



UNIVERSIDADE FEDERAL DO ESTADO DO RIO DE JANEIRO
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

ANÁLISE DE MANIFESTAÇÕES POPULARES BRASILEIRAS NO TWITTER
DURANTE A PANDEMIA DO COVID-19

Yuri Luz de Almeida

Orientadores

Dr^a. Adriana Cesário de Faria Alvim

Dr^a. Vânia Maria Félix Dias

RIO DE JANEIRO, RJ - BRASIL
MARÇO DE 2022



ANÁLISE DE MANIFESTAÇÕES POPULARES BRASILEIRAS NO TWITTER
DURANTE A PANDEMIA DO COVID-19

Yuri Luz de Almeida

DISSERTAÇÃO APRESENTADA COMO REQUISITO PARCIAL PARA
OBTENÇÃO DO TÍTULO DE MESTRE PELO PROGRAMA DE PÓS-
GRADUAÇÃO EM INFORMÁTICA DA UNIVERSIDADE FEDERAL DO
ESTADO DO RIO DE JANEIRO (UNIRIO). APROVADA PELA COMISSÃO
EXAMINADORA ABAIXO ASSINADA.

Em conformidade com a Resolução nº 5.257 de 25/03/2020 e a Ordem de Serviço
PROPGPI nº 3 de 2/07/2020, esta ata vai somente por mim assinada, atestando que a
defesa ocorreu com a participação dos componentes abaixo listados.

Aprovada por:

Adriana Cesário de Faria Alvim, D.Sc. - UNIRIO

Vânia Maria Félix Dias, D.Sc. - UNIRIO

Mariano Gomes Pimentel, D.Sc. - UNIRIO

Maria Claudia Silva Boeres, D.Sc. - UFES

RIO DE JANEIRO, RJ – BRASIL
MARÇO DE 2022

Catálogo informatizada pelo(a) autor(a)

- | | |
|------|---|
| L979 | <p>Luz de Almeida, Yuri</p> <p>Análise de Manifestações Populares Brasileiras no Twitter durante a Pandemia do COVID-19 / Yuri Luz de Almeida. -- Rio de Janeiro, 2022. 109 f.</p> <p>Orientadora: Adriana Cesário de Faria Alvim.
Coorientadora: Vânia Maria Félix Dias.
Dissertação (Mestrado) - Universidade Federal do Estado do Rio de Janeiro, Programa de Pós-Graduação em Informática, 2022.</p> <p>1. Teoria dos Grafos. 2. Análise de Redes Sociais. 3. Twitter. 4. Manifestações populares. 5. COVID-19. I. Cesário de Faria Alvim, Adriana, orient. II. Félix Dias, Vânia Maria, coorient. III. Título.</p> |
|------|---|

Eu dedico esta dissertação à minha afilhada Cecília. Meu estudo e meu trabalho são em busca da construção de um Brasil melhor para ela e as crianças de hoje e do futuro.

Agradecimentos

Eu agradeço aos meus pais por todo apoio, incentivo e amor que me fizeram chegar até aqui e me animam a sempre buscar evoluir como pessoa e profissional. Como também, agradeço aos meus irmãos pelo companheirismo que dá graça à vida. E agradeço a toda a minha família pelo respeito e apreço que tenho pela educação.

Sou grato ao Jonathan que me deu suporte e forças para seguir nos estudos, além de me amparar em todos os momentos nessa jornada. Aos meus amigos, em especial Uína e Emanuel, que entenderam minha ausência e, ao mesmo tempo, estiveram comigo em muitos momentos desses últimos anos. Igualmente a Carroll, que além de minha amiga é minha coordenadora no trabalho, por não me deixar desistir dos meus objetivos e me fazer valorizar cada etapa nessa caminhada.

Eu também gostaria de fazer um agradecimento aos meus amigos do Ciência sem Fronteiras que me fizeram enxergar a graduação de outra forma e, dessa maneira, me influenciaram a explorar a vida acadêmica.

Por último, mas não menos importante, eu quero agradecer a todos os professores que me fizeram chegar até aqui, destaco cinco deles em especial: prof.^a Dilza Szwarcman e prof.^a Rosana da Paz, minhas orientadora e coordenadora da graduação em Análise e Desenvolvimento de Sistemas na UEZO, que me incentivaram a seguir estudando e iniciar o mestrado; prof. Rodrigo Santos que foi um grande parceiro nas conquistas alcançadas nesse mestrado e com quem pude aprender muito sobre como fazer pesquisa; e prof.^a Adriana C. F. Alvim e prof.^a Vânia Dias, minhas orientadora e co-orientadora do mestrado, que me ofereçam todo o aprendizado, direcionamento e apoio necessário no curso, nas publicações e nesta minha formação.

ALMEIDA, YURI LUZ DE **Análise das Manifestações Populares Brasileiras no Twitter durante a Pandemia do COVID-19**. UNIRIO, 2022. 110 páginas. Dissertação de Mestrado. Programa de Pós-Graduação em Informática, UNIRIO.

RESUMO

Nas últimas décadas, houve uma expansão significativa da Internet entre a população mundial e, em conjunto com esse crescimento, as redes sociais se consolidaram como tendência e se estabeleceram como plataformas sociais e políticas devido às suas funcionalidades, baixos custos e crescente número de usuários. Diante desse cenário, é muito importante identificar, analisar e trazer conhecimento sobre as manifestações populares que têm ocorrido nas redes sociais para entender essa nova era dos ativismos sociais e políticos em plataformas digitais. Neste trabalho, por meio de análise de dados e conceitos da Teoria de Grafos, junto a observações qualitativas, foram estudadas manifestações populares no Twitter com o objetivo central de trazer conhecimento sobre como as manifestações populares têm sido organizadas na plataforma de rede social Twitter e têm contribuído para as transformações no debate político e em pautas sociais. Ademais, no final de 2019, surgiu uma doença infecciosa causada por um coronavírus (COVID-19) descoberto na China que se espalhou para todos os países causando uma pandemia global. Portanto, esta pesquisa tem como contexto a pandemia da COVID-19, mas deve-se destacar que não foram analisadas necessariamente manifestações diretamente relacionadas à pandemia da COVID-19. Contudo, foram selecionadas para análise manifestações populares no Twitter relacionadas a variados tópicos. Por fim, a partir da metodologia definida para esta pesquisa, foi possível identificar as estruturas de rede e atores das manifestações selecionadas para análise e que estas possuem algumas semelhanças entre si.

Palavras-chave: Teoria dos Grafos, Análise de Redes Sociais, manifestações populares, Twitter, COVID-19.

ABSTRACT

In the last decades, there has been a significant expansion of the Internet among the world population and, in addition with this growth, social networks have consolidated as a trend and established themselves as social and political platforms due to their functionalities, low costs and growing number of users. Given this scenario, it is very important to identify, analyze, and bring knowledge about the popular manifestations that have occurred on social networks in order to understand this new era of social and political activism on digital platforms. In this work, through data analysis and Graph Theory concepts, along with qualitative observations, popular manifestations on Twitter were studied with the central objective of bringing knowledge about how virtual popular manifestations have been organized and contributed to changes in the political debate and in social agendas. Furthermore, in late 2019, an infectious disease caused by a coronavirus (COVID-19) discovered in China emerged and spread to all countries causing a global pandemic. Therefore, this research has the COVID-19 pandemic as its context, but it should be noted that manifestations directly related to the COVID-19 pandemic were not necessarily analyzed. However, popular manifestations on Twitter related to varied topics were selected for analysis. Finally, from the methodology defined for this research, it was possible to identify the network structures and actors of the manifestations selected for analysis and that they have some similarities among themselves.

Keywords: Graph Theory, Social Network Analysis, popular manifestations, Twitter, COVID-19.

Sumário

1	Introdução	17
1.1	Definição do Problema	17
1.2	Objetivos da Pesquisa	18
2	Referencial Teórico	20
2.1	Fundamentação Teórica	20
2.1.1	Teoria dos Grafos	20
2.1.2	Análise de Redes Sociais	21
2.1.3	Topologias de Redes Sociais	24
2.2	Trabalhos Relacionados	26
3	Metodologia	29
3.1	Visão Geral	29
3.2	Coleta de Dados	31
3.2.1	Ferramentas de Coleta de Dados	31
3.2.1.1	Rede Social Twitter	31
3.2.1.2	Banco de Dados Relacional MySQL	32
3.2.1.3	Website GetDayTrends	33

3.2.2	Estratégias de Seleção de Termos e Hashtags para as Coletas de Dados	33
3.2.2.1	Estratégia a Partir dos Assuntos do Momento	33
3.2.2.2	Estratégia a Partir de Capas de Jornais Impressos	34
3.2.2.3	Estratégia a Partir de Portais de Notícias	34
3.2.2.4	Discussão Sobre as Estratégias de Seleção de Termos e Hashtags Experimentadas	35
3.3	Processamento de Dados	36
3.3.1	Mineração dos Dados Coletados	36
3.3.2	Modelagem dos Dados em Grafos	37
3.3.3	Visualização dos Grafos	41
4	Análise dos Dados	43
4.1	Manifestações Sobre o Novo FUNDEB	43
4.1.1	Contexto	43
4.1.2	Coletas de Dados	44
4.1.3	Apresentação do Grafo de Menções G_1	45
4.1.4	Apresentação das Métricas Aplicadas ao Componente Conexo CG_1	46
4.1.5	Análise dos Resultados	51
4.2	Manifestações Sobre a <i>Breque dos Apps</i>	53
4.2.1	Contexto	53
4.2.2	Coletas de Dados	54
4.2.3	Apresentação do Grafo de Menções G_2	55
4.2.4	Apresentação das Métricas Aplicadas ao Componente Conexo CG_2	56

4.2.5	Análise dos Resultados	60
4.3	Manifestações Sobre os Movimentos #Somos70PorCento e #Somos57MILHOES	61
4.3.1	Contexto	61
4.3.2	Coletas de Dados	62
4.3.3	Apresentação do Grafo de Menções G_3	62
4.3.4	Apresentação das Métricas Aplicadas ao Componente Conexo CG_3	64
4.3.5	Análise dos Resultados	69
4.4	Manifestações Sobre o Auxílio Emergencial	71
4.4.1	Contexto	71
4.4.2	Coletas de Dados	72
4.4.3	Apresentação do Grafo de Menções G_4	73
4.4.4	Apresentação das Métricas Aplicadas ao Componente Conexo CG_4	74
4.4.5	Análise dos Resultados	79
4.5	Manifestações Sobre a Realização do ENEM 2020	80
4.5.1	Contexto	80
4.5.2	Coletas de Dados	81
4.5.3	Apresentação do Grafo de Menções G_5	82
4.5.4	Apresentação das Métricas Aplicadas ao Componente Conexo CG_5	82
4.5.5	Análise dos Resultados	86
5	Discussão Sobre os Resultados	90
6	Conclusão	94

- A Estrutura das Tabelas para Armazenamento dos Resultados das Coletas no SGBD MySQL**
- B Dados Complementares da Seção 4.1**
- C Dados Complementares da Seção 4.2**
- D Dados Complementares da Seção 4.3**
- E Dados Complementares da Seção 4.4**
- F Dados Complementares da Seção 4.5**

Lista de Figuras

2.1	Representação visual das três topologias de rede propostas por Baran [1].	25
2.2	Distribuição de grau em redes de modelo aleatório (A), de pequeno-mundo (B) e sem escala (C). Os exemplos dos três modelos de redes foram gerados com 1000 vértices pela ferramenta ELIXIR [2].	26
3.1	Fluxograma da metodologia da pesquisa.	30
3.2	No tweet apresentado nesta captura de tela, o usuário @danielcara gera uma menção ao usuário @cnte_oficial a partir de uma citação direta no texto do tweet.	37
3.3	No tweet apresentado nesta captura de tela, o usuário @carlacgoliveira gera uma menção ao usuário @onyxllorenzoni a partir de uma resposta ao tweet do @onyxllorenzoni.	38
3.4	No tweet apresentado nesta captura de tela, o usuário @ranysbaby gera uma menção ao usuário @tretanotrampo a partir de um retweet do tweet do @tretanotrampo.	38
3.5	No tweet apresentado nesta captura de tela, o usuário @vo_dize gera uma menção ao usuário @jairbolsonaro a partir de um comentário ao tweet do @jairbolsonaro.	39
4.1	Distribuição do tamanho dos componentes no Grafo G_1 (eixos ajustados para escala logarítmica).	46
4.2	Distribuição da centralidade de grau no Componente CG_1	48
4.3	Representação visual do Componente CG_1	50

4.4	Distribuição do tamanho dos componentes no Grafo G_2 (eixos ajustados para escala logarítmica).	55
4.5	Distribuição da centralidade de grau no Componente CG_2	56
4.6	Representação visual do Componente CG_2	59
4.7	Distribuição do tamanho dos componentes no Grafo G_3 (eixos ajustados para escala logarítmica).	63
4.8	Distribuição da centralidade de grau no Componente CG_3	64
4.9	Representação visual do Componente CG_3	67
4.10	Distribuição do tamanho dos componentes no Grafo G_4 (eixos ajustados para escala logarítmica).	74
4.11	Distribuição da centralidade de grau no Componente CG_4	76
4.12	Representação visual do Componente CG_4	77
4.13	Distribuição do tamanho dos componentes no Grafo G_5 (eixos ajustados para escala logarítmica).	83
4.14	Distribuição da centralidade de grau no Componente CG_5	83
4.15	Representação visual do Componente CG_5	86

Lista de Tabelas

4.1	Descrição dos dados das coletas de tweets relacionadas ao FUNDEB. . .	45
4.2	Vértices com os dez maiores valores de grau de entrada no Componente CG_1	47
4.3	Vértices com os dez maiores valores de grau de saída no Componente CG_1 . 47	
4.4	Vértices com os dez maiores valores de centralidade de autovetor no Componente CG_1	48
4.5	Vértices com os dez maiores valores de centralidade de proximidade no Componente CG_1	49
4.6	Vértices com os dez maiores valores de centralidade de intermediação no Componente CG_1	50
4.7	Resumo das métricas de grafo para o Grafo G_1 e o Componente CG_1 . . .	51
4.8	Descrição das coletas de tweets relacionadas à greve dos entregadores dos aplicativos.	55
4.9	Vértices com os dez maiores valores de grau de entrada no Componente CG_2	57
4.10	Vértices com os dez maiores valores de grau de saída no Componente CG_2 . 57	
4.11	Vértices com os dez maiores valores de centralidade de autovetor no Componente CG_2	57
4.12	Vértices com os dez maiores valores de centralidade de proximidade no Componente CG_2	58

4.13	Vértices com os dez maiores valores de centralidade de intermediação no Componente CG_2	59
4.14	Resumo das métricas de grafo para o Grafo G_2 e o Componente CG_2 . . .	60
4.15	Descrição das coletas de tweets relacionados aos movimentos #Somos70PorCento e #Somos57MILHOES.	63
4.16	Vértices com os dez maiores valores de grau de entrada no Componente CG_3	65
4.17	Vértices com os dez maiores valores de grau de saída no Componente CG_3 . . .	65
4.18	Vértices com os dez maiores valores de centralidade de autovetor no Componente CG_3	66
4.19	Vértices com os dez maiores valores de centralidade de proximidade no Componente CG_3	66
4.20	Vértices com os dez maiores valores de centralidade de intermediação no Componente CG_3	68
4.21	Resumo das métricas de grafo para o Grafo G_3 e o Componente CG_3 . . .	68
4.22	Descrição das coletas de tweets relacionados às manifestações sobre o auxílio emergencial.	73
4.23	Vértices com os dez maiores valores de grau de entrada no Componente CG_4	75
4.24	Vértices com os dez maiores valores de grau de saída no Componente CG_4 . . .	75
4.25	Vértices com os dez maiores valores de centralidade de autovetor no Componente CG_4	75
4.26	Vértices com os dez maiores valores de centralidade de proximidade no Componente CG_4	76
4.27	Vértices com os dez maiores valores de centralidade de intermediação no Componente CG_4	78
4.28	Resumo das métricas de grafo para o Grafo G_4 e o Componente CG_4 . . .	78
4.29	Descrição das coletas de tweets relacionados às manifestações sobre o ENEM 2020.	82

