dessa forma, para cada usuário, foi criada uma lista de *hashtags* usadas em todos os seus *tweets* e calculada sua polaridade ($P(H)$), dada pela Equação 1:

$$P(H) = \frac{|H_R| - |H_L|}{|H|}, \quad (1)$$

onde H é a lista com todas as *hashtags* usadas por um usuário, i.e., $H = H_R + H_L + H_?$, $H_R \in H$ as *hashtags* com posicionamento de direita, $H_L \in H$ as *hashtags* com posicionamento de esquerda e $H_? \in H$ as *hashtags* com posicionamento incerto. O resultado de $P(H)$ é um valor no intervalo contínuo $[-1.0; +1.0]$, em que valores positivos e próximos de +1.0 representam uma inclinação maior para a direita, valores negativos e próximos de -1.0 representam uma inclinação maior para a esquerda e valores próximos de 0.0

representam um posicionamento neutro ou não posicionamento. Uma escala de cores, apresentada na Figura 1, foi criada para facilitar a visualização dos resultados. Quanto mais vermelho claro, a polaridade é mais próxima de -1.0, quanto mais azul claro, a polaridade é mais próxima de +1.0 e quanto mais preto escuro, a polaridade é mais próxima de 0.0 - essa mesma escala é usada para todas as representações visuais apresentadas neste artigo. A Figura 2 apresenta a distribuição de usuários pela sua respectiva polaridade, dada por $P(H)$, por meio da qual é possível verificar visualmente a presença de uma curva bimodal, proporcionando forte indício de polaridade política, segundo [9].

Figura 1: Escala de cores de polaridade política.**Figura 2: Distribuição do número de usuários pela polaridade política.**

Cada usuário teve seu posicionamento calculado a partir do resultado de $P(H)$, sendo L (esquerda) os usuários com $P(H) \in [-1; -1/3]$, N (neutro) os usuários com $P(H) \in [-1/3; 1/3]$ e R (direita) os usuários com $P(H) \in]1/3; 1]$. A Tabela 4 mostra a distribuição de usuários de acordo com seu posicionamento.

3.4 Rede de *retweets*

Após identificar as polaridades e posicionamentos de todos os usuários, foi criada uma rede ligando um usuário ao outro por meio de *retweets*. A rede foi representada na forma de um grafo não-direcionado $G = (V, E)$, onde cada usuário é um nó $v \in V$ e o *retweet* uma aresta $e_{i,j} \in E$ entre o usuário v_i que postou o *tweet* original e o usuário v_j que fez o *retweet*. O peso da aresta representa a quantidade de *retweets* entre um usuário e outro. Arestas ligando um nó a ele próprio, ou seja, de um usuário que fez um *retweet* do seu próprio *tweet*, foram eliminadas da rede e, posteriormente, eliminados os nós isolados da rede.

Tabela 2: Hashtags-semente classificadas em incerto, direita e esquerda por três avaliadores independentes. Todas possuem concordância entre os três avaliadores.

Posicionamento	Hashtags
? (incerto)	#AndradeJaEra, #Eleicoes2018, #FolhaFakeNews, #UOLnasUrnas, #Brasil, #InformacaocontraoAchismo, #HoraDaVirada, #VemProDebate, #ViraVoto, #VotoEmCedula, #RodaViva, #AFalhaéCafonérrima, #G1, #SeuVotoMePõeEmRisco, #FAKE, #NãoAceitaremosFraude, #SanatórioGeral, #EAgoraTSE, #Folha, #NaoAceitaremosFraude, #IbopeFake, #viravoto, #democracia, #SuasticaFake, #FakeNews, #Brazil, #SomosTodosReginaDuarte, #delegadofrancischini, #Eleicoes2018 e #BrasilDecide.
R (direita)	#NasRuasComBolsonaro, #BolsonaroSim, #MudaBrasil17, #Bolsonaro17, #Bolsonaro, #BrasilComBolsonaro17, #NordesteÉ17, #BolsonaroPresidente, #NemVemQueNaoTemPT, #PTNuncaMais, #BolsonaroPresidenteEleito, #LulaTaPresoBabaca, #HaddadNãoÉCristão, #B17, #Nordeste17, #OLulaTaPresoBabaca, #EleSim, #bolsonaro17, #PTNão, #PTnão, #FolhaP*****DoPT (Conteúdo impróprio), #AcabouPiranhagemPT, #BolsonaroPresidente17, #elesim, #bolsonaro e #PTnao.
L (esquerda)	#Haddad13, #Caixa2doBolsonaro, #HaddadPresidente, #HaddadSim, #EleNao, #BrasilViraHaddad, #EleNão, #AgoraÉHaddad, #BolsonaroNão, #ViraVotoHaddad13, #CassaçãoDoBolsonaro, #LulaLivre, #bolsonaroCagao, #Caixa2DoBolsonaro, #Haddad, #elenao, #ViraVirouHaddad13, #ELENAO, #MaisLivrosMenosArmas e #haddadpresidente.