4.30	Vértices com os dez maiores valores de grau de entrada no Componente CG_5	84
4.31	Vértices com os dez maiores valores de grau de saída no Componente CG_5	84
4.32	Vértices com os dez maiores valores de centralidade de autovetor no Componente CG_5	85
4.33	Vértices com os dez maiores valores de centralidade de proximidade no Componente CG_5	85
4.34	Vértices com os dez maiores valores de centralidade de intermediação no Componente CG_5	87
4.35	Resumo das métricas de grafo para o Grafo G_5 e para o Componente CG_5	87
A.1	Estrutura das tabelas de tweets coletados (separadas por assunto selecionado)	
A.2	Estrutura da tabela de tweets referenciados (retweetados, citados ou respondidos)	
A.3	Estrutura da tabela de usuários	
B.1	Tweets coletados sobre o FUNDEB com mais interações	
C.1	Tweets coletados sobre o Breque dos Apps com mais interações	
D.1	Tweets coletados sobre os movimentos #Somos70porcento e #Somos57MILHOES com mais interações	
E.1	Tweets coletados sobre o auxílio emergencial com mais interações	
F.1	Tweets Coletados sobre o ENEM 2020 com mais interações	

Lista de Siglas

CLT	Consolidação das Leis do Trabalho
CNTE	Confederação Nacional dos Trabalhadores em Educação
ENEM	Exame Nacional do Ensino Médio
FIES	Fundo de Financiamento Estudantil
FUNDEB	Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
MEC	Ministério da Educação
PEC	Proposta de Emenda à Constituição
PROUNI	Programa Universidade para Todos
PSOL	Partido Socialismo e Liberdade
SINTESE	Sindicato dos Trabalhadores em Educação Básica da Rede Oficial do Estado de Sergipe
SISU	Sistema de Seleção Unificada
STF	Superior Tribunal Federal
UNDIME	União Nacional dos Dirigentes Municipais de Educação

1. Introdução

1.1 Definição do Problema

Nas últimas décadas, houve uma expansão da Internet entre a população mundial e, conforme demonstrado em pesquisa feita pela Pew Research Center [3], em conjunto com esse crescimento e avanço da Internet, novas tecnologias, como as redes sociais online, emergiram e ampliaram sua relevância social. As redes sociais se consolidaram e se estabeleceram como plataformas sociais ao possibilitar a conexão facilitada entre pessoas, permitindo a troca de ideias e compartilhamento de informações, entre outras funcionalidades sociais. Da mesma maneira, as plataformas de redes sociais têm se tornado um espaço importante para o conhecimento político [4]. Dessa forma, o impacto das redes sociais na sociedade, bem como a capacidade de influenciar a visão política por meio das redes sociais, tem sido inegável.

No Brasil, pesquisa realizada pela plataforma DataReportal em 2019 [5] mostra que 70% da população (cerca de 149,1 milhões de pessoas) está conectada à Internet e 66% dos brasileiros (cerca de 140 milhões de pessoas) são usuários ativos de alguma rede social. Também, como consequência a esse alcance, políticos, partidos, organizações e movimentos sociais estão intensificando o uso das redes sociais como canal de comunicação para anunciar ações, divulgar realizações, reivindicar pautas, criar debate, demonstrar apoio, como também, realizar manifestações.

Cohen [6] afirma que as redes sociais não são aleatórias, sua formação e desenvolvimento são ditados por uma combinação de muitos processos e influências diferentes como processos naturais, condições humanas, condições econômicas, seleção natural, dentre outros processos. Dessa forma, faz-se necessário o estudo e análise das redes sociais para trazer mais entendimento sobre as manifestações populares online a partir de identificação, análise e produção de conhecimento sobre elas. Este conhecimento é de suma importância para assimilar essa nova era de ativismos sociais e atividades políticas

em plataformas digitais. Portanto, propõe-se com esta pesquisa analisar manifestações populares brasileiras em plataforma de rede social e trazer informações sobre como elas se estruturam em rede e transmitem suas mensagens nesse ambiente virtual. No caso deste trabalho, serão analisadas manifestações populares em uma plataforma de rede social no formato de microblog, o Twitter.

1.2 Objetivos da Pesquisa

As manifestações em redes sociais podem ser de diversas temáticas, entre elas, político, partidário e social. As manifestações populares selecionadas para análise nesta pesquisa são relacionadas a essas variadas temáticas. Portanto, buscou-se observar os assuntos em destaque na plataforma de rede social Twitter, independente de suas temáticas, para trazer conhecimento sobre as redes de manifestações populares geradas.

É importante destacar que, ao longo deste trabalho, surgiu uma doença infecciosa causada por um coronavírus (COVID-19) descoberto na China no final de 2019. Essa doença logo se espalhou para todos os países causando uma pandemia global. A pandemia da COVID-19 tem produzido, além da crise sanitária, impactos sociais, prejuízos econômicos e disseminação de desinformação. Portanto, para esta pesquisa, o contexto da pandemia da COVID-19 deve ser considerado. Dentro desse contexto, foi evidente a intensificação do uso das redes sociais para diversos fins devido ao isolamento social necessário e estabelecido pelas autoridades públicas, como também o surgimento e destaque de pautas que não haveriam na ausência de uma situação sanitária igual à da pandemia da COVID-19. Portanto, a pandemia da COVID-19 será observada como contexto neste estudo, mas deve-se destacar que não foram analisadas necessariamente manifestações diretamente relacionadas com a pandemia da COVID-19.

Por meio de análise de dados e conceitos da Teoria de Grafos, em conjunto com observação qualitativa das manifestações populares no Twitter, este trabalho possui três objetivos principais, descritos a seguir:

1. Compreender os fluxos de informação em manifestações populares realizadas no Twitter por meio da Teoria de Grafos, avaliando como os participantes (atores) se estruturam em rede para difundir movimentos, causas, ideias, posições, opiniões, etc.;
2. Identificar quais grupos estão presentes nas manifestações;
3. Identificar quais atores possuem relevância nas manifestações, provocando, por

meio do alcance e repercussão de suas publicações, importância na formação da opinião pública.

Procura-se com esta pesquisa ajudar a entender como as manifestações populares virtuais têm contribuído para as transformações no debate político ao longo do ano de 2020 — primeiro ano da pandemia da COVID-19. Por fim, pretende-se produzir informações sobre as militâncias políticas em ambiente virtual no Brasil trazendo luz sobre suas dinâmicas, apontando seus principais atores e fornecendo contribuições científicas a pesquisadores que investigam política nas redes sociais. Desta maneira, a análise de variadas manifestações durante um momento significativo e de grande impacto para a população brasileira, como no caso da pandemia da COVID-19, buscando identificar os atores políticos mais importantes e analisando a distribuição das mensagens e estrutura das redes, é a principal contribuição deste trabalho.

Esta dissertação está organizada da seguinte forma: a Seção 2 traz o referencial teórico para as análises; a Seção 3 relata a metodologia que conduzirá a produção desse trabalho; a Seção 4 apresenta os dados coletados das manifestações selecionadas para estudo e, em seguida, as análises sobre cada resultado obtido; a Seção 5 discute e relaciona os resultados das análises; por fim, a Seção 6 conclui este trabalho.

2. Referencial Teórico

Neste capítulo, serão descritas as definições e notações básicas utilizadas como fundamentação teórica neste trabalho. Na Seção 2.1, apresentam-se conceitos da área de *Teoria dos Grafos*, conceitos e métricas da área de *Análise de Redes Sociais*, e, de forma resumida, as *topologias e modelos de redes* usados como base para as análises desta pesquisa. Na Seção 2.2 são apresentados trabalhos relacionados à presente pesquisa.

2.1 Fundamentação Teórica

2.1.1 Teoria dos Grafos

Um *grafo* $G = (V, E)$ é um conjunto finito não vazio V e um conjunto E finito de pares não ordenados de elementos distintos de V . Os elementos de V são os *vértices* e os de E são as *arestas* de G . Por definição, o *tamanho* de um grafo é $|V| + |E|$. Cada aresta $e \in E$ será denotada pelo par de vértices $e = (u, v)$ que a forma. Os vértices u e v são as *extremidades* de e , denominados *adjacentes*. Define-se *grau* de um vértice $v \in V$, denotado por $\text{grau}(v)$, como sendo o número de vértices adjacentes a v .

Um *grafo direcionado*, ou *digrafo*, é um conjunto finito não vazio V e um conjunto E finito de pares ordenados de vértices distintos. Portanto, num digrafo cada aresta possui uma única direção de u para v . Diz-se que u é o vértice de saída e v é o vértice de entrada. Seja $G = (V, E)$ um digrafo e um vértice $v \in V$, o *grau de entrada* de $v \in V$ corresponde ao número de arestas que entram em v , e o *grau de saída* de v é o número de arestas que saem de v . Usualmente, denota-se por $n = |V|$ e por $m = |E|$.

Um *grafo completo* é um grafo em que todos os pares de vértices são adjacentes. A *densidade* d de um grafo não direcionado corresponde a razão entre a quantidade m de arestas do grafo e a quantidade n de arestas de um grafo completo, isto é, $d = 2m/n(n-1)$.

1). A densidade para grafos direcionados, como os utilizados nesta pesquisa, é calculada por $d = m/n(n-1)$.

Um *caminho* em um grafo é uma sequência finita de vértices $v_1, \dots, v_k$ tal que $(v_i, v_{i+1}) \in E$ para todo i no intervalo $1, \dots, k-1$. O caminho mais curto possível entre dois vértices é chamado *caminho geodésico*. A *distância* entre dois vértices de um grafo é indicada pelo caminho geodésico entre esses vértices.

Um *subgrafo* $G_2 = (V_2, E_2)$ de um grafo $G_1 = (V_1, E_1)$ é um grafo tal que $V_2 \subseteq V_1$ e $E_2 \subseteq E_1$. Usa-se a notação $G_2 \subseteq G_1$. Além disso, se G_2 possuir toda aresta de G_1 tal que ambos u e v estejam em V_2 , então G_2 é o *subgrafo induzido* pelo subconjunto de vértices V_2 . Diz-se que V_2 induz G_2 .

Um grafo G é denominado *conexo* se há um caminho entre todo par de vértices; se o grafo G é direcionado, ele é denominado de *fortemente conexo* para essa condição. Caso contrário, G é dito *desconexo* se é um grafo não-direcionado e dito *não-fortemente conexo* se é um grafo direcionado. Para grafos não-direcionados, um *componente conexo* é um subgrafo conexo maximal do grafo G em relação a esta propriedade; naturalmente, um grafo *desconexo* possui no mínimo dois componentes conexos. Para grafos direcionados, um *componente fortemente conexo* é um subgrafo fortemente conexo maximal do grafo G em relação a esta propriedade; diz-se, de forma trivial, um grafo *não-fortemente conexo* aquele que possui mais de um componente.

Um *grafo ponderado* $G = (V, E, w)$ (nas arestas) é um grafo, com uma função w , que associa um número real não negativo $w(e)$ a cada aresta $e \in E$, denotado por *peso da aresta* e .

Ao longo deste texto, e doravante, a palavra grafo terá a conotação de grafo direcionado. O mesmo é válido para as métricas que serão sempre consideradas métricas de grafos direcionados. Com ressalva ao termo componente conexo, que é usualmente aplicado para grafos não-direcionados, mas, por simplicidade, este será usado para definir componentes não-fortemente conexos. Daqui por diante, essas informações serão omitidas ao longo do texto.

2.1.2 Análise de Redes Sociais

Uma *rede social* pode ser definida como um conjunto de *atores*, como pessoas, grupos e organizações, com alguns relacionamentos ou interações entre eles. As redes sociais são geralmente modeladas por grafos, onde os vértices representam os atores e as arestas representam as relações estabelecidas entre eles.

A Análise de Redes Sociais pode ser descrita como um estudo das relações humanas com o auxílio da Teoria dos Grafos para obter informações sobre redes sociais em um determinado contexto [7]. Além das técnicas estatísticas usuais para análise de dados, a Análise de Redes Sociais traz outras métricas para investigação das redes sociais que ajudam a compreender as dependências entre atores sociais nos dados, caracterizando seus comportamentos e efeitos em uma rede.

Nos próximos parágrafos serão apresentadas as métricas de Análise de Redes Sociais aplicadas nesta pesquisa. Para todas as métricas, considere um grafo $G = (V, E)$, como definido na Seção 2.1.1, usado para modelar uma rede social. São consideradas *métricas locais* aquelas que não há necessidade de pesquisar todo o gráfico para calcular seu valor.

A *centralidade de grau* de um vértice é o grau desse vértice. É uma métrica local porque depende apenas do número de arestas que um vértice possui. Para grafos direcionados, a centralidade de grau é calculada como seu grau. A centralidade de grau do vértice v é dada por:

$$CG(v) = g(v)$$

Onde $g(v)$ é o grau do vértice v .

Ao avaliar a influência de um vértice para um grafo, é possível que a centralidade de grau sozinha não apresente corretamente a importância dos vértices presentes no grafo em estudo. Portanto, outras medidas de centralidade em grafos são consideradas para ampliar as informações para análises.

A *centralidade de autovetor* estima quais são os vértices centrais de um grafo a partir de adjacências a outros vértices centrais do grafo. A centralidade de autovetor de um vértice v é dada por:

$$CE(v) = \frac{1}{\lambda} \sum_{u \in N(v)} CE(u) = \frac{1}{\lambda} \sum_{u \in V} a_{uv} C(u)$$

Onde $\lambda \neq 0$ é uma constante, $N(v)$ é o conjunto de vértices vizinhos a v e $a_{uv} = 1$ se o vértice v é adjacente ao vértice u , caso contrário, $a_{uv} = 0$.

A *centralidade de proximidade* é relacionada à distância total de um vértice aos demais vértices do grafo. A centralidade de proximidade comumente é dada por [8,9]:

$$CP(v) = \frac{n-1}{\sum_{v \neq u} d(u, v)}$$

Onde $v \in V$, $n = |V|$ e $d(u, v)$ é a distância entre os vértices u e v .

A *centralidade de intermediação* calcula quantos caminhos geodésicos entre todos os pares de vértices do grafo passam através de um determinado vértice. A centralidade de intermediação do vértice v é dada por:

$$CI(v) = \sum_{u, x \neq v} \frac{g_{uvx}}{g_{ux}}$$

Onde g_{ux} é o número de caminhos geodésicos entre u e x , e g_{uvx} é o número desses caminhos geodésicos que passam por v .

O *coeficiente de agrupamento* é uma propriedade dos grafos que indica o quanto os vértices estão conectados. O coeficiente de agrupamento de um determinado vértice v é calculado a partir da razão entre as arestas existentes e as arestas possíveis entre os vértices adjacentes desse vértice. Para grafos direcionados, o *coeficiente de agrupamento local* do vértice v pode ser dado como:

$$CA(v) = \frac{|\{e_{ux} : u, x \in N(v), e_{ux} \in E\}|}{k_v(k_v - 1)}$$

Onde $N(v)$ é o conjunto de vértices vizinhos a v e k_v corresponde ao número de vértices vizinhos ao vértice v .

O *coeficiente de agrupamento global*, que mede o quanto os vértices se agrupam em toda a rede, é denotado por:

$$CA = \frac{1}{n} \sum_{v \in V} CA(v)$$

Onde n é o número de vértices na rede e CA é a média, calculada sobre os vértices de um grafo G , dos coeficientes de agrupamento local $CA(v)$. Por definição, $0 \leq CA(v) \leq 1$ e $0 \leq CA \leq 1$.

Comunidades em um grafo são subconjuntos de vértices em que a densidade entre os vértices é maior que a densidade entre vértices de outros subconjuntos [10]. Em uma rede social, as comunidades podem representar grupos sociais que possuem características ou interesses em comum. A *detecção de comunidades* é um problema que se refere à divisão do grafo em subconjuntos de vértices mais densamente conectados. Como linha de pesquisa em redes, a detecção de comunidades por algoritmos de busca, por meio das arestas do grafo, busca identificar a melhor representação da relação entre os atores do grafo.

A *modularidade* Q de um grafo representa a propriedade do grafo possuir comunidades de vértices altamente conectados entre si e pode ser dada por:

$$Q = \frac{1}{2m} \sum_{uv} \left[A_{uv} - \frac{k_u k_v}{2m} \right] \delta_{uv} (c_u c_v)$$

Onde A_{uv} representa o peso da aresta entre u e v (para $u, v \in E$), $k_u = \sum_v A_{uv}$ é a soma dos pesos das arestas conectadas ao vértice u , c_u é a comunidade à qual o vértice u é atribuído, $m = \frac{1}{2} \sum_{uv} A_{uv}$ e, se $u = v$, a função (δ_{uv}) é igual 1, caso contrário ela é igual 0. Otimizar a modularidade em teoria resulta no melhor agrupamento possível dos vértices de uma determinada rede. Porém, para grafos grandes é inviável computacionalmente passar por todas as conexões possíveis dos vértices em comunidades, portanto, com esta finalidade, são aplicados algoritmos heurísticos para *otimização da modularidade*.

Diz-se que uma relação E , entre os indivíduos de V , é explícita quando $(u, v) \in E$ e pelo menos um dos extremos de (u, v) estabelece explicitamente uma conexão com o outro; por exemplo, o usuário u menciona deliberadamente o usuário v em uma postagem em uma rede social [11]. Relações explícitas são fornecidas por links explícitos existentes entre usuários; por exemplo, o usuário u envia uma mensagem privada para o usuário v pelos correios de uma rede social. Diz-se que E é implícita se $(u, v) \in E$ pode ser apontada a partir de um conjunto de outras conexões, como efeito colateral de ações tomadas por u e v , *e.g.*, visitas de perfil.

2.1.3 Topologias de Redes Sociais

Baran [1] sugere três topologias de redes: *a centralizada, a descentralizada e a distribuída*. A diferenciação entre essas topologias de redes pode ser observada na Figura 2.1. A rede centralizada (A) é apresentada de forma estrelar em que um vértice está presente em todas as conexões estabelecidas com os demais vértices. A rede descentralizada (B) estabelece pequenos grupos de vértices conectados que, por sua vez, estão interconectados aleatoriamente entre si. Por último, a rede distribuída (C) é organizada a partir de vários vértices que estabelecem conexões de tal maneira que equilibra a quantidade de vértices e arestas.

As três topologias de redes apresentadas no parágrafo anterior não focam na complexidade das redes. Nos próximos parágrafos, serão apresentados brevemente os três modelos de redes complexas mais comuns da chamada “Ciência das Redes” [12]. Esses modelos descrevem o processo de criação de redes complexas.

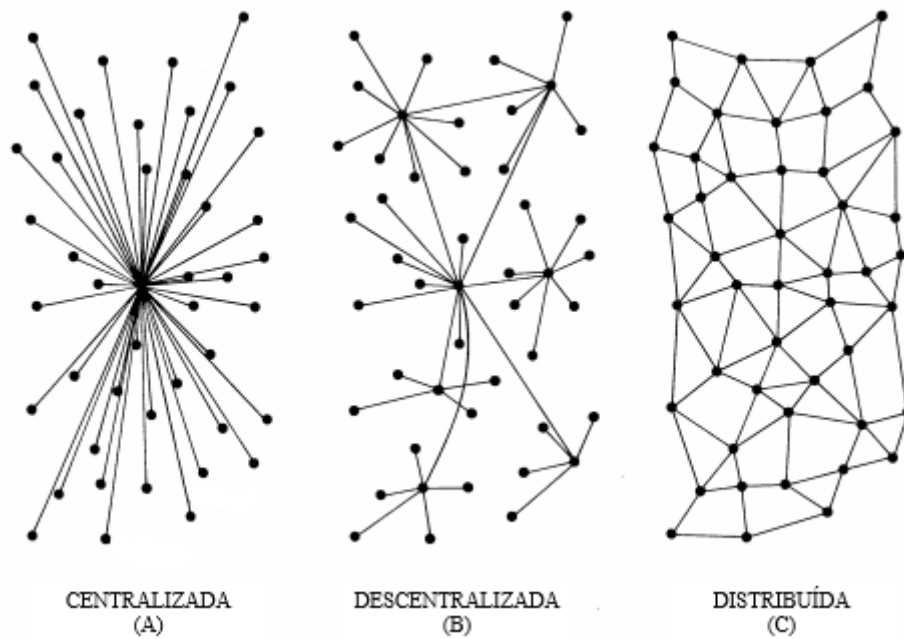


Figura 2.1: Representação visual das três topologias de rede propostas por Baran [1].

As *redes aleatórias* (em inglês, random networks) são representadas por grafos em que uma aresta depende de uma distribuição binomial ou Poisson dos graus dos vértices de probabilidade para existir. Porém, é o modelo de rede mais simples. Vale ressaltar que ainda há controvérsias sobre a medida em que os modelos aleatórios representam redes reais, talvez a maior delas seja que grafos aleatórios possuem uma distribuição de graus do tipo Poisson, sem grandes discrepâncias de valor de grau dos vértices, e, por sua vez, redes reais normalmente apresentam distribuição de cauda longa (*right-skewed distribution*), nesse caso, a maioria dos vértices possui valor de grau baixo e poucos vértices possuem valor de grau alto. Além disso, na maioria das redes reais os vértices se agrupam em comunidades, comportamento que grafos aleatórios não apresentam.

Sobre a formação de redes reais, Watts e Strogatz afirmam que:

"Embora grafos regulares e grafos aleatórios sejam idealizações úteis, muitas redes reais ficam em algum lugar entre os extremos da ordem e da aleatoriedade." [12]

As *redes de pequeno-mundo* (em inglês, small-world networks), apresentadas por Watts e Strogatz [13], trazem um modelo simples que pode se aproximar do meio-termo: uma rede real onde as conexões originais são substituídas por outras aleatórias com probabilidade $0 \leq \phi \leq 1$. Nesse modelo de redes, há caminhos curtos entre quaisquer dois vértices, semelhante às redes aleatórias, mas com a tendência de se formar comunidades,

como em redes reais. Entre as vantagens do modelo pequeno-mundo em relação a outros modelos de rede, podem ser citadas: a alta transitividade compatível com redes reais, propriedade medida a partir do coeficiente de agrupamento da rede; o efeito de pequeno-mundo (*small-world effect*) que faz com que o comprimento de caminho entre qualquer par de vértices na rede seja pequeno.

Em qualquer rede real, alguns vértices são mais altamente conectados do que outros [12]. As *redes sem escala*, descritas pelo modelo de Barabási-Albert, utilizam um método em que quanto maior o grau do vértice maior a probabilidade deste vértice receber novas arestas. Os matemáticos Barabási, Albert e Jeong apelidaram essas redes sem escala (em inglês, *scale-free networks*), por analogia com situações onde aparecem leis de potência e nenhuma escala característica única pode ser atribuída. A razão disso é que uma lei de potência é a única distribuição que tem a mesma forma independentemente da escala em que se olhe [14].

A diferença entre as distribuições de grau dos três modelos de redes complexas apresentados é ilustrada por meio da Figura 2.2.

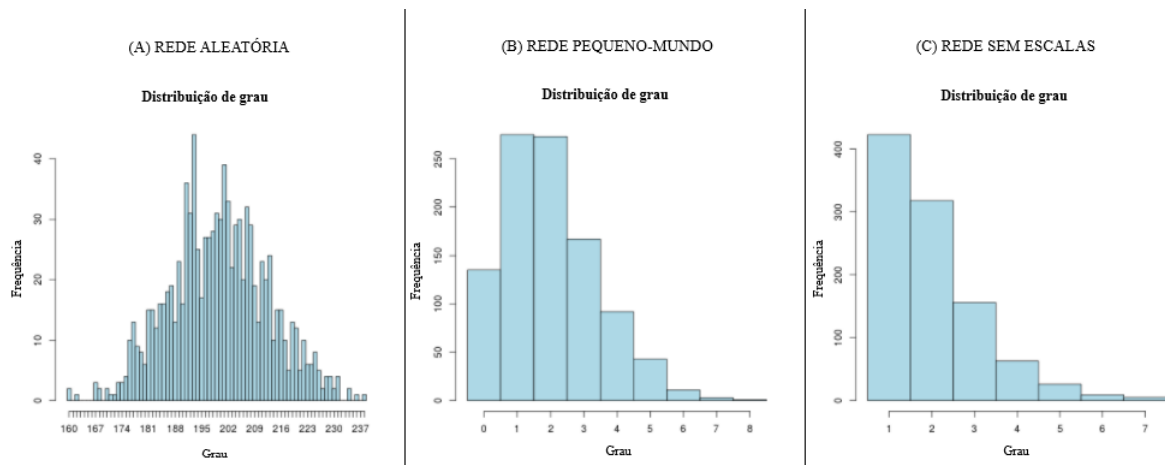


Figura 2.2: Distribuição de grau em redes de modelo aleatório (A), de pequeno-mundo (B) e sem escala (C). Os exemplos dos três modelos de redes foram gerados com 1000 vértices pela ferramenta ELIXIR [2].

2.2 Trabalhos Relacionados

Nesta seção, serão apresentados alguns estudos que analisaram manifestações populares e demais ações coletivas no Twitter ou em outras redes sociais, e esses estudos serão comparados, posteriormente, com os objetivos desta pesquisa na Seção 5 — Discussão sobre os Resultados.

Sandra González-Bailón e Ning Wang [15] analisaram a estrutura de uma rede de comunicação no Twitter sobre uma campanha internacional — *United for Global Change* — lançada em maio de 2012. Para avaliar a rede, as trocas de mensagens foram observadas, como também examinada a coesão da rede com métodos de detecção de comunidade e identificação dos usuários que intermediaram a difusão de informações. As descobertas feitas no estudo de González-Bailón e Wang fornecem evidências de centralização nas redes, de fragmentação na dinâmica da comunicação online, e de desigualdade e sub-exploração em oportunidades de mediação. Além disso, as análises mostraram que apenas uma minoria de usuários facilita a disseminação de informação na comunicação de manifestações, prejudicando a conectividade global. González-Bailón e Wang afirmam que um crescente número de pesquisas sugere que as redes online não são estruturas descentralizadas com as quais os movimentos sociais são frequentemente comparados. Contudo, as autoras argumentam que essa discussão teórica deve avançar com base em uma longa tradição de pesquisa em rede e descobertas empíricas.

Jordana George e Dorothy E. Leidner [16] se referem ao ativismo digital como o uso de tecnologias digitais por um indivíduo, grupo ou organização para promover mudanças políticas, como também a oportunidades de ação política direta e de maior impacto com menos recursos. Em estudo realizado, as autoras buscaram explorar e construir uma teoria sobre o uso e o impacto do ativismo digital para participação política. George e Leidner defendem que o ativismo digital permitiu que órgãos políticos ultrapassassem fronteiras e ganhassem novos adeptos em tempo recorde e, de tal maneira, os sistemas de informação se estabeleceram como parte do cenário político. Como resultado do estudo, é concluído que, conforme as tecnologias se espalham e mais pessoas as usam, pode-se esperar mais incidentes relacionados ao ativismo digital. Por fim, as autoras afirmam que o nível de ativismo digital que vemos hoje é provavelmente um precursor de grandes ações empreendidas pelos cidadãos do futuro.

S.J. Park et al. [17] exploraram comparativamente a estrutura de rede, padrão de interação e distribuição geográfica de usuários envolvidos em redes de comunicação do movimento *Occupy Wall Street*, no Twitter e no YouTube. Os resultados do estudo mostram que os usuários do Twitter geraram uma rede *hub-and-spoke* — o que indica que as informações eram mais provavelmente organizadas por vários usuários centrais na rede. Os usuários centrais não apenas desempenharam papéis de formadores de opinião na disseminação de informações, mas também criaram pontes para outras pequenas comunidades na rede. Portanto, pode haver limitações em compartilhar todas as mensagens sem alguns atores-chave que ligam diferentes comunidades e é provável que as informações sejam controladas por tais formadores de opinião. Os resultados destacam os diferentes papéis

das redes sociais na difusão de informações políticas. Por meio de técnicas analíticas de mineração de dados e análise de redes, é demonstrado que Twitter e YouTube têm papéis distintos como plataformas de rede social na organização da ação coletiva. Os resultados também têm implicações teóricas para a compreensão da ecologia da web, identificando a estrutura de rede distinta e o padrão de interação dos usuários em duas plataformas de mídia social relevantes do Twitter e YouTube [18].

Agathangelou et al. [19] examinaram a conectividade e interações dentro e entre redes políticas no Twitter na campanha para três votações diferentes realizadas na Grécia em 2015: a eleição parlamentar de 25 de janeiro, o referendo de 5 de julho, e a eleição antecipada de 20 de setembro. No estudo, foram exploradas a dinâmica interna de partidos de todo o espectro político, tão como a estrutura da rede e o sentimento nas relações entre usuários. Além disso, a ação política durante as campanhas foi comparada com a ação política rotineira. Os resultados do estudo mostraram que as discussões políticas não ocorreram apenas em câmaras de eco como outros estudos apontam, os atores políticos nas redes consideram a turbulência e a volatilidade como oportunidade política para ganhar visibilidade. O estudo também revelou como as divisões emergentes relacionadas à crise, nesse caso, a grande crise grega, afetaram a estrutura das redes e facilitaram a convergência entre as arestas das redes políticas. Por fim, foi identificado que as redes políticas passaram de orientadas a partidos para orientadas a questões.

Esses trabalhos sustentam a ideia de que as redes sociais têm sido fortemente exploradas como plataformas digitais para manifestações populares. Em geral, os trabalhos citados consideraram manifestações específicas ou campanhas eleitorais, observaram a disseminação de informações nas redes e analisaram as estruturas de rede formadas. Em contrapartida, neste estudo, serão analisadas manifestações diversas dentro de um contexto específico — a pandemia global da COVID-19 — além de não haver campanhas eleitorais entre as manifestações selecionadas. Apesar do ano de 2020, ano em que os dados deste estudo foram coletados, ser um ano de eleições municipais no Brasil, faz-se essa escolha, de não selecionar nenhuma campanha eleitoral para análise, porque se entende que, com o advento das plataformas de redes sociais e a expansão de seu uso pela população, as campanhas eleitorais têm deixado de ocorrer apenas durante o período eleitoral e tem se tornado contínuas ao longo dos quatro anos que separam os pleitos. Portanto, identificar os atores políticos em evidência nas manifestações populares nas redes sociais é de suma importância para analisar as corridas eleitorais dos pleitos seguintes em futuras pesquisas.

3. Metodologia

Este capítulo apresenta a metodologia utilizada nesta dissertação e está organizado em três seções. A primeira seção, Seção 3.1, apresenta uma visão geral sobre a metodologia utilizada. A segunda seção, Seção 3.2, descreve o processo de coleta de dados a partir das ferramentas e estratégias abordadas na pesquisa. A terceira seção, Seção 3.3, relata os procedimentos executados para mineração dos dados, modelagem em grafos e visualização dos grafos obtidos, além de apresentar as métricas selecionadas para a análise dos dados que será feita no Capítulo 4.

3.1 Visão Geral

Segundo as classificações apresentadas por Silva e Menezes [20], este estudo pode ser enquadrado como uma pesquisa quali-quantitativa sobre as redes sociais. A pesquisa é, por um lado, quantitativa, pois busca traduzir as informações em números e analisá-las por meio da Teoria dos Grafos; e, por outro lado, a pesquisa é qualitativa, pois, a partir da observação do objeto de estudo, a plataforma de rede social Twitter, e de fatos ocorridos na sociedade, faz a relação entre as informações obtidas de modo a buscar o entendimento de fenômenos complexos específicos de natureza social e cultural.

Fachin [21] afirma que o projeto de pesquisa é uma sequência de etapas estabelecidas pelo pesquisador, em que se direciona a metodologia a ser aplicada no desenvolvimento da pesquisa. Dessa forma, para guiar o desenvolvimento da dissertação, foram definidas as seguintes etapas da metodologia:

1. Etapa de coleta de dados sistemática: é dividida em dois passos. O primeiro passo desta etapa é a elaboração e experimentação de estratégias para a coleta de dados para a pesquisa. Esse passo visa obter a melhor maneira de acompanhar os acontecimentos sociais durante a pandemia da COVID-19, e suas repercussões no Twitter,

e identificar os termos e “hashtags” (em português, palavras-chave) que devem ser selecionados para as coletas de dados. As estratégias de coleta produzidas nesse passo são usadas como base para o segundo passo em que são coletadas, diariamente, publicações do Twitter relacionadas aos termos e hashtags selecionados. Essas coletas são agrupadas por assuntos e filtradas a partir das observações obtidas durante o processo executado no passo anterior. O objetivo desta primeira etapa é obter um procedimento de acompanhamento dos assuntos levantados durante a pandemia da COVID-19 e da evolução das manifestações online sobre esses assuntos. Em seguida, a partir das observações feitas, efetuar a coleta de dados no Twitter.