Tabela 3: Sumarização das Hashtags classificadas entre incerto (?), direita (R) e esquerda (L) com a metodologia proposta.

Posicionamento	Hashtags	Percentual
Incerto (?)	3.623	4, 6%
Direita (R)	49.718	63, 2%
Esquerda (L)	25.308	32, 2%
Total	78.649	100%

Tabela 4: Sumarização dos usuários classificados de acordo com o seu posicionamento político.

Posicionamento	Usuários	Percentual
Esquerda (L)	12.141	37,36%
Neutro (N)	1.169	3,60%
Direita (R)	19.186	59,04%
Total	32.496	100%

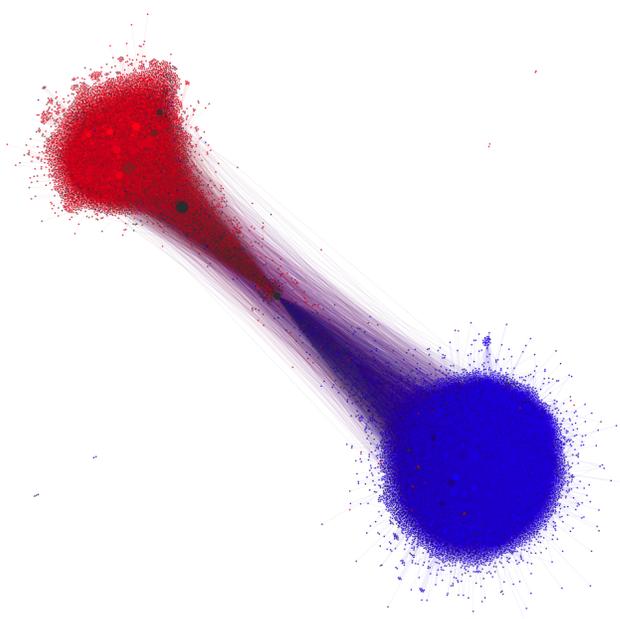
4 RESULTADOS

Nesta seção são apresentados os resultados obtidos para a rede de *retweets* apresentada anteriormente. A Seção 4.1 apresenta uma visão geral dessa rede. A Seção 4.2 analisa a importância de usuários na rede com métricas de centralidade. A Seção 4.3 apresenta resultados para as duas maiores comunidades de usuários formadas na rede, incluindo a descrição dos usuários mais centrais e a distribuição de polaridades em cada uma.

4.1 Visão geral da rede de *retweets*

Ao total, a rede consistiu de 32.086 nós, 515.212 arestas e grau médio de 32, 11 (quantidade média de arestas ligadas a cada nó na rede). Uma visualização da rede foi gerada usando o *software* Gephi [3], na qual a polaridade dos usuários foi representada pela escala de cores da Figura 1 e o grau de intermediação de cada nó representada pelo seu diâmetro (ver Figura 3).

A rede de *retweets* foi composta majoritariamente por usuários com posicionamento de direita, representando 59, 67% do total, depois por usuários de esquerda, representando 36, 84% do total e por apenas 3, 49% de usuários neutros, como mostra a Tabela 5.

Figura 3: Rede de *retweets*.**Tabela 5: Posicionamento político dos usuários na rede de *retweets*.**

Posicionamento	Usuários	Percentual
Esquerda (L)	11.820	36,84%
Neutro (N)	1.121	3,49%
Direita (R)	19.145	59,67%
Total	32.086	100%

4.2 Análise de centralidade

Para analisar a relevância dos usuários, foram extraídas as métricas de centralidade de grau e intermediação da rede de *retweets*. A centralidade de grau refere-se à quantidade de arestas ligadas a um nó, nesse caso, representando a quantidade de *retweets* que um usuário fez ou a quantidade de *tweets* desse usuário de que foram feitos *retweets* por outros usuários [8]. Já a centralidade de

Tabela 6: Usuários com maior valor para as métricas de centralidade de grau e intermediação na rede de *retweets* completa.