2. Etapa de processamento de dados: após finalizar o período de coleta de dados sistemática, são selecionados os principais assuntos para análise. Os dados obtidos relacionados aos assuntos selecionados são processados por meio de mineração de dados e modelados em grafos. Em conjunto, são extraídas métricas e geradas visualizações gráficas para análises posteriores.
3. Análise dos dados: os grafos obtidos após o processamento de dados são analisados individualmente a partir de suas métricas e visualizações para adquirir informações sobre seus vértices, topologias e modelos de redes visando trazer conhecimento sobre as manifestações online relacionadas aos assuntos escolhidos. Por fim, as análises feitas são avaliadas em conjunto de modo a observar semelhanças e particularidades entre os grafos analisados.

A seguir, a Figura 3.1 mostra um fluxograma ilustrando o processo de metodologia resumido acima.

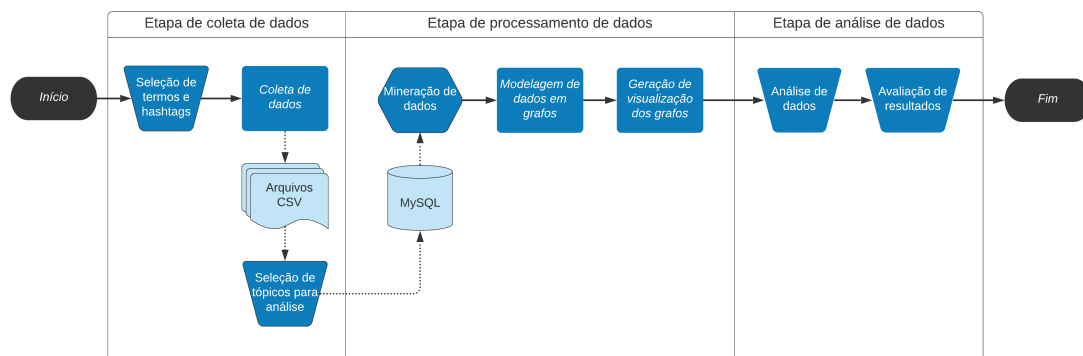


Figura 3.1: Fluxograma da metodologia da pesquisa.

3.2 Coleta de Dados

Nesta seção, organizada em duas subseções, será apresentado o processo de coleta de dados executado nesta pesquisa. A Seção 3.2.1 descreve as ferramentas selecionadas de coleta e armazenamento de dados, conforme a ordem em que foram utilizadas na pesquisa. A Seção 3.2.2 relata as estratégias de seleção de termos e hashtags experimentadas nesta pesquisa para a coleta de dados.

3.2.1 Ferramentas de Coleta de Dados

3.2.1.1 Rede Social Twitter

De acordo com o Portal DataReportal [5], as três plataformas de rede social mais usadas no Brasil são, em ordem de popularidade, YouTube, Facebook e WhatsApp. Entretanto, apesar de não estar nem entre as cinco plataformas de rede social preferidas dos brasileiros, o Twitter tem se mostrado muito importante para o cenário político em geral [22–24]: sejam campanhas eleitorais, propagandas governamentais ou manifestações sociais. Por esta razão, o Twitter foi escolhido para ser analisado na atual pesquisa.

Como plataforma de rede social, o Twitter permite que cada usuário publique pequenas mensagens de texto com até 280 caracteres, também conhecidas como “tweets”, contendo imagens, vídeos, endereços de rede (URL) ou palavras-chave livremente escolhidas sobre assuntos que são chamadas de “hashtags”. As interações que um usuário pode fazer com um tweet são: compartilhar o tweet, ação popularmente conhecida como “retweetar”; comentar um tweet, que é um compartilhamento de tweet com comentário; responder um tweet; e favoritar um tweet, mais conhecido como curtir. Todas essas interações citadas afetam a descoberta e o alcance de uma publicação no Twitter. Além disso, no Twitter, os usuários podem mencionar diretamente uns aos outros marcando outro ID de perfil de usuário em um tweet.

Interfaces de programação (em inglês, *Application Programming Interface* - API) são interfaces que disponibilizam acesso ao conjunto de serviços de um aplicativo, ou de uma plataforma Web, e *endpoints* de uma API são os URLs dos serviços disponíveis pelo software em questão. O Twitter oferece uma API com diversos *endpoints* para a coleta de dados da rede social, como tweets e dados públicos dos usuários do Twitter. Para a coleta de tweets, a API permite extrair em formato JSON os dados e conteúdos de tweets publicados no período de dias e qualquer termo ou hashtag estabelecidos como parâmetros de filtragem. Desta maneira, é possível salvar os dados coletados em um arquivo ou base de dados de livre escolha para futuros processamentos.

É importante destacar que, no início de 2021, a API do Twitter recebeu uma grande atualização, sendo lançada a API do Twitter v2. A API do Twitter v2 incluiu funções novas e avançadas, além de adicionar um nível de acesso especial para pesquisadores acadêmicos que aprimorou os recursos para esse grupo e possibilitou obter, no caso deste estudo, mais detalhes dos debates públicos [25].

A atualização da API do Twitter possibilitou, por meio de seus novos recursos e acesso especial para pesquisadores, a melhoria na coleta de dados. Esse fato pôde ser observado nesta pesquisa visto que haviam sido identificadas limitações no conjunto de dados obtido com a primeira versão da API do Twitter. Portanto, como a API do Twitter v2 aconteceu quando a pesquisa já estava na etapa de análise dos dados, foi preciso retornar para a etapa de coleta de dados, a etapa inicial da metodologia desta pesquisa, a fim de corrigir os dados limitados. Assim, devido à limitação tecnológica encontrada no decorrer desta pesquisa e a necessidade de correção dos dados, decidiu-se que as manifestações relacionadas à pandemia da COVID-19 abordadas nesta pesquisa se limitariam àquelas ocorridas no ano de 2020.

Em geral, os dados coletados para esta pesquisa são referentes ao autor do tweet (nome, identificador, biografia e verificação), data e hora de publicação do tweet, como também o texto do tweet, número identificador do tweet, hashtags utilizadas no tweet, quantidade de curtidas, quantidade de respostas, quantidade de comentários e quantidade de retweets do tweet.

3.2.1.2 Banco de Dados Relacional MySQL

Bancos de dados relacionais são bancos de dados baseados no modelo relacional, um modelo de dados que representa os dados em tabelas. Em um banco de dados relacional, cada linha de uma tabela é um registro, as colunas representam os atributos dos dados e cada registro pode ou não ter um valor para cada atributo definido. Sistema Gerenciador de Banco de Dados Relacional que operam banco de dados relacionais também são conhecidos como SGBDR. O SGBDR escolhido para esta pesquisa foi o MySQL devido ao seu reconhecido bom desempenho no processamento de grandes volumes de dados.

Os dados obtidos nas coletas por meio da API do Twitter foram salvos em arquivos de formato JSON e, posteriormente, esses arquivos de dados foram importados para tabelas do MySQL. A estrutura das tabelas aplicadas no armazenamento dos dados no SGBDR MySQL é apresentada no Apêndice A.

3.2.1.3 Website GetDayTrends

Os “Assuntos do Momento” (em inglês, *Trending Topics*) no Twitter são uma seleção dos assuntos — representados por termos ou hashtags — mais comentados durante um determinado período no Twitter em todo mundo, ou em localizações específicas, como países ou cidades. O Twitter avalia, por meio de um algoritmo, quais termos ou hashtags estão em tendência nos últimos minutos na plataforma para destacar o que é considerado em evidência.

GetDayTrends.com [26] é um *website* que registra os Assuntos do Momento no Twitter ao longo de um dia. Neste *website*, os Assuntos do Momento no Twitter são registrados e exibidos até os 50 primeiros itens, organizados por data, hora e região. Por meio do GetDayTrends.com, é possível obter as listas históricas dos Assuntos do Momento do Twitter, e então selecionar os termos e hashtags para as coletas posteriores de tweets. Por este fato, este *website* foi utilizado nesta pesquisa para auxiliar a seleção dos assuntos a serem buscados na execução da coleta de dados.

3.2.2 Estratégias de Seleção de Termos e Hashtags para as Coletas de Dados

Esta seção tem o objetivo de apresentar as estratégias aplicadas para a seleção de termos e hashtags utilizados nas coletas de dados realizadas na presente pesquisa, bem como a avaliação do desempenho de cada estratégia apresentada a respeito de seu desempenho em relação ao objetivo proposto.

Devido à natureza sociopolítica do objeto de estudo desta pesquisa, as manifestações populares nas redes sociais, realizar a seleção dos termos e hashtags por meio exclusivo dos Assuntos do Momento do Twitter não seria interessante, dado que a falta de contexto sobre as questões em debate no dia poderia produzir escolhas de termos e hashtags sem a diversidade de posicionamentos necessária para analisar cada discussão. Por esta razão, foram elaboradas três estratégias: (i) Estratégia a partir dos Assuntos do Momento, (ii) Estratégia a partir de capas de jornais impressos e (iii) Estratégia a partir de portais de notícias. As três estratégias foram experimentadas por sete dias, cada uma delas, durante o período de 24/05/2020 até 17/06/2020. A seguir, apresentam-se cada uma das três estratégias.

3.2.2.1 Estratégia a Partir dos Assuntos do Momento

Esta primeira estratégia foi organizada exclusivamente por meio do acompanhamento dos Assuntos do Momento do Twitter. O período de experimentação desta estratégia foi

de 09/06/2020 até 17/06/2020. Durante esse período, os termos e hashtags que possuíam relação com assuntos políticos e sociais do dia, presentes nos Assuntos do Momento, eram registrados para a coleta de tweets. Para essa estratégia, a janela temporal utilizada foi a das 24 horas do dia em análise.

3.2.2.2 Estratégia a Partir de Capas de Jornais Impressos

Esta segunda estratégia foi organizada por meio do acompanhamento dos Assuntos do Momento do Twitter em conjunto com as manchetes das capas dos três principais jornais do país: O Globo, Folha de São Paulo e Estadão. Essas mídias foram selecionadas pois são os jornais de maior circulação digital no Brasil, de acordo com o Instituto Verificador de Comunicação [27]. O período de aplicação deste experimento foi de 24/05/2020 até 31/05/2020. Durante esse período, a cada dia, as capas dos jornais selecionados eram documentadas e, a partir dos assuntos em destaque, buscou-se termos e hashtags presentes nos Assuntos do Momento para a coleta de tweets. Para essa estratégia, experimentou-se três janelas de tempo para seleção dos Assuntos do Momento considerando o possível horário da última tiragem do dia das capas dos jornais, 6:00 da manhã. A seguir, as três janelas de tempo propostas para a segunda estratégia:

1. Primeira janela de tempo: as 24 horas anteriores ao horário da última tiragem do dia das capas dos jornais;
2. Segunda janela de tempo: as 24 horas seguintes ao horário da última tiragem do dia das capas dos jornais;
3. Terceira janela de tempo: as 12 horas anteriores e 12 horas seguintes ao horário da última tiragem do dia das capas dos jornais;

3.2.2.3 Estratégia a Partir de Portais de Notícias

A terceira e última estratégia foi organizada por meio do acompanhamento dos Assuntos do Momento do Twitter em conjunto com as manchetes dos portais de notícias mais populares no Brasil — Globo, UOL e R7 — conforme o levantamento feito durante o primeiro trimestre de 2020 [28]. Além desses portais foram acrescentados os portais de notícias BBC Brasil e El País Brasil à estratégia em experimentação. Esses últimos portais citados são renomadas mídias internacionais com páginas sobre o Brasil que costumam apresentar repercussão no Twitter. O período de experimentação desta estratégia foi de 01/06/2020 até 08/06/2020. Durante esse período, registros das páginas iniciais dos portais de notícias selecionados eram feitos três vezes ao dia — de manhã, de tarde e de noite — e, a partir dos assuntos em destaque nas páginas, eram buscados termos e

hashtags presentes nas 50 primeiras posições dos Assuntos do Momento para a coleta de tweets. Para essa estratégia, a janela temporal utilizada foi das 24 horas do dia em análise.

3.2.2.4 Discussão Sobre as Estratégias de Seleção de Termos e Hashtags Experimentadas

A primeira estratégia, baseada nos Assuntos do Momento (Seção 3.2.2.1), foi mais veloz na execução em comparação com as outras duas estratégias experimentadas. Porém, caso não haja acompanhamento das publicações no Twitter pelo pesquisador durante o dia, ela pode deixar de ser eficiente. Isso acontece já que não há outro meio para analisar os fatos ocorridos, como capas de jornais e portais de notícias, e, dessa maneira, não há como relacionar com os termos e hashtags presentes nos Assuntos do Momento do dia.

Para a segunda estratégia, baseada em capas de jornais impressos (Seção 3.2.2.2), foi possível notar durante o período de experimentação que as capas de jornais podem tanto refletir os assuntos que estiveram em destaque no Twitter no dia anterior, como podem produzir os debates do dia na rede social observada. Essas condições vão depender do horário e da fonte da notícia em questão. Por exemplo, quando uma notícia acontece ao longo do dia anterior, as capas de jornais trazem a informação e sua repercussão na sociedade civil, entretanto, quando a notícia é dada com exclusividade pelo jornal, as redes sociais debatem sobre ela durante parte do dia. Portanto, as duas primeiras opções de janelas de tempo não apresentaram bons resultados porque, devido à condição de momento em que ocorre o fato noticiado, não é possível encontrar um padrão entre essas janelas de tempo para o registro dos Assuntos do Momento. A terceira janela de tempo também não apresentou um resultado satisfatório, visto que não foi possível superar o problema de abrangência encontrado com as duas janelas de períodos de 24 horas. Caso a notícia aconteça logo após as 6:00 da manhã e, dessa maneira, não esteja presente nas capas de jornais e a repercussão da notícia no Twitter se estenda por todo dia, essa janela temporal acaba detectando apenas parte do debate sobre a notícia em questão. Seria possível contornar o problema de pouca abrangência temporal a partir da aplicação de uma janela de tempo com as 24 horas anteriores e as 24 horas seguintes. Porém, a utilização de uma faixa temporal tão extensa na estratégia de coleta torna o processo bastante demorado de ser realizado e nem um pouco eficiente para a pesquisa, já que a janela de tempo para observação dos Assuntos do Momento é de 48 horas no total. Ademais, muitos assuntos em destaque nas redes sociais não são repercutidos pelos jornais até que se desenvolva em um acontecimento fora da internet. Portanto, a segunda estratégia não obteve bons resultados porque alguns debates só são observados após o desenrolar dos fatos.

Por fim, para a terceira estratégia, baseada em portais de notícias (Seção 3.2.2.3), foi

possível notar durante a experimentação desta estratégia que os portais de notícias acompanham mais dinamicamente os assuntos do dia, uma vez que podem ser atualizados conforme os fatos vão acontecendo. Os portais de notícias, com mais antecedência que os jornais impressos, abordam, de alguma maneira, assuntos que estão em evidência nas redes sociais. Apesar de algumas vezes os portais de notícias não registrarem o debate e toda sua extensão nas redes sociais, eles permitem detectar mais precipitadamente o assunto a ser registrado e coletado. Em comparação com as demais estratégias experimentadas, a terceira estratégia se mostrou mais eficiente, dado que é possível registrar debates desde o início e muitos dos assuntos são abordados pelas diversas seções presentes nas páginas dos portais. Provavelmente, isso se deve ao espaço disponível e capacidade de atualização desses *websites*.

Pelas características apresentadas, e em comparação com as estratégias anteriormente avaliadas, a estratégia adotada para conduzir as seleções de termos e hashtags das coletas para a referente pesquisa é a terceira estratégia, que utiliza os portais de notícias.

3.3 Processamento de Dados

Nesta seção, dividida em três subseções, será apresentado como os dados coletados foram processados para análise. A Seção 3.3.1 relata como os dados coletados e selecionados foram tratados para modelagem em grafos; a Seção 3.3.2 descreve como os dados tratados foram modelados em grafos e apresenta as métricas utilizadas nas análises; e a Seção 3.3.3 apresenta como as visualizações dos grafos foram geradas.

3.3.1 Mineração dos Dados Coletados

Para a mineração dos dados, foram executados dois procedimentos: a remoção de tweets duplicados e a normalização dos identificadores dos usuários.

As remoções dos tweets duplicados nas tabelas de cada assunto para análise foram executadas a partir de um comando SQL de exclusão de dados por meio do MySQL. Esse comando tratou de identificar, e remover, os registros que possuíam identificadores de tweet repetidos, mantendo apenas o primeiro registro encontrado para cada identificador de tweet.

Em computação, *case sensitive* define se as escritas maiúsculas e minúsculas são tratadas como distintas ou equivalentes. As ferramentas para modelagem de grafos são *case sensitive* na formação dos vértices, portanto os identificadores dos usuários mencionados

nos tweets e os identificadores dos usuários autores dos tweets tiveram suas escritas normalizadas em letras minúsculas com o propósito de evitar a duplicação de vértices na fase de modelagem de dados em grafos.

3.3.2 Modelagem dos Dados em Grafos

Sejam @u e @v dois usuários do Twitter. Diz-se que existe uma *menção* do usuário @u para o usuário @v caso aconteça uma das seguintes situações:

1. Citação direta no texto do tweet: quando o usuário @u cita o usuário @v no corpo do texto de um tweet t, *e.g.*, o tweet da Figura 3.2.
2. Resposta a um tweet: quando o usuário @u responde a um tweet t do usuário @v, *e.g.*, o tweet da Figura 3.3.
3. *Retweet*: quando o usuário @u compartilha um tweet t do usuário @v sem escrever comentário, *e.g.*, o tweet da Figura 3.4.
4. Comentário a um tweet: quando o usuário @u compartilha um tweet t do usuário v adicionado um comentário sobre o tweet, *e.g.*, o tweet da Figura 3.5.

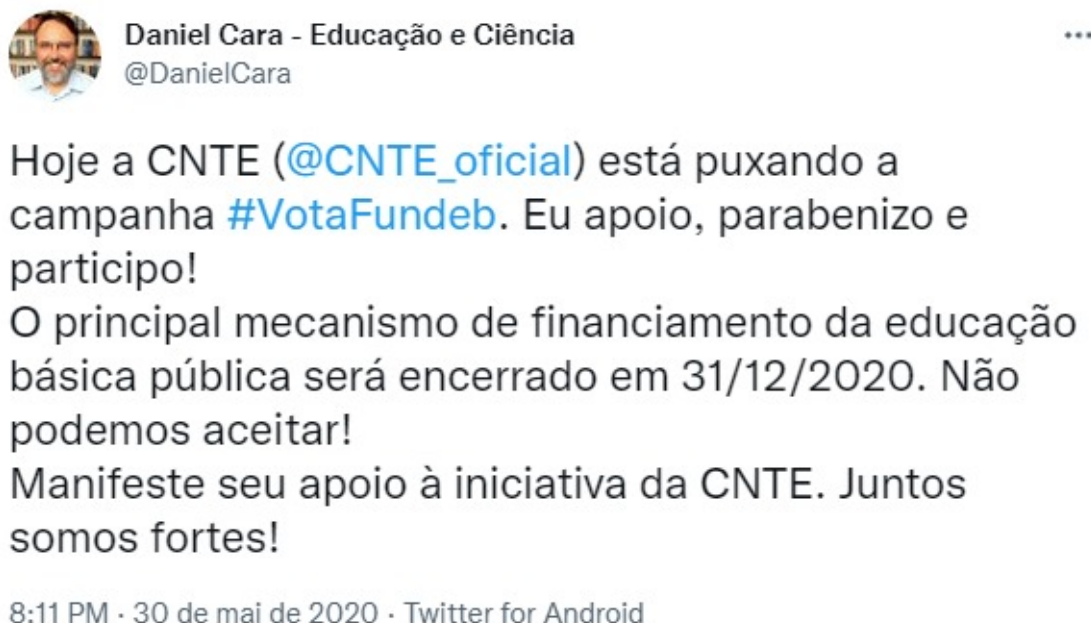


Figura 3.2: No tweet apresentado nesta captura de tela, o usuário @danielcara gera uma menção ao usuário @cnte_oficial a partir de uma citação direta no texto do tweet.

Um *Grafo de Menções* $G = (V, E)$ é um grafo direcionado no qual cada $v \in V$ corresponde a um usuário @v da rede coletada. Existe uma aresta direcionada $e = (u, v)$



Figura 3.3: No tweet apresentado nesta captura de tela, o usuário *@carlacgoliveira* gera uma menção ao usuário *@onyxlorenzoni* a partir de uma resposta ao tweet do *@onyxlorenzoni*.



Figura 3.4: No tweet apresentado nesta captura de tela, o usuário *@ranysbaby* gera uma menção ao usuário *@tretanotrampo* a partir de um retweet do tweet do *@tretanotrampo*.

em E se e somente se o usuário @u mencionou o usuário @v no tweet t. Cada tweet t



Figura 3.5: No tweet apresentado nesta captura de tela, o usuário @vo_dize gera uma menção ao usuário @jairbolsonaro a partir de um comentário ao tweet do @jairbolsonaro.

pode definir uma ou mais arestas caso haja menção a mais de um usuário. O peso $w(u, v)$ da aresta $e = (u, v)$ corresponde a quantidade de arestas paralelas entre o vértice u e o vértice v .

Como exemplo do Grafo de Menções, neste tweet publicado por @usponline: “O @minsaude oficiou a @fiocruz para divulgar cloroquina como tratamento precoce para a covid-19. Sem evidências científicas, porém, medicamentos não podem fazer parte de protocolos oficiais” há três vértices e duas arestas direcionadas devido às menções diretas no tweet. Aresta e_1 onde usponline é o vértice u_1 e minsaude é o vértice v_1 ; aresta e_2 onde usponline é o vértice u_2 e fiocruz é o vértice v_2 .

NetworkX é um pacote Python que permite a criação, manipulação e estudo das estruturas, dinâmicas e funções de redes complexas [29]. Após a coleta e mineração dos dados, foi adotado este pacote computacional para a modelagem dos dados, presentes nas tabelas da pesquisa, em grafos. A partir dos recursos do NetworkX, as métricas de grafos apresentadas nas Seções 2.1.1 e 2.1.2 foram calculadas para cada grafo obtido a partir das etapas anteriores da metodologia deste estudo. As métricas de grafos são relacionadas à aceção que elas possuem para os grafos obtidos conforme descrito a seguir:

1. Quantidade de vértices e arestas: correspondem ao tamanho do grafo.
2. Coeficiente de agrupamento global: mede a tendência dos vértices do grafo formarem grupos caracterizados por conexões mais densas. Essa métrica é útil para analisar o nível de coesão entre os vértices do grafo.
3. Densidade do grafo: é aplicada para avaliar a interconectividade no grafo. Densidade de um grafo de valor igual a 0 (zero), o valor mínimo, expressa que não há arestas entre os vértices deste grafo; densidade do grafo de valor igual a 1, o valor máximo, significa que o grafo é completo, portanto, todo vértice é adjacente a todos os outros vértices do grafo.

As métricas dos vértices são relacionadas à acepção que elas possuem para os vértices obtidos como descrito em seguida:

1. Grau de entrada: representa a quantidade de menções (arestas) que um usuário (vértice) recebeu de outros usuários nos tweets coletados que fazem parte do universo do grafo em análise. Por definição, o grau de entrada é calculado considerando-se o direcionamento das arestas.
2. Grau de saída: representa a quantidade de menções (arestas) que um usuário (vértice) fez a outros usuários nos tweets coletados que fazem parte do universo do grafo em análise. Por definição, o grau de saída é calculado considerando-se o direcionamento das arestas.
3. Centralidade de grau: contabiliza o número de arestas incidentes a um dado vértice do grafo. Por definição, são consideradas para o cálculo da centralidade de grau arestas de todos os direcionamentos, independente do vértice ser origem ou destino destas. Para centralidade de grau, valores mais altos significam que o vértice é mais central e tem posição mais importante na rede.
4. Centralidade de autovetor: aponta que um vértice está conectado a muitos vértices que possuem valores altos de centralidade de autovetor. Desta maneira, quanto maior o valor de centralidade de autovetor que um vértice possui, mais ampla é a influência deste vértice no grafo. Essa métrica ajuda a esclarecer quais são os vértices mais relevantes no grafo além dos valores de centralidade de grau.
5. Centralidade de intermediação: auxilia na identificação dos vértices que atuaram como intermediadores entre comunidades diferentes. Por possuir o papel de conector entre diversos caminhos do grafo, os vértices com maiores valores de centralidade de intermediação são chamados *gatekeepers* (porteiros em português).

6. Centralidade de proximidade: indica o quão próximo é um usuário (vértice) de todos os outros usuários (vértices) da rede (grafo) analisada. Portanto, quanto mais alto o valor de proximidade, mais rápido a informação transmitida pelo vértice alcança outros vértices no grafo.

Além dessas métricas, também foram realizadas consultas SQL na base de dados no MySQL para a identificação dos tweets mais compartilhados em cada assunto analisado. Essa contabilização permite observar, nas redes de cada assunto em análise, quais foram os sentimentos e discursos mais disseminados.

Normalmente, o maior componente conexo é utilizado para as análises sobre grafos. Entretanto, para esta pesquisa, foram calculados os tamanhos de todos os componentes conexos de cada grafo e, em seguida, seus tamanhos foram comparados de modo a avaliar se havia mais de um componente conexo com tamanho relevante nos grafos para análise. Esta verificação foi realizada apenas para garantir que nenhum componente conexo de tamanho expressivo nos grafos obtidos foi ignorado, evitando assim a possibilidade de exclusão de qualquer vértice importante das análises realizadas.

3.3.3 Visualização dos Grafos

O Gephi é um software de código aberto para análise de gráficos e redes que usa um mecanismo de renderização 3D para exibir redes grandes em tempo real permitindo a exploração das mesmas [30]. O Gephi fornece acesso fácil aos dados da rede e permite espacializar, filtrar, navegar, manipular e agrupar seus elementos [30]. Esse programa foi utilizado na geração da apresentação visual de cada grafo (ou componentes conexos) e detecção de comunidades.

O algoritmo de Louvain é um método heurístico para detectar a estrutura de comunidades de grandes redes [31]. Baseado na otimização da modularidade, segundo Blondel et al., o algoritmo de Louvain demonstrou superar outros métodos de detecção de comunidades conhecidos, em tempo de computação, entregando boa qualidade e precisão das comunidades detectadas [31]. Por este motivo, com o propósito de calcular a modularidade dos grafos e detectar suas comunidades, foi escolhido o algoritmo de Louvain, que está implementado no Gephi. Para execução do algoritmo em todos os grafos, foram mantidos os valores padrão dos parâmetros, onde: parâmetro “Aleatório” é verdadeiro, parâmetro “Usar pesos das arestas” é verdadeiro e o parâmetro “Resolução” é igual a 1,0. Após a execução, foi atribuída coloração aos vértices conforme as comunidades em que eles foram alocados.

Em seguida, foi atribuído, para cada vértice, tamanho visual proporcional à sua centralidade de grau. Os tamanhos estão em uma escala de 1 (tamanho mínimo) a 50 (tamanho máximo).

O Gephi disponibiliza algoritmos para trabalhar a disposição gráfica das redes de modo a obter uma melhor visualização delas. Entre esses algoritmos, dois foram selecionados: OpenOrd e ForceAtlas 2. Esses dois algoritmos foram aplicados de forma combinada a todas as redes estudadas para produzir uma versão gráfica otimizada para as análises. Com exceção do parâmetro “Gravidade mais forte” do ForceAtlas 2, configurado como verdadeiro para todos os grafos, os demais parâmetros de cada algoritmo foram ajustados segundo as características de cada grafo.

No próximo capítulo serão detalhadas as análises de cada um dos grafos formados a partir da coleta dos dados dos assuntos selecionados. Após as análises individuais, será feita uma análise conjunta de todas as redes, destacando os padrões observados ou as características a serem destacadas.

4. Análise dos Dados

Este capítulo é destinado à análise dos dados coletados do Twitter e está organizado por manifestações sobre os assuntos selecionados. Os assuntos selecionados para as análises nesta pesquisa foram: 1) Novo FUNDEB (Seção 4.1); 2) Breque dos Apps (Seção 4.2); 3) Movimentos #Somos70PorCento e #Somos57MILHOES (Seção 4.3); 4) Aprovação e pagamento do auxílio emergencial (Seção 4.4); e 5) Realização do ENEM 2020 (Seção 4.5). Tais assuntos foram selecionados para análise devido a sua importância para a esfera pública ou sua relevância no contexto da pandemia da COVID-19.

4.1 Manifestações Sobre o Novo FUNDEB

4.1.1 Contexto

O Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (FUNDEB) é uma política de financiamento da educação pública nacional em vigor desde janeiro de 2007 [32]. Esse fundo especial tem por finalidade promover a redistribuição de recursos, desde reformas em escolas até salários de professores da educação básica pública, da creche ao ensino médio, de todos os entes federativos brasileiros [32]. O FUNDEB é especialmente importante para os estados e municípios que não conseguem investir o mínimo exigido por aluno.

Por estar para expirar por lei no final do ano de 2020 [33], uma Proposta de Emenda à Constituição (PEC) com a proposta de renovação, ampliação e consolidação do FUNDEB começou a ser debatida no Congresso Nacional a partir de 2015. A relatoria da PEC na Câmara Federal e no Senado Nacional ficaram, respectivamente, a cargo da deputada Professora Dorinha e do senador Flávio Arns.

Ao iniciar o ano de 2020, a urgência de aprovação do novo FUNDEB se torna evi-

dente devido ao seu prazo de validade. Embora o Governo Federal, sob comando de Jair Bolsonaro, não tenha se manifestado publicamente sobre o FUNDEB, a base governista no Legislativo foi orientada a barrar a votação das PECs relacionadas ao novo FUNDEB. Apesar da falta de apoio do Governo Federal, a matéria foi aprovada com ampla maioria, em dois turnos, na Câmara dos Deputados em 21 de julho de 2020 sob a presidência do deputado Rodrigo Maia e, posteriormente, no Senado Federal em 25 de agosto de 2020 sob a presidência do senador Davi Alcolumbre, submetendo uma derrota ao poder executivo. Manifestações online ocorreram periodicamente entre os meses de maio de 2020 e agosto de 2020 até a aprovação da matéria na Câmara dos Deputados e, em seguida, no Senado Federal.

De modo a contextualizar tais manifestações, foi gerada uma lista dos dez tweets com mais interações¹ (Apêndice B) a partir dos tweets obtidos nas coletas realizadas para esta análise. Sete dos dez tweets com mais interações são em apoio ao novo FUNDEB, sendo que seis destes foram publicados por políticos de oposição ao governo de Jair Bolsonaro, como, por exemplo, Marcelo Freixo e Guilherme Boulos. Em geral, os tweets citados demonstram apoio à aprovação do novo FUNDEB, explicam a importância do FUNDEB ou comemoram a aprovação da PEC do novo FUNDEB.

4.1.2 Coletas de Dados

Os termos e hashtags utilizados para a coleta de tweets relacionados ao debate acerca do novo FUNDEB, como também suas respectivas datas de coletas, realizadas no período entre 29/05/2020 e 26/08/2020, são descritos na Tabela 4.1, em que a primeira coluna (#) mostra o número de identificação da coleta, a segunda coluna indica o período selecionado para a coleta, e a terceira coluna relaciona os termos e hashtags observados nas publicações.

A coleta número 10 obteve um tweet de humor publicado no dia 20/07/2020 devido ao termo “educação” que era uma das palavras do texto. Esse tweet foi muito compartilhado durante os dias seguintes a sua publicação, o que fez o vértice *ryanlucazz*, autor do tweet, se tornar o vértice de maior grau de entrada no grafo que será apresentado na próxima seção. Buscou-se refazer a coleta número 10 sem o termo “educação” para remover esse vértice e os demais vértices associados que não estão relacionados às manifestações do Novo FUNDEB. Entretanto, ao realizar a mudança na coleta, foi notado que os apoiadores do Governo Federal buscaram não citar diretamente o FUNDEB em seus tweets e sim usar palavras relacionadas ao tema, como “educação”. Desta maneira, ao remover o termo

¹ Para esta listagem, foram consideradas como interações retweets, respostas e comentários.