Usuário	P(H)	Tipo	L%	N%	R%	Grau
JoelAlexandreM	0,8	Ativista	0,01%	0,25%	99,74%	61.729
Jouberth19	0,9	Ativista	0,02%	0,12%	99,86%	21.546
Ro_Moller	0,7	Ativista	0,05%	0,23%	99,72%	18.893
RenovaMidia	0,6	Mídia livre	0,1%	0,35%	99,55%	16.327
J_LIVRES	-0,5	Mídia livre	95,83%	3,53%	0,64%	16.083
maepatriota	0,8	Ativista	0,05%	0,18%	99,77%	13.110
RoseBarros__	1,0	Ativista	0,02%	0,24%	99,74%	10.871
OsvladPereira	0,9	Ativista	0,02%	0,21%	99,77%	10.132
delucca	-0,9	Ativista	97,77%	2,05%	0,18%	9.570
LulaOficial	-0,8	Personalidade	98,29%	1,42%	0,29%	8.634
Usuário	P(H)	Tipo	L%	N%	R%	Intermediação
g1	0,0	Mídia tradicional	46,95%	16,08%	36,97%	0,117877
J_LIVRES	-0,5	Mídia livre	95,83%	3,53%	0,64%	0,069908
JoelAlexandreM	0,8	Ativista	0,01%	0,25%	99,74%	0,054183
Ro_Moller	0,7	Ativista	0,05%	0,23%	99,72%	0,038414
JornalOGlobo	0,0	Mídia tradicional	76,24%	15,55%	8,2%	0,037053
Jouberth19	0,9	Ativista	0,02%	0,12%	99,86%	0,037045
delucca	-0,9	Ativista	97,77%	2,05%	0,18%	0,036635
maepatriota	0,8	Ativista	0,05%	0,18%	99,77%	0,030962
LulaOficial	-0,8	Personalidade	98,29%	1,42%	0,29%	0,030535
RenovaMidia	0,6	Mídia livre	0,1%	0,35%	99,55%	0,029635

intermediação de um nó v é a soma da fração de todos os pares de caminhos mais curtos que passam por v [8]. Na rede de *retweets* em estudo, podemos olhar para essa métrica como uma indicação dos usuários que podem atuar como pontes na propagação de conteúdo entre grupos distintos (não necessariamente com posicionamento político distinto).

Dessa forma, para a rede de *retweets*, foram obtidos os 10 usuários com maior valor para as métricas de centralidade de grau e intermediação, cujos valores são apresentados na Tabela 6. Cada usuário na tabela foi classificado de acordo com seu tipo (ativista, personalidade, político, mídia livre ou mídia tradicional). Os ativistas são usuários que defendem um posicionamento, partido ou candidato. Personalidades são celebridades ou pessoas de notável conhecimento público. Políticos são candidatos ou indivíduos com cargo político. Mídias livres são canais de comunicação que permitem a expressão de opiniões de forma livre e espontânea. E, por fim, mídias tradicionais são canais de comunicação que focam na transmissão de fatos procurando evitar viés de posicionamento. Além dessa classificação, foi obtido o percentual de trocas de mensagens com vizinhos de cada nó, de acordo com seu posicionamento, cujos dados são apresentados nas colunas $L\%$, $N\%$ e $R\%$.

Dentre os usuários, destaca-se @JoelAlexandreM, que teve os maiores valores para as duas métricas de centralidade. Enquanto essa pesquisa estava sendo realizada, esse usuário teve seu nome alterado para @PresidenBr no site do Twitter. Mesmo assim, para efeitos de análise, foi mantido seu nome original. Na rede social, ele se apresenta como “Presidente” e usa sua conta para divulgar ações e discursar em defesa do presidente Bolsonaro, muitas vezes em tom sarcástico para atacar seus opositores. Apesar do nome, essa não é uma conta oficial e teve sua criação em janeiro de 2018, no mesmo ano das eleições, sendo a conta mais nova dentre todas as outras com maior valor para cada métrica de centralidade na rede de *retweets*.