Tabela 4.1: Descrição dos dados das coletas de tweets relacionadas ao FUNDEB.

#	Período de publicação	Termos e hashtags observados
1	De 29/05/2020 até 31/05/2020	#VotaFundeb
2	De 30/05/2020 até 01/06/2020	#VotaFundeb
3	De 22/06/2020 até 24/06/2020	#VotaFundeb
4	De 29/06/2020 até 01/07/2020	#VotaFundeb
5	De 06/07/2020 até 08/07/2020	#VotaFundeb
6	De 13/07/2020 até 15/07/2020	#VotaFundeb
7	De 18/07/2020 até 20/07/2020	Fundeb
8	De 19/07/2020 até 21/07/2020	Fundeb, #AprovaFundeb, #Vota-Fundeb e #VotaFundebJa
9	De 20/07/2020 até 22/07/2020	Fundeb, #AprovaFundeb e #APROVAFUNDEBJA
10	De 21/07/2020 até 23/07/2020	Educação, Fundeb, #APROVAFUNDEBJA, #fundebaprovado
11	De 28/07/2020 até 30/07/2020	#AprovaFundeb
12	De 29/07/2020 até 31/07/2020	#AprovaFundeb
13	De 11/08/2020 até 13/08/2020	#AprovaFundeb
14	De 17/08/2020 até 19/08/2020	#FundebComCAQ
15	De 18/08/2020 até 20/08/2020	#AprovaFundeb
16	De 19/08/2020 até 21/08/2020	#AprovaFundeb e #FundebCom-CAQ
17	De 24/08/2020 até 26/08/2020	Fundeb e #AprovaFundeb

“educação” da referida coleta, são perdidos vértices importantes para as manifestações, portanto, optou-se por manter a coleta original e desconsiderar o vértice *ryanlucazz* da análise.

4.1.3 Apresentação do Grafo de Menções G_1

A partir do conjunto de dados reunido pelas coletas exibidas na Tabela 4.1, foi obtido um Grafo de Menções G_1 relacionado às manifestações em torno do FUNDEB. O Grafo G_1 possui 156.049 vértices e 466.557 arestas. Além disso, G_1 não é fortemente conexo e tem, nessas condições, 4.104 componentes que não são fortemente conexos.

A distribuição do tamanho dos componentes conexos do Grafo G_1 pode ser observada na Figura 4.1. É possível notar que há em G_1 um componente conexo consideravelmente maior em relação ao número de vértices que os demais componentes. Essa informação mostra que um componente conexo concentra as conexões do grafo. Por esta razão, a análise prosseguirá sobre o maior componente não-fortemente conexo de G_1 , que será identificado como CG_1 . Neste Componente CG_1 , há 145.764 vértices e 459.916 arestas.

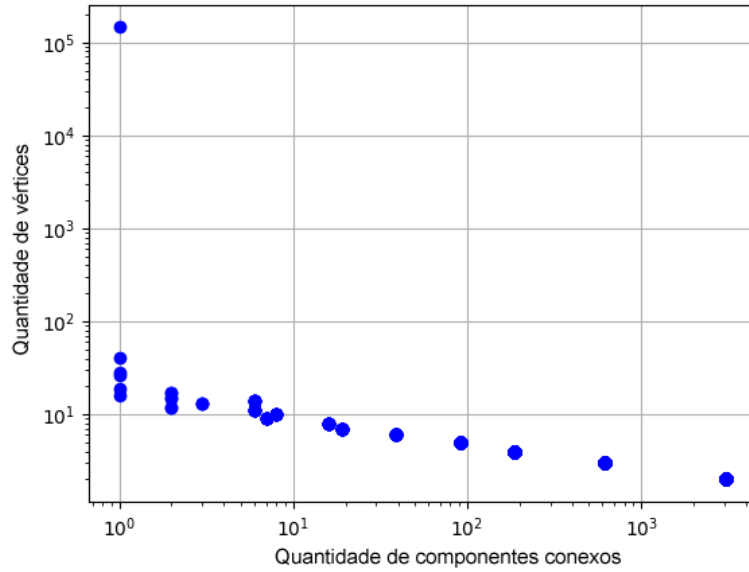


Figura 4.1: Distribuição do tamanho dos componentes no Grafo G_1 (eixos ajustados para escala logarítmica).

4.1.4 Apresentação das Métricas Aplicadas ao Componente Conexo CG_1

Para o Componente Conexo CG_1 , a centralidade de grau máximo de um vértice é igual a 34.229 e a centralidade de grau mínimo de um vértice é igual a 1, tendo os vértices centralidade de grau médio igual a 6,3104.

De modo a avaliar a tendência de distribuição dos valores de centralidade de grau, foi gerado um histograma desses valores distribuídos pelo número de vértices presentes no Componente CG_1 (Figura 4.2). Para esse histograma, os valores do eixo x foram normalizados dividindo-se o valor da centralidade de grau do vértice pelo grau máximo possível, dado por $n - 1$, em um grafo não-direcionado obtido a partir de CG_1 , onde n é o número de vértices em CG_1 ; e os valores do eixo y foram ajustados para escala logarítmica.

Para entender quais são os perfis dos vértices com os maiores valores de centralidade de grau foram geradas duas listas: uma lista apresentando os dez vértices com maiores valores de grau de entrada (Tabela 4.2) no Componente CG_1 e outra lista contendo os dez vértices com maiores valores de grau de saída (Tabela 4.3) no Componente CG_1 .

Para ajudar a esclarecer quais são os outros vértices mais relevantes no grafo, além dos valores de centralidade de grau, foram calculadas as centralidades de autovetor dos vértices presentes no Componente CG_1 . A relevância, nesse caso, se dá por meio de suas conexões com outros vértices centrais no grafo. Os dez vértices com os maiores valores

Tabela 4.2: Vértices com os dez maiores valores de grau de entrada no Componente CG_1 .

Vértice	Grau de entrada	Grau de saída	Centralidade de grau
ryanlucazz	34.229	0	34.229
marcelofreixo	12.264	13	12.277
guilhermeboulos	6.301	4	6.305
gfuza_oficial	5.385	0	5.385
haddad_fernando	4.983	2	4.985
biakicis	4.653	1	4.654
samiabomfim	4.647	20	4.667
jandira_feghali	4.633	2	4.635
allantercalivre	4.468	5	4.473
jairbolsonaro	4.113	0	4.113

Tabela 4.3: Vértices com os dez maiores valores de grau de saída no Componente CG_1 .

Vértice	Grau de saída	Grau de entrada	Centralidade de grau
cnte_oficial	745	1.277	2.022
manojoliveira1	621	109	730
socritiki	556	1	557
squerubina	550	9	559
synthese	540	48	588
paulomariante	540	46	586
undime	507	114	621
marleneprudente	503	62	565
camp_educacao	487	970	1.457
cirofuturo	472	19	491

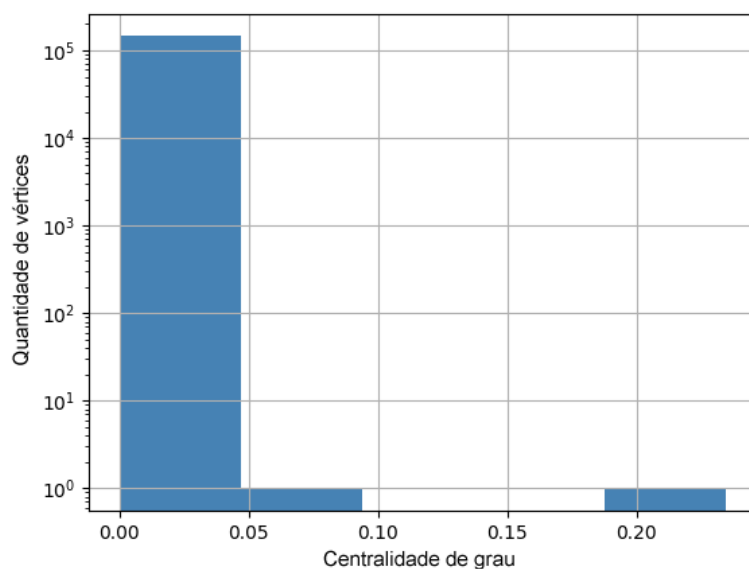


Figura 4.2: Distribuição da centralidade de grau no Componente CG_1 .

de centralidade de autovetor são mostrados na Tabela 4.4.

Tabela 4.4: Vértices com os dez maiores valores de centralidade de autovetor no Componente CG_1 .

Vértice	Centralidade de autovetor	Grau de entrada	Grau de saída	Centralidade de grau
profdorinha	0,2354	2.214	8	2.222
cnte_oficial	0,1803	1.277	745	2.022
danielcara	0,1709	3.705	64	3.769
senadofederal	0,1625	1.316	10	1.326
ubesoficial	0,1494	1.736	177	1.913
uneoficial	0,1490	2.597	63	2.660
rodrigomaia	0,1489	3.208	0	3.208
camp_educacao	0,1485	970	487	1.457
camaradeputados	0,1374	1.418	10	1.428
arnsflavio	0,1311	347	10	357

Com o intuito de avaliar os principais vértices na troca de informações no Componente CG_1 , foi gerada a listagem dos dez maiores valores de centralidade de proximidade no Componente CG_1 . Essa listagem é exibida na Tabela 4.5.

Buscando entender a conectividade entre os vértices no Componente CG_1 , são apresentados os dados sobre o coeficiente de agrupamento médio no Componente CG_1 , que é de 0,1373, e sobre a densidade no Componente CG_1 , que é de 0,00002. Como o Componente CG_1 não é fortemente conexo, não foi possível calcular o diâmetro e a distância média.

Tabela 4.5: Vértices com os dez maiores valores de centralidade de proximidade no Componente CG₁.

Vértice	Centralidade de proximidade	Grau de entrada	Grau de saída	Centralidade de grau
marleneprudente	0,0290	62	503	565
manajuoliveira1	0,0284	109	621	730
izacrispi	0,0280	63	334	397
socritiki	0,0278	1	556	557
xiicaaa	0,0277	3	285	288
squerubina	0,0277	9	550	559
eri_leite	0,0275	29	376	405
avanci1983	0,0273	189	222	411
luladilma5	0,0272	507	301	808
carlos_juniorpt	0,0271	6	254	260

Com a intenção de compreender como o Componente CG₁ pode ser dividido em comunidades, foi aplicado, por meio do Gephi, o algoritmo Louvain no Componente CG₁. Foram obtidas 76 comunidades a partir desse algoritmo. A Figura 4.3 mostra a representação visual do resultado obtido. Os vértices têm seus identificadores exibidos com tamanho de fonte proporcional às suas centralidades de grau, e quanto mais centrais ao grafo, mais próximos ao centro do círculo são posicionados. As comunidades são posicionadas próximas às comunidades que tiveram mais relações. As principais comunidades detectadas podem ser identificadas na Figura 4.3 a partir das cores com as definições dadas a seguir. As definições foram inferidas a partir da avaliação de amostras dos vértices atribuídos às comunidades identificadas pelo algoritmo Louvain.

1. Vermelho, laranja e rosa: comunidades compostas por vértices favoráveis ao FUNDEB e de oposição ao Governo Federal.
2. Amarelo e roxo: comunidades compostas por vértices tanto favoráveis quanto neutros ao FUNDEB e de posicionamentos variados em relação ao Governo Federal.
3. Azul: comunidades compostas por vértices neutros ou contrários ao FUNDEB e da base do Governo Federal.
4. Verde: comunidades compostas por vértices contrários ao FUNDEB e apoiadores do Governo Federal.

Para entender quais foram os principais responsáveis pela transferência de informação entre as comunidades do Componente CG₁, foi gerada a Tabela 4.6 que exibe os dez vértices com maior valor de centralidade de intermediação.

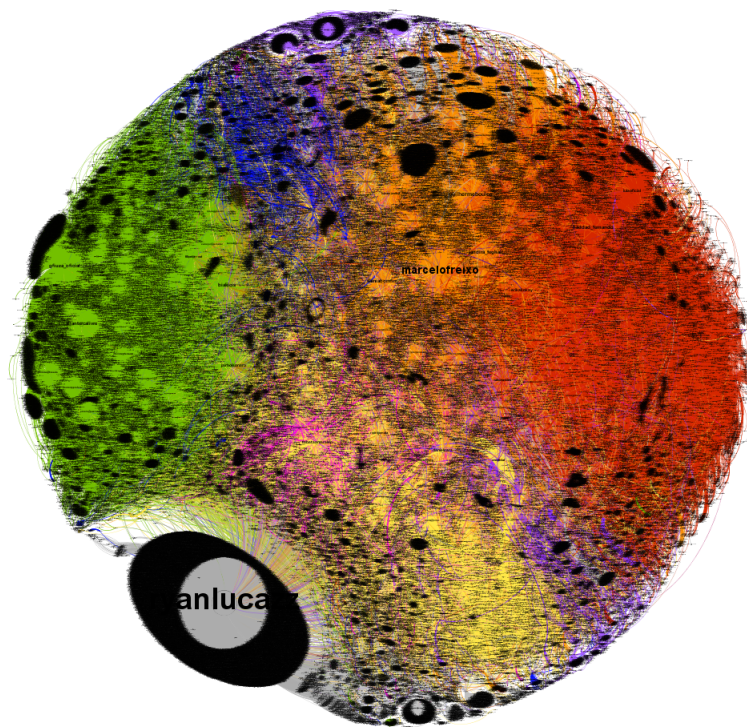


Figura 4.3: Representação visual do Componente CG_1 .

Tabela 4.6: Vértices com os dez maiores valores de centralidade de intermediação no Componente CG_1

Vértice	Centralidade de intermediação	Grau de entrada	Grau de saída	Centralidade de grau
cnte_oficial	0,0098	1.277	745	2.022
danielcara	0,0069	3.705	64 ff	3.769
camp_educacao	0,0059	970	487	1.457
marcelofreixo	0,0045	12.264	13	12.277
ubesoficial	0,0043	1.736	177	1.913
cpersoficial	0,0039	504	289	793
majorvitorhugo	0,0035	2.058	8	2.066
renatosimoespt	0,0035	350	225	575
fa1rynight	0,0035	32	7	39
sandrasanmag	0,0028	10	17	27

A Tabela 4.7 mostra um resumo das métricas de grafo para o Grafo G_1 e para o Componente CG_1 .

Tabela 4.7: Resumo das métricas de grafo para o Grafo G_1 e o Componente CG_1 .

	Grafo G_1	Componente CG_1
Vértices	156.049	145.764
Arestas	466.557	459.916
Centralidade de grau mínima	1	1
Centralidade de grau máxima	34.229	34.229
Centralidade de grau média	5,9796	6,3104
Coeficiente de agrupamento médio	0,0519	0,1373
Densidade	0,00002	0,00002
Comunidades	-	76

4.1.5 Análise dos Resultados

Para o Componente CG_1 é observado, a partir da Figura 4.2, que a distribuição de grau se assemelha a cauda longa, uma característica de redes reais. Os dados sobre a distribuição e média de centralidade de grau apresentados na Seção 4.1.4 mostram que a origem, ou o destino, da maioria das arestas existentes se concentra em poucos vértices presentes no Componente CG_1 , *i.e.*, as militâncias presentes nesta rede se concentram em torno de poucos atores. Essa informação é importante para entender, a partir da concentração das arestas em poucos vértices, o fluxo de informação na rede em análise. A concentração em valores baixos de centralidade de grau vista no Componente CG_1 é uma propriedade do modelo de redes sem escala, apresentado na Seção 2.1.3.