Destaca-se também que, dentre os usuários com maior centralidade de grau, foram encontradas duas mídias livres, sendo elas

@J_LIVRES e @RenovaMidia, de posicionamentos de esquerda e direita, respectivamente. Já dentre os usuários com maior centralidade de intermediação, foram encontradas duas mídias tradicionais, representadas pelos usuários @g1 e @JornalOGlobo, ambos com posicionamento neutro, e duas mídias livres, representadas pelos usuários @J_LIVRES e @RenovaMidia, que também estavam dentre as com maior centralidade de grau, como já mencionado. Esses dados sugerem que as mídias livres tenham ganhado relevância no debate político, pelo fato de apresentarem posicionamento claro e não neutro em relação às mídias tradicionais, que tendem a se manter neutras.

4.3 Comunidades

Uma vez que foi identificado um indício de polarização na rede de *retweets*, dado pela distribuição bimodal de usuários em relação à sua polaridade (Figura 2), o próximo passo foi identificar as comunidades formadas pelos usuários dentro da rede de *retweets* para verificar a existência de grupos mais ou menos coesos de usuários com mesmo posicionamento. Para detectar comunidades na rede de *retweets*, foi usado o algoritmo de propagação de rótulos (*Label Propagation*) [6], que funciona inicializando cada um dos nós com uma classificação única e, a partir disso, de forma iterativa, atribui a cada nó a classificação que aparece com maior frequência na sua vizinhança. Em caso de empate, é escolhida uma classificação aleatória entre as que empataram. O algoritmo para quando todos os nós deixam de ter sua classificação modificada. Como resultado, foram detectadas 35 comunidades, dentre as quais apenas as duas maiores juntas incluíam 99,76% de todos os usuários, sendo a maior composta majoritariamente por usuários de direita (97,15%) e a segunda maior majoritariamente por usuários de esquerda (93,82%), como mostra a Tabela 7.

As duas maiores comunidades serão tratadas a seguir por meio do seu respectivo posicionamento majoritário, sendo a comunidade azul a que possui a maioria dos membros de direita e a

Tabela 7: Posicionamento político dos usuários nas duas maiores comunidades.

Comunidade	Usuários	L%	N%	R%
azul	19.645	1,02%	1,83%	97,15%
vermelha	12.364	93,82%	6,02%	0,16%

Tabela 8: Características das redes das duas maiores comunidades.

Comunidade	Nós	Arestas	Grau médio
azul	19.645	386.028	39,30
vermelha	12.364	124.922	20,20

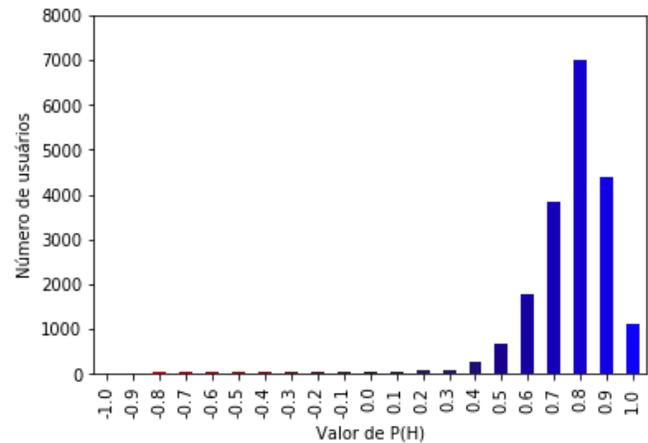
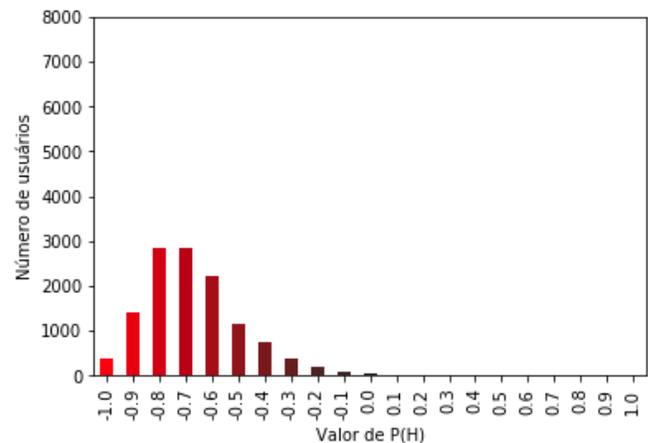
comunidade vermelha a que possui a maioria dos membros de esquerda. Para as análises, foram gerados dois subgrafos a partir da rede de *retweets* para cada uma das comunidades, criando uma rede para a comunidade azul e outra para a comunidade vermelha, cujas características são apresentadas na Tabela 8.