Observa-se, na Tabela 4.2, que, entre os vértices com os maiores valores de grau de entrada, figuram: o presidente Jair Bolsonaro (*jairbolsonaro*); os deputados da oposição Marcelo Freixo (*marcelofreixo*), Sâmia Bomfim (*samiabomfim*) e Jandira Feghali (*jandira_feghali*); a deputada governista Bia Kicis (*biakicis*); os políticos da oposição Guilherme Boulos (*guilhermeboulos*) e Fernando Haddad (*haddad_fernando*); o jornalista bolsonarista Guilherme Fiuza (*gfiuza_oficial*); e o blogueiro de extrema-direita, dono da página Terça Livre e investigado pela Polícia Federal por disseminação de notícias falsas e apoio a atos antidemocráticos, Allan dos Santos (*allantercalivre*) [34]. Portanto, 70% dos vértices que figuram entre os mais citados pertencem a políticos com cargos federais na atual legislatura (2018–2022) ou que disputaram eleições majoritárias recentemente. Também é importante destacar que, dentre esses políticos, 71,4% são representantes da oposição, e que a base governista ficou caracterizada pelos jornalistas apoiadores do Presidente. Este fato deve estar relacionado à falta de um posicionamento por parte do Governo Federal sobre o apoio ou não pela aprovação do novo FUNDEB. Desta maneira, os deputados e políticos do governo ficaram sem evidência nas manifestações online sobre a matéria.

Na Tabela 4.3, os vértices que aparecem com os maiores valores de grau de saída são referentes a civis ou movimentos sociais de trabalhadores da educação básica do Brasil, como a Confederação Nacional dos Trabalhadores em Educação - CNTE (*cnte_oficial*), União Nacional dos Dirigentes Municipais de Educação - UNDIME (*undime*), Sindicato dos Trabalhadores em Educação Básica da Rede Oficial do Estado de Sergipe - SINTESE (*sintese*) e a Campanha Nacional pelo Direito à Educação (*camp_educacao*). Isto indica que 40% dos mobilizadores mais ativos pelo novo FUNDEB são vértices de contas do Twitter vinculadas ao ativismo pelos direitos dos educadores e melhoria da educação brasileira.

A partir dos dados sobre centralidade de grau, grau de entrada e grau de saída, apresentados nas Tabelas 4.2 e 4.3, é possível considerar que o fluxo de informação na militância sobre o novo FUNDEB inicia-se a partir de grupos ou pessoas interessados em pautas relacionadas à educação brasileira como, por exemplo, os usuários *camp_educacao* e *cnte_oficial* e destina-se, diretamente ou indiretamente, a políticos responsáveis por pausar, articular e aprovar a PEC como, por exemplo, os usuários *marcelofreixo* e *biakicis*.

Pode ser observado na Tabela 4.4 que, entre os dez vértices com os maiores valores de centralidade de autovetor, metade deles são de políticos ou instituições políticas diretamente responsáveis pela tramitação da PEC sobre o novo FUNDEB: Professora Dorinha (*profdorinha*), Senado Federal (*senadofederal*), Rodrigo Maia (*rodrigomaia*), Câmara dos Deputados (*camaradeputados*) e Flávio Arns (*arnsflavio*). Como esses vértices são relatores das PECs e representantes da Câmara Federal e do Senado Nacional, explica-se a relevância desses vértices no Componente CG₁. Há também vértices relacionados a representantes estudantis, como a União Nacional dos Estudantes - UNE (*uneoficial*), ou de trabalhadores da Educação no âmbito nacional, como a CNTE (*cnte_oficial*), demonstrando a importância dessas organizações na manifestação. Mas é importante destacar aqui a relevância que Daniel Cara (*danielcara*) — professor da Universidade de São Paulo (USP), dirigente da Campanha Nacional pelo Direito à Educação e político filiado ao PSOL — teve pois, apesar de não aparecer entre os vértices com maiores valores de grau de entrada ou grau de saída, ele surge como o terceiro vértice mais relevante no Componente CG₁ de acordo com a centralidade de autovetor.

Nota-se que 40% dos vértices com maiores valores de centralidade de proximidade, apresentados na Tabela 4.5, são de vértices de usuários comuns que possuem valores de grau de saída elevados para o Componente CG₁, como, por exemplo, os usuários *marleneprudente* e *manojuliveira1*. Essa informação indica que os usuários civis buscaram citar, comentar e responder outros usuários durante as manifestações sobre o novo FUNDEB. Ao analisar as publicações no Twitter dos usuários listados com maiores valores

de centralidade de proximidade, é identificado que todos são de oposição ao Governo Federal e apoiaram o FUNDEB. Então os vértices com maiores valores de centralidade de proximidade tiveram um papel importante para divulgar e apoiar a causa durante as manifestações online.

Observando-se os baixos valores de coeficiente de agrupamento médio (0,1373) e densidade (0,00002) do Componente CG_1 , apresentados na Seção 4.1.4, é possível concluir que os vértices do Componente CG_1 não formaram muitas conexões entre si. Incluindo na análise os dados obtidos sobre grau de entrada, grau de saída e centralidade de proximidade no Componente CG_1 , entende-se, a partir dos perfis dos vértices listados, que os vértices anônimos como, por exemplo, *marlenepudente* e *manojuliveira1*, buscaram mencionar outros vértices que são atores políticos como, por exemplo, *marcelofreixo* e *biakicis*, e poucos vértices realizaram debates entre si.

É possível ver, a partir da Figura 4.3, que a maior parte do Componente CG_1 é formada pelas comunidades de apoio ao novo FUNDEB. As comunidades contrárias ao FUNDEB se limitaram aos apoiadores do presidente Jair Bolsonaro e vértices de atores de extrema-direita [34], como *allantercalivre* e *biakicis*. Por fim, pode-se também concluir que, como os resultados das votações da matéria no Legislativo, a comunidade do governo (a área verde) sofreu uma derrota dentro desse debate no Twitter visto que possui uma presença menor no gráfico e isolamento das demais comunidades detectadas, dificultando a disseminação de seus discursos para outros vértices. A derrota dos governistas no debate nesta plataforma de rede social corresponde com a derrota do governo no tema em questão, já que o FUNDEB foi aprovado apesar da posição contrária dos principais atores governistas na Câmara e no Senado. De tal maneira, o debate no Twitter refletiu a realidade.

4.2 Manifestações Sobre a *Breque dos Apps*

4.2.1 Contexto

Nos últimos cinco anos, houve um crescimento acentuado do desemprego no Brasil [35]. Muitos trabalhadores, diante das dificuldades enfrentadas para conseguir empregos formais nesses anos, encontraram como alternativa para seu sustento o trabalho informal oferecido por aplicativos de entregas de comida ou mercadorias como, por exemplo, iFood, Rappi e Uber Eats. Entretanto, esse serviço prestado aos aplicativos de entrega não gera vínculo formal, portanto, os trabalhadores ficam desamparados das condições mínimas de trabalho garantidas pela Consolidação das Leis do Trabalho (CLT). Essa relação de trabalho sustentou a queda do desemprego apresentada em 2019 [36].

Diante da pandemia da COVID-19, bares e restaurantes foram proibidos de receber clientes em suas lojas em diversas cidades do país como medida para reduzir a contaminação entre a população [37]. Como alternativa, muitos estabelecimentos buscaram a entrega de comida por meio de aplicativos para o período em que estiveram interditados. Dessa maneira, a demanda de serviço para os entregadores, como também riscos de contaminação, aumentaram. As insatisfações já existentes entre esses trabalhadores e as condições precárias de segurança oferecida aos entregadores pelas empresas de entregas resultaram em uma mobilização nacional da categoria dos entregadores de aplicativos por melhoria de trabalho. Essa mobilização resultou nas manifestações chamadas “Breque dos Apps” que ocorreram em dois dias: a primeira no dia 1 de julho de 2020 [38] e a segunda no dia 25 de julho de 2020 [39]. A Breque dos Apps contou com a participação da categoria dos trabalhadores de entregas em carreatas de motocicletas e dos clientes dos aplicativos de entregas em mobilizações online nas redes sociais.

Ao longo desta seção, serão analisados os dados referentes às manifestações online ocorridas no Twitter acerca da Breque dos Apps. Antes de iniciar a análise dos dados, é importante compreender quais foram os discursos mais compartilhados durante a manifestação sobre a Breque dos Apps para contextualizar a análise. Para tal, foi gerada uma lista dos dez tweets com mais interações² a partir dos tweets extraídos nas coletas realizadas. Essa lista pode ser encontrada no Apêndice C. Todos os dez tweets com mais interações são em apoio à Breque dos Apps, sendo quatro deles publicados pelo Paulo Galo, conhecido também como Galo de Luta, líder dos Entregadores Antifascistas e um dos principais nomes das greves dos entregadores de aplicativos. Paulo Galo foi um notório porta-voz dessa manifestação. Em geral, os dez tweets com mais interações demonstram apoio às manifestações da Breque dos Apps, justificam o movimento e dão instruções de como ajudar na Breque dos Apps.

4.2.2 Coletas de Dados

Os termos e hashtags utilizados para a coleta de tweets relacionados a esse debate, como também suas respectivas datas de coletas, realizadas no período entre 29/06/2020 e 26/07/2020, estão descritos na Tabela 4.8, em que a primeira coluna (#) mostra o número de identificação da coleta, a segunda coluna indica o período selecionado para a coleta, e a terceira coluna relaciona os termos e hashtags observados nas publicações.

²Para esta listagem, foram consideradas como interações retweets, respostas e comentários.

Tabela 4.8: Descrição das coletas de tweets relacionadas à greve dos entregadores dos aplicativos.

#	Período de publicação	Termos e hashtags observados
1	De 29/06/2020 até 01/07/2020	#AmanhaTemBrequeDosApps e #BrequeDosApps
2	De 30/06/2020 até 02/07/2020	#BrequeDosApps
3	De 23/07/2020 até 25/07/2020	#BrequeDosApps
4	De 24/07/2020 até 26/07/2020	#BrequeDosApps

4.2.3 Apresentação do Grafo de Menções G_2

Dado o conjunto de dados resultado das coletas, foi obtido o Grafo de Menções G_2 relacionado às militâncias sobre a Breque dos Apps que possui 83.244 vértices e 185.365 arestas únicas. A partir da ferramenta NetworkX, obtém-se a informação de que o Grafo G_2 não é fortemente conexo e possui 521 componentes que não são fortemente conexos. Dessa forma, faz-se necessário observar o tamanho dos componentes do Grafo G_2 para selecionar quais componentes analisar mais detalhadamente. Com esse objetivo, um gráfico foi gerado (Figura 4.4) para avaliar a distribuição de tamanho dos componentes conexos do Grafo G_2 e possibilitar a escolha dos componentes a serem analisados.

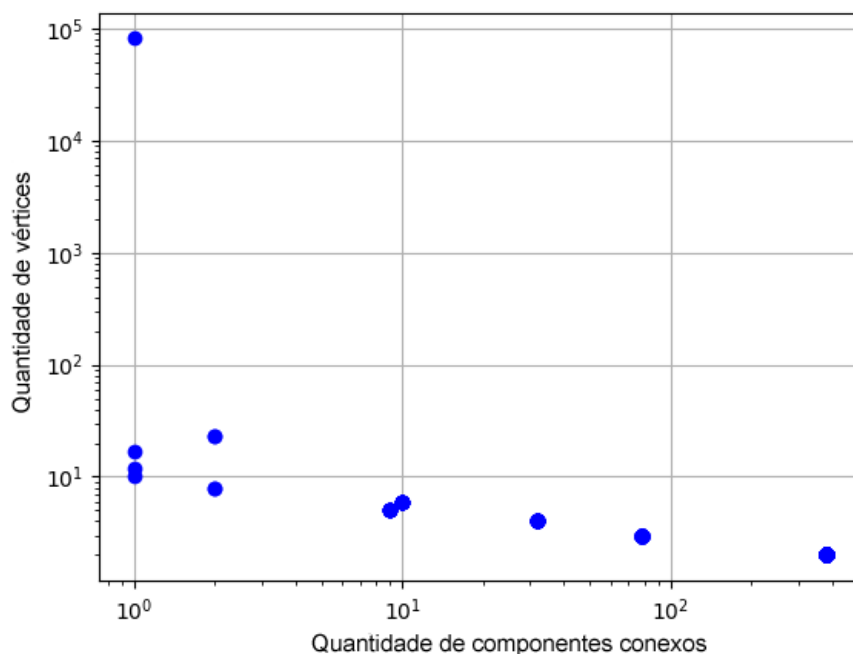


Figura 4.4: Distribuição do tamanho dos componentes no Grafo G_2 (eixos ajustados para escala logarítmica).

A partir da Figura 4.4, observa-se que há no Grafo G_2 um componente que se destaca dos demais componentes na quantidade de vértices no Grafo G_2 , concentrando muitos

vértices. Portanto, a análise prosseguirá sobre esse componente do Grafo G_2 , que será identificado como CG_2 . Há 81.908 vértices e 184.500 arestas presentes no Componente CG_2 .

4.2.4 Apresentação das Métricas Aplicadas ao Componente Conexo CG_2

A centralidade de grau máximo de um vértice no Componente CG_2 é 17.318, a centralidade de grau mínimo de um vértice no Componente CG_2 é 1 e a centralidade de grau médio dos vértices no Componente CG_2 é 4,5051. Para melhor visualização da distribuição dos valores de centralidade de grau, foi gerado um histograma desses valores normalizados pelo número de vértices presentes no Componente CG_2 (Figura 4.5). Para esse histograma, os valores do eixo x foram normalizados dividindo-se o valor da centralidade de grau do vértice pelo grau máximo possível, dado por $n - 1$, em um grafo não-direcionado obtido a partir de CG_2 , onde n é o número de vértices em CG_2 ; e os valores do eixo y foram ajustados para escala logarítmica.

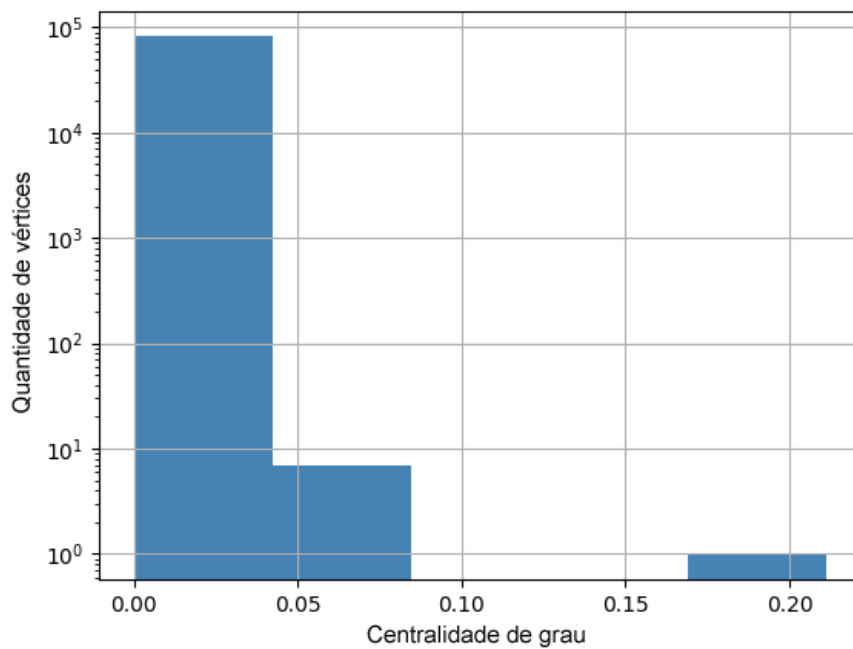


Figura 4.5: Distribuição da centralidade de grau no Componente CG_2 .

Foram geradas duas listas: uma lista apresentando os dez vértices com maiores valores de grau de entrada (Tabela 4.9) no Componente CG_2 e outra lista contendo os dez vértices com maior valor de grau de saída (Tabela 4.10) no Componente CG_2 . Também foram calculadas as centralidades de autovetor dos vértices presentes no Componente CG_2 . Os dez vértices com os maiores valores de centralidade de autovetor estão relacionados na

Tabela 4.11.

Tabela 4.9: Vértices com os dez maiores valores de grau de entrada no Componente CG₂.

Vértice	Grau de entrada	Grau de saída	Centralidade de grau
galodeluta	17.270	48	17.318
tretanotrampo	6.239	160	6.399
gduvivier	5.633	2	5.635
marcelofreixo	5.358	6	5.364
nathfinancas	4.933	5	4.938
gabrielaz	4.753	5	4.758
manueladavila	3.693	1	3.694
yurimarcas	3.664	0	3.664
andrezadelgado	3.309	0	3.309
midianinja	3.057	16	3.073

Tabela 4.10: Vértices com os dez maiores valores de grau de saída no Componente CG₂.