A comunidade azul se apresentou altamente homogênea em relação ao posicionamento dos usuários - apenas 1,83% dos nós eram neutros e 1,02% de esquerda, como mostra a Tabela 7. Já a comunidade vermelha, comparativamente, também se apresentou homogênea, mas com uma maior quantidade de usuários neutros, representando 6,02% do total, contra os 1,83% da comunidade azul. Essa distribuição pode indicar uma maior afinidade de usuários neutros com a esquerda política.

Para verificar a coesão das comunidades, foi calculado o coeficiente médio de agrupamento de cada comunidade [18], uma métrica que indica o grau com que os nós de um grafo tendem a se agrupar. Para a comunidade azul, o coeficiente médio de agrupamento foi de $3,47 \cdot 10^{-3}$, enquanto para a comunidade vermelha foi de $4,56 \cdot 10^{-3}$. Isso indica que a comunidade vermelha, mesmo tendo uma maior concentração de usuários neutros em relação à azul, ainda continua sendo a mais coesa em termos de coeficiente médio de agrupamento.

As Figuras 4 e 5 apresentam a distribuição de usuários pela polaridade ($P(H)$) na comunidade azul e vermelha, respectivamente. Apesar das duas distribuições terem um formato similar, é possível observar um incremento na concentração de usuários entre as polaridades +0,7 e +0,8 na comunidade azul de 83,11% (de 3.825 para 7.004), o que, comparativamente, não é evidente na comunidade vermelha, entre as polaridades -0,7 e -0,8, com uma redução de 0,17% (de 2.850 para 2.845). Além disso, pode-se observar que a proporção de usuários com polaridades +0,9 e +1,0 é maior do que a proporção de usuários com polaridades -0,9 e -1,0.

Da mesma forma como foram calculadas as métricas de centralidade na rede de *retweets*, também foram calculadas essas métricas para as comunidades azul e vermelha, como é apresentado nas tabelas 9 e 10, respectivamente. Os usuários mais centrais da comunidade azul tiveram uma maior representatividade na rede completa, em relação aos usuários mais centrais da comunidade vermelha. Esse fato ligado à disparidade da métrica de centralidade de grau entre os usuários mais centrais da comunidade azul em relação à comunidade vermelha, leva à intuição de que houve um maior

Figura 4: Distribuição de usuários pela polaridade política na comunidade azul.**Figura 5: Distribuição de usuários pela polaridade política na comunidade vermelha.**

engajamento na comunidade azul, mesmo havendo baixa coesão em relação ao coeficiente médio de agrupamento.

Foram identificados dois políticos (@humbertocostapt e @DeputadoFederal) e duas personalidades públicas (@LulaOficial e @zehdeabreu) dentre os usuários mais centrais na comunidade vermelha. Em comparação, nenhum político ou personalidade pública foi identificado dentre os usuários mais centrais na comunidade azul. Essa diferença mostra que usuários ativistas tiveram um papel mais ativo no engajamento de usuários de direita, enquanto, na comunidade vermelha, o engajamento se dividiu entre ativistas, políticos e personalidades públicas.

Assim como na rede de *retweets* completa, o usuário com maior valor para todas as métricas de centralidade na comunidade azul foi @JoelAlexandreM, um ativista de direita. Já na comunidade vermelha, o usuário com maior valor para todas as métricas de centralidade foi @J_LIVRES, uma mídia livre de esquerda, com uma

Tabela 9: Usuários com maior valor para as métricas de centralidade de grau e intermediação na rede da comunidade azul.

Usuário	P(H)	Tipo	L%	N%	R%	Grau
JoelAlexandreM	0,8	Ativista	0,01%	0,24%	99,75%	61.722
Jouberth19	0,9	Ativista	0,0%	0,12%	99,87%	21.543
Ro_Moller	0,7	Ativista	0,01%	0,22%	99,77%	18.884
RenovaMidia	0,6	Mídia livre	0,04%	0,32%	99,64%	16.312
maepatriota	0,8	Ativista	0,02%	0,18%	99,79%	13.107
RoseBarros_	1,0	Ativista	0,01%	0,22%	99,77%	10.868
OsvladPereira	0,9	Ativista	0,0%	0,21%	99,79%	10.130
paulacamara_	1,0	Ativista	0,0%	0,34%	99,66%	8.252
Mark_Senk	0,9	Ativista	0,0%	0,28%	99,72%	5.017
SaabBrasil	0,8	(não encontrado)	0,02%	0,21%	99,77%	4.346
Usuário	P(H)	Tipo	L%	N%	R%	Intermediação
JoelAlexandreM	0,8	Ativista	0,01%	0,24%	99,75%	0,088806
Jouberth19	0,9	Ativista	0,0%	0,12%	99,87%	0,063074
Ro_Moller	0,7	Ativista	0,01%	0,22%	99,77%	0,061618
maepatriota	0,8	Ativista	0,02%	0,18%	99,79%	0,050183
g1	0,0	Mídia tradicional	6,57%	14,98%	78,45%	0,045127
RenovaMidia	0,6	Mídia livre	0,04%	0,32%	99,64%	0,042680
paulacamara_	1,0	Ativista	0,0%	0,34%	99,66%	0,042463
RoseBarros_	1,0	Ativista	0,01%	0,22%	99,77%	0,037164
OsvladPereira	0,9	Ativista	0,0%	0,21%	99,79%	0,034974
M4RC3LO_O	0,5	Ativista	0,0%	0,26%	99,74%	0,024957

Tabela 10: Usuários com maior valor para as métricas de centralidade de grau e intermediação na rede da comunidade vermelha.

Usuário	P(H)	Tipo	L%	N%	R%	Grau
J_LIVRES	-0,5	Mídia livre	96,54%	3,39%	0,07%	15.939
delucca	-0,9	Ativista	97,97%	2,03%	0,0%	9.543
LulaOficial	-0,8	Personalidade	98,6%	1,39%	0,01%	8.591
zehdeabreu	-0,9	Personalidade	97,98%	2,01%	0,01%	7.912
DeputadoFederal	-0,5	Político	97,86%	2,07%	0,08%	7.649
OBrasilFeliz	-0,8	Ativista	98,56%	1,43%	0,02%	6.312
humbertocostapt	-0,9	Político	98,83%	1,17%	0,0%	4.375
DilmaResiste	-0,8	Ativista	99,01%	0,99%	0,0%	4.346
diImabr	-0,8	Ativista	97,22%	2,78%	0,0%	3.920
joomikhail	-0,8	Ativista	98,09%	1,88%	0,03%	3.881
Usuário	P(H)	Tipo	L%	N%	R%	Intermediação
J_LIVRES	-0,5	Mídia livre	96,54%	3,39%	0,07%	0,114840
delucca	-0,9	Ativista	97,97%	2,03%	0,0%	0,098838
kallilolv	-0,9	Ativista	96,43%	3,57%	0,0%	0,063711
zehdeabreu	-0,9	Personalidade	97,98%	2,01%	0,01%	0,063635
LulaOficial	-0,8	Personalidade	98,6%	1,39%	0,01%	0,062630
diImabr	-0,8	Ativista	97,22%	2,78%	0,0%	0,056179
DeputadoFederal	-0,5	Político	97,86%	2,07%	0,08%	0,053371
anabeat05826201	-1,0	Ativista	96,02%	3,82%	0,16%	0,050370
OBrasilFeliz	-0,8	Ativista	98,56%	1,43%	0,02%	0,046576
joomikhail	-0,8	Ativista	98,09%	1,88%	0,03%	0,038888

alta disparidade de valores em comparação com o segundo colocado para cada uma das métricas na comunidade vermelha. Esse resultado oferece indícios de que os usuários de direita tiveram maior engajamento com mídias livres do que com mídias tradicionais no debate político durante o período que antecedeu as votações para segundo turno. É importante notar que, ao estudar a rede completa, as mídias tradicionais tiveram um papel fundamental para fazer uma ponte entre usuários com posicionamento político diferentes. Como pode-se observar na Tabela 6, somente as mídias tradicionais proporcionaram engajamento de um número significativo de usuários de esquerda e de direita com conteúdo neutro. Além disso, essas foram as mídias que mais engajaram com o público neutro, sendo 16,08% para o usuário @g1 e 15,55% para o usuário @JornalOGlobo.

Em relação às mídias livres, somente uma apresentou um leve engajamento com o público neutro, sendo ela o usuário @J_LIVRES, uma mídia livre de esquerda, que se engajou com 3,53% de usuários neutros e 95,83% de usuários de esquerda. Já a outra mídia livre, representada pelo usuário @RenovaMidia, de direita, se engajou majoritariamente com o público de direita (99,55%). O restante dos usuários, incluindo ativistas, personalidades e políticos, engajaram apenas o público mais alinhado com seu respectivo posicionamento, fato fortemente associado ao conceito de homofilia [15].

5 DISCUSSÃO GERAL

Os resultados apresentados mostram uma grande polarização da rede estudada, na qual é notável o papel que algumas mídias tradicionais desempenharam na conexão de usuários com posicionamento

políticos opostos. As pessoas ainda parecem buscar nas grandes mídias tradicionais informações confiáveis, ainda que o jornalismo tradicional enfrente uma crise de confiança em que todos parecem considerá-lo enviesado na direção oposta das próprias convicções. Diante de um contexto com um elevado número de notícias falsas, veiculadas principalmente pelas redes sociais, como foi o caso da eleição estudada [16], grandes mídias tradicionais podem ter ganhado um papel ainda mais relevante, representado uma alternativa segura para compartilhamento de conteúdo por parte de alguns usuários com diferentes posicionamentos políticos.

Os resultados mostram que, no Brasil, as redes políticas de pessoas mais alinhadas com a direita tiveram uma atuação mais forte do que as redes políticas de pessoas mais alinhadas com a esquerda. Esse fato pode estar ligado a um possível movimento de ruptura com os 15 anos de governo do Partido dos Trabalhadores, representado pela esquerda política. Além disso, os achados deste estudo podem ter implicações mais amplas para o estudo de movimentos políticos conservadores que vêm ocorrendo pelo mundo, como foi o caso do Brexit, por exemplo.

Apesar dos resultados, existem possíveis limitações no estudo. A remoção dos usuários com menos de 10 *retweets*, para capturar usuários mais ativos e aferir seu engajamento, pode ter levado à remoção de grandes veículos jornalísticos, uma vez que não é comum que estes realizem *retweet* de mensagens. Sabe-se também que a população de usuários do Twitter não representa de forma justa toda a sociedade, na qual usuários idosos e de baixa renda, por exemplo, tendem a ser sub-representados. Além disso, a classificação de usuários de acordo com seu posicionamento entre esquerda, direita e neutro, tende a limitar o entendimento de uma realidade muito mais complexa, onde coexistem usuários que ao mesmo tempo são favoráveis ou contra determinadas características de cada um dos lados. A existência de um maior volume de usuários de direita em relação a usuários de esquerda na rede, assim como o posicionamento político dos avaliadores das *hashtags*, podem ter exercido influência nos resultados. Não menos importante, vários autores relatam que a rede social do Twitter não favorece o diálogo político por causa do limite de 280 caracteres, portanto, de forma empírica, essa rede pode ser mais propensa à formação de "câmaras de eco", dada a dificuldade em redigir argumentos completos em poucos caracteres.

6 CONCLUSÕES

Este trabalho analisou em escala nacional o fenômeno da polarização política na eleição presidencial de 2018 do Brasil. Para isso foi criada uma rede com base no engajamento de usuários em mensagens de cunho político transmitidas no Twitter. Isso permitiu estudar o grau de polarização entre os usuários. Observa-se que usuários com maior alinhamento político com a "direita" são a maioria na rede estudada, sendo ainda, no geral, mais atuantes na rede. Além disso, a rede estudada é bastante polarizada, composta por essencialmente dois grandes grupos com alinhamentos políticos distintos. Observa-se, ainda, que a mídia tradicional apresentou um papel de destaque como ponte entre usuários com diferentes posições políticas. Esses e outros resultados deste estudo são importantes para entender melhor o fenômeno da polarização política em redes sociais *online*, um passo importante para a criação de

mecanismos que possibilitem mitigar o efeito de "câmara de eco", várias vezes reportado por outros pesquisadores [1, 5, 12].

AGRADECIMENTOS

Este trabalho foi parcialmente apoiado pelo projeto CNPq-URBCOMP (processo 403260/2016-7), CAPES, FAPESP e Fundação Araucária.

REFERÊNCIAS

- [1] Pablo Barberá. 2015. *How Social Media Reduces Mass Political Polarization. Evidence from Germany, Spain, and the U.S.* Technical Report. www.pablobarbera.com
- [2] Pablo Barberá, John T. Jost, Jonathan Nagler, Joshua A. Tucker, and Richard Bonneau. 2015. Tweeting From Left to Right: Is Online Political Communication More Than an Echo Chamber? *Psychological Science* 26, 10 (2015), 1531–1542. <https://doi.org/10.1177/0956797615594620> PMID: 26297377.
- [3] Mathieu Bastian, Sebastien Heymann, and Mathieu Jacomy. 2009. Gephi: An Open Source Software for Exploring and Manipulating Networks. In *Proc. of International AAAI Conference on Weblogs and Social Media*. San Jose, CA, USA. <http://www.aaai.org/ocs/index.php/ICWSM/09/paper/view/154>
- [4] Dan Bernhardt, Stefan Krassa, and Mattias Polborn. 2008. Political polarization and the electoral effects of media bias. *Journal of Public Economics* 92, 5 (2008), 1092 – 1104. <https://doi.org/10.1016/j.jpubeco.2008.01.006>
- [5] M Conover, J Ratkiewicz, and M Francisco. 2011. Political polarization on twitter. In *Proc. of International Conference on Weblogs and Social Media*. Barcelona, Spain, 89–96. <https://doi.org/10.1021/ja202932e>
- [6] Gennaro Cordasco and Luisa Gargano. 2010. Community detection via semi-synchronous label propagation algorithms. In *Proc. of IEEE International Workshop on: Business Applications of Social Network Analysis (BASNA)*. IEEE, Bangalore, India, 1–8.
- [7] DATAFOLHA. 2018. *Intenção de voto para presidente da República – 2º turno – 18/10*. Retrieved October 18, 2018 from <http://media.folha.uol.com.br/datafolha/2018/10/19/692a4086c399805ae503454cf8ed0d361V.pdf>
- [8] David Easley and Jon Kleinberg. 2010. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press.
- [9] Morris P. Fiorina and Samuel J. Abrams. 2008. Political Polarization in the American Public. *Annual Review of Political Science* 11, 1 (jun 2008), 563–588. <https://doi.org/10.1146/annurev.polisci.11.053106.153836>
- [10] Gareth Fowler, Arthur Larok, Pawet Marczewski, Vijayan MJ, Ghia Nodia, Natalia Shapovalova, Janjira Sombatpoonsiri, Marisa von Bulow, and Ozge Zihnioglu. 2018. *The Mobilization of Conservative Civil Society*. Carnegie Endowment for International Peace.
- [11] Todd Graham, Marcel Broersma, Karin Hazelhoff, and Guido van 't Haar. 2013. Between Broadcasting Political Messages and Interacting With Voters. *Information, Communication & Society* 16, 5 (2013), 692–716. <https://doi.org/10.1080/1369118X.2013.785581>
- [12] Anatoliy Gruzd. 2014. Investigating Political Polarization on Twitter: A Canadian Perspective. *Policy and Internet* 6, 1 (2014), 28–45. <https://doi.org/10.1002/1944-2866.POI354>
- [13] Aric Hagberg, Pieter Swart, and Daniel S Chult. 2008. *Exploring network structure, dynamics, and function using NetworkX*. Technical Report. Los Alamos National Lab. (LANL).
- [14] IBOPE. 2018. *Pesquisa de opinião pública sobre assuntos políticos/administrativos*. Retrieved October 11, 2018 from http://www.ibopeinteligencia.com/arquivos/JOB_0011-10_BRASIL%20-%20Relat%C3%B3rio%20de%20tabelas.pdf
- [15] Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology* 27, 1 (2001), 415–444. <https://doi.org/10.1146/annurev.soc.27.1.415>
- [16] Alessandra Monnerat, Matheus Riga, and Pedro Ramos. 2018. *Fake news devem causar impacto em eleições de 2018*. Retrieved May 12, 2019 from <http://infograficos.estadao.com.br/focas/politico-em-construcao/materia/fake-news-devem-causar-impacto-em-eleicoes-de-2018>
- [17] A. J. Morales, J. Borondo, J. C. Losada, and R. M. Benito. 2015. Measuring political polarization: Twitter shows the two sides of Venezuela. *Chaos* 25, 3 (2015). <https://doi.org/10.1063/1.4913758> arXiv:arXiv:1505.04095v1
- [18] Jari Saramäki, Mikko Kivelä, Jukka-Pekka Onnela, Kimmo Kaski, and Janos Kertesz. 2007. Generalizations of the clustering coefficient to weighted complex networks. *Physical Review E* 75, 2 (2007), 027105.
- [19] Cass R Sunstein. 2002. The law of group polarization. *Journal of political philosophy* 10, 2 (2002), 175–195.
- [20] Xiaojin Zhu, Zoubin Ghahramani, and John D Lafferty. 2003. Semi-supervised learning using gaussian fields and harmonic functions. In *Proc. of the 20th International conference on Machine learning*. Washington, DC, USA, 912–919.