Vértice	Grau de saída	Grau de entrada	Centralidade de grau
izabellula	416	15	431
luninha_da_van	238	10	248
jussaracarneiro	195	9	204
silmarq	191	20	211
mata_a_dentro	182	11	193
odio_nao	181	0	181
rogsiqueira	179	8	187
tretanotrampo	160	6.239	6.399
luladilma5	140	295	435
robson_m_matos	136	27	163

Tabela 4.11: Vértices com os dez maiores valores de centralidade de autovetor no Componente CG₂.

Vértice	Centralidade de autovetor	Grau de entrada	Grau de saída	Centralidade de grau
galodeluta	0,4928	17.270	48	17.318
tretanotrampo	0,4857	6.239	160	6.399
gabrielaz	0,1615	4.753	5	4.758
andrezadelgado	0,1265	3.309	0	3.309
midianinja	0,1220	3.057	16	3.073
sarahcaos	0,1171	1.766	22	1.788
florababylon	0,1126	517	29	546
dimitravulcana	0,1056	364	17	381
transcurecer	0,1043	2.533	40	2.573
acacio1871	0,1041	540	34	574

Com a intenção de entender sobre a troca de informações entre os vértices, foi gerada a listagem dos dez maiores valores de centralidade de proximidade no Componente CG₂. Essa listagem está exibida na Tabela 4.12.

Tabela 4.12: Vértices com os dez maiores valores de centralidade de proximidade no Componente CG₂.

Vértice	Centralidade de proximidade	Grau de entrada	Grau de saída	Centralidade de grau
marceloyerion	0,0132	6	72	78
luninha_da_van	0,0131	10	238	248
izabellula	0,0131	15	416	431
tadeu_alves	0,0129	4	71	75
marcia07453765	0,0129	5	62	67
gaylulista2	0,0128	2	52	54
luladilma5	0,0127	295	140	435
marise_vaz	0,0126	4	69	73
silmarq	0,0124	20	191	211
croismo	0,0124	1	136	137

Para o Componente CG₂, o coeficiente de agrupamento médio é 0,1661 e a densidade do componente é 0,00003. Não foi possível calcular a distância média e diâmetro do Componente CG₂ pois ele não é fortemente conexo.

Além disso, foi aplicado, por meio do Gephi, o algoritmo Louvain no Componente CG₂ para detecção de comunidades. Foram obtidas 45 comunidades a partir desse algoritmo. A Figura 4.6 mostra a representação visual do resultado obtido pelo algoritmo executado. Os vértices têm seus identificadores exibidos com tamanho de fonte proporcional às suas centralidades de grau, e quanto mais centrais ao grafo, mais próximos ao centro do círculo são posicionados. As comunidades são posicionadas próximas às comunidades que tiveram mais relações. As principais comunidades detectadas podem ser identificadas na Figura 4.6 a partir das cores com as definições dadas a seguir. As definições foram inferidas a partir da avaliação de amostras dos vértices atribuídos às comunidades identificadas pelo algoritmo Louvain.

1. Vermelho e roxo: comunidades compostas por vértices representantes e apoiadores da Breque dos Apps.
2. Rosa e laranja: comunidades compostas por grupos políticos apoiadores da Breque dos Apps.
3. Marrom: comunidades compostas por militantes do movimento negro apoiadores da Breque dos Apps.

4. Amarelo: comunidades compostas por vértices relacionados à cobertura jornalística da Breque dos Apps.

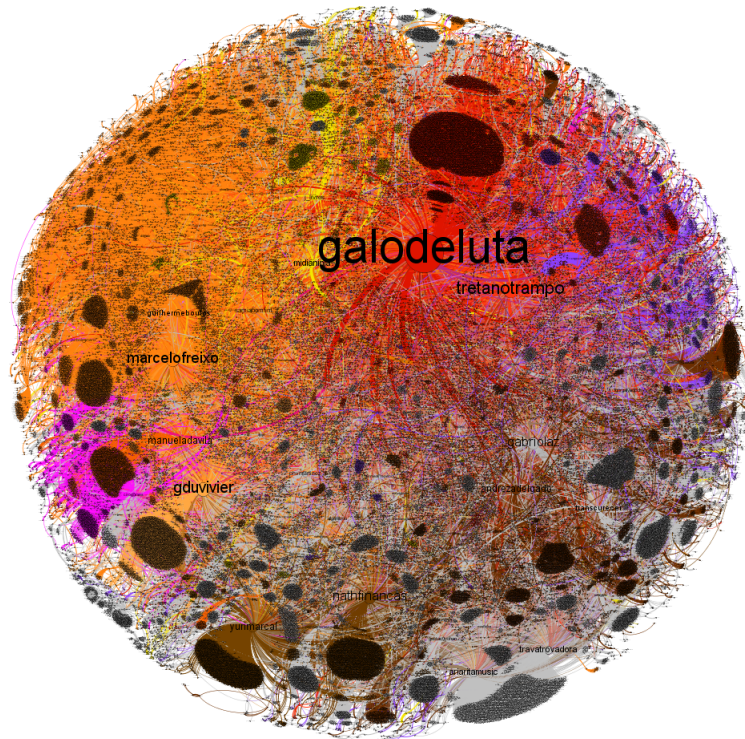


Figura 4.6: Representação visual do Componente CG₂.

Para entender quais foram os principais responsáveis pela transferência de informação entre as comunidades do Componente CG₂, foi gerada a Tabela 4.13 que exibe os dez vértices com maior valor de centralidade de intermediação.

Tabela 4.13: Vértices com os dez maiores valores de centralidade de intermediação no Componente CG₂

Vértice	Centralidade de intermediação	Grau de entrada	Grau de saída	Centralidade de grau
tretanotrampo	0,0229	6.239	160	6.399
galodeluta	0,0189	17.270	48	17.318
luladilma5	0,0096	295	140	435
silmarq	0,0061	20	191	311
articularedes	0,0031	338	34	372
manueladavila	0,0030	3.693	1	3.694
transcurecer	0,0028	2.533	40	2.573
lelecoutz	0,0026	61	23	84
izabellula	0,0025	15	416	431
gabriolaz	0,0025	4.753	5	4.758

A Tabela 4.14 mostra um resumo das métricas de grafo para o Grafo G_2 e para o Componente CG_2 .

Tabela 4.14: Resumo das métricas de grafo para o Grafo G_2 e o Componente CG_2 .

	Grafo G_2	Componente CG_2
Vértices	83.244	81.908
Arestas	185.365	184.500
Centralidade de grau mínima	1	1
Centralidade de grau máxima	17.318	17.318
Centralidade de grau média	4,4535	4,5051
Coefficiente de agrupamento médio	0,0674	0,1661
Densidade	0,00003	0,00003
Comunidades	-	45

4.2.5 Análise dos Resultados

Para o Componente CG_2 é possível observar, na Figura 4.5, distribuição de cauda longa para centralidades de grau, como a que foi vista para o Componente CG_1 (Figura 4.2). Essa distribuição é a esperada para redes reais.

É possível notar na Tabela 4.9 que há um vértice com grau de entrada muito superior aos demais, o ativista Paulo Galo (*galodeluta*), indicando que esse vértice foi um ator central nas manifestações online. Paulo Galo e o segundo vértice em grau de entrada, Treta no Trampo (*tretanotrampo*), são entregadores de aplicativos que ajudaram a organizar os movimentos de entregadores de aplicativos e promoveram as greves dos entregadores. Com exceção do ator e apresentador de TV Gregório Duvivier (*gduvivier*), do deputado federal Marcelo Freixo (*marcelofreixo*) e da página de jornalismo independente Mídia Ninja (*midianinja*), os demais vértices são usuários de pessoas negras verificadas pelo Twitter, como *andrezadelgado* e *yurimarc*. Essa situação demonstra uma relação próxima entre a causa da categoria dos entregadores e o ativismo negro no Brasil.

Sobre os vértices com os maiores valores de grau de saída, nota-se, a partir da análise dos vértices listados na Tabela 4.10, que todos são de páginas de civis identificados com pensamentos de esquerda. Também pode ser observado, na Tabela 4.11, que 90% dos vértices com os valores de centralidade de autovetor mais altos, possuindo assim maior importância na rede do Componente CG_2 , são referentes aos entregadores de aplicativos, representantes do movimento negro ou perfis de pessoas que se identificam com a esquerda. É importante observar essa circunstância de que nenhum representante político ou civil da direita figura entre os vértices com maior grau de entrada, maior grau de saída ou maior centralidade de autovetor. Essa informação indica que a Breque dos Apps foi

uma manifestação organizada pelos movimentos de entregadores de aplicativos, apoiada por diversos setores da esquerda brasileira e não sofreu nenhum dissenso relevante.

A partir dos dados sobre distribuição da centralidade de grau, grau de entrada, grau de saída e centralidade de autovetor, apresentados nas Tabelas 4.9, 4.10 e 4.11, é possível considerar que o fluxo de informação nas manifestações sobre a Breque dos Apps inicia-se a partir de grupos ou pessoas de esquerda e destina-se, diretamente ou indiretamente, aos entregadores e seus apoiadores de esquerda mais conhecidos no Twitter.

Observando-se os valores de coeficiente de agrupamento médio (0,0003) e densidade (0,1661), apresentados na Seção 4.2.4, é possível avaliar que os vértices do Componente CG_2 não estabeleceram conexões estreitas entre si. Associando esses dados aos dados obtidos de grau de entrada, grau de saída, centralidade de autovetor e de centralidade de proximidade para os vértices no Componente CG_2 , entende-se que os vértices nesse componente buscaram apoiar vértices com mais relevância no debate e não trocaram informações entre eles.

A partir da Figura 4.6, nota-se que as comunidades que representam os entregadores dos aplicativos e mídias como, por exemplo, *galodeluta* e *midianinja*, estão mais ao centro, devido a sua centralidade em relação a todo o Componente CG_2 , nas cores vermelho e amarelo. A comunidade marrom, referente ao movimento negro, encontra-se bem espalhada e misturada a outras comunidades, demonstrando visualmente a militância e divulgação de seus representantes nessa manifestação. As comunidades de cor laranja e rosa, que englobam os perfis de pessoas e grupos de esquerda em apoio à greve dos entregadores, compõem o restante do gráfico. Observando-se os dados da Tabela 4.13, nota-se por vértices como, por exemplo, *transcurecer* e *gabrielaz*, a importância dos vértices dos representantes dos entregadores e do movimento negro nessa manifestação: foram centrais no debate, próximos aos demais vértices e intermediadores entre diversas comunidades de apoiadores.

4.3 Manifestações Sobre os Movimentos #Somos70PorCento e #Somos57MILHOES

4.3.1 Contexto

No dia 28/05/2020, foi divulgada uma pesquisa Datafolha que apresentava a avaliação popular sobre o governo do presidente Jair Bolsonaro [40]. A pesquisa Datafolha apontou que o governo era aprovado por 33% da população. Nesse contexto, Eduardo Moreira —

empresário, engenheiro, escritor e palestrante carioca reconhecido por seu conhecimento econômico — fez uma reflexão sobre as pessoas que reprovavam o governo federal, que em números arredondados equivaleriam a 70% da população, após um comentário de um seguidor durante uma transmissão ao vivo em seu canal no YouTube [41]. A partir desse momento, Eduardo Moreira lança o movimento #Somos70PorCento em oposição ao presidente Bolsonaro. Durante os dias seguintes, o movimento #Somos70PorCento figurou entre os Assuntos do Momento no Twitter e recebeu o apoio de várias personalidades [41].

Por efeito da grande repercussão do movimento #Somos70PorCento em oposição a Jair Bolsonaro, os apoiadores do presidente reagiram com a hashtag #Somos57MILHOES. A hashtag faz alusão ao número de votos que Bolsonaro recebeu no segundo turno da eleição presidencial de 2018: 57,8 milhões de votos [42].

De modo a contextualizar o debate em torno das hashtags #Somos70PorCento e #Somos57MILHOES, foi gerada uma lista dos dez tweets com mais interações³ a partir dos tweets extraídos nas coletas realizadas para essa análise que pode ser visualizada no Apêndice D. Entre os tweets com mais interações, quatro tweets são em apoio ao movimento #Somos70PorCento e seis são em apoio o movimento #Somos57MILHOES. Os perfis dos autores dos tweets são variados: políticos, blogueiros, influenciadores virtuais, empresários e artistas.

4.3.2 Coletas de Dados

Os termos e hashtags utilizados para a coleta de tweets relacionados a essas manifestações online, como também suas respectivas datas de coletas, realizadas no período entre 29/05/2020 e 03/06/2020, estão descritos na Tabela 4.15, em que a primeira coluna (#) mostra o número de identificação da coleta, a segunda coluna indica o período selecionado para a coleta, e a terceira coluna relaciona os termos e hashtags observados nas publicações.

4.3.3 Apresentação do Grafo de Menções G_3

A partir do conjunto de dados obtido pelas coletas exibidas na Tabela 4.15, foi gerado o Grafo de Menções G_3 relacionado aos movimentos #Somos70PorCento e #Somos57MILHOES. O Grafo G_3 possui 177.572 vértices e 1.167.922 arestas únicas. Contudo, o Grafo G_3 não é fortemente conexo e possui 890 componentes que não são for-

³Para esta listagem, foram consideradas como interações retweets, respostas e comentários.

Tabela 4.15: Descrição das coletas de tweets relacionados aos movimentos #Somos70PorCento e #Somos57MILHOES.

#	Período de publicação	Termos e hashtags observados
1	De 29/05/2020 até 31/05/2020	#Somos70porcento e #Somos57MILHOES
2	De 30/05/2020 até 01/06/2020	#Somos70porcento, #somos70milhoes, #Somos70porcento, #Somos57MILHOES e #Somos57MLHOES
3	De 31/05/2020 até 02/06/2020	#Somos70porcento, #Somos70porcento, #70PorCentoContraOFascismo, #Somos57MILHOES e #Somos57MLHOES
4	De 01/06/2020 até 03/06/2020	#Somos70porcento, #Somos70porcento e #Somos57MILHOES

temente conexos. Portanto, é necessário observar o tamanho dos componentes conexos do Grafo G_3 para selecionar quais componentes analisar mais detalhadamente. Com esse objetivo, um gráfico foi gerado (Figura 4.7) para esclarecer como é a distribuição de tamanho dos componentes conexos do Grafo G_3 e possibilitar a escolha de quais desses componentes serão analisados.

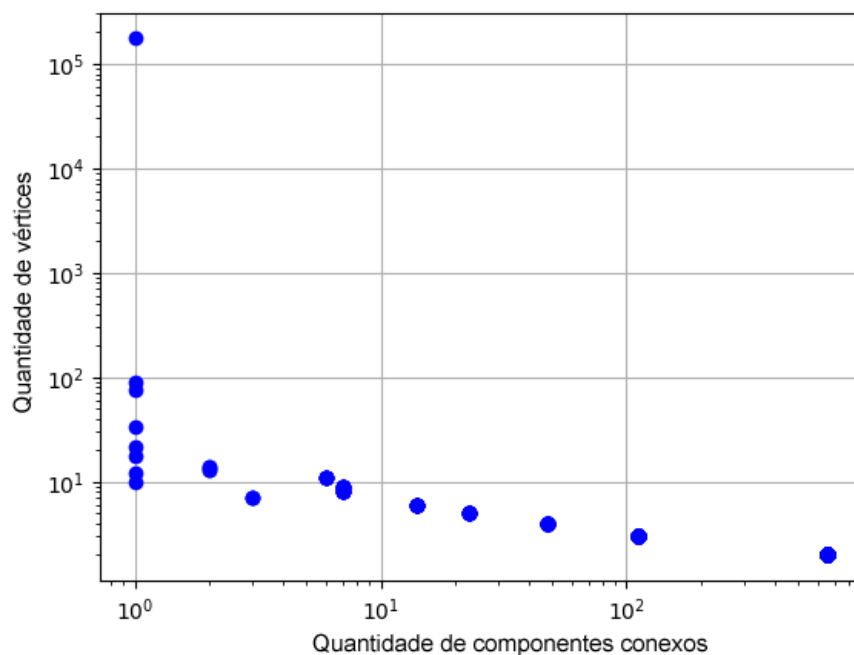


Figura 4.7: Distribuição do tamanho dos componentes no Grafo G_3 (eixos ajustados para escala logarítmica).

Dada a Figura 4.7, nota-se que há no Grafo G_3 um componente conexo com uma quantidade superior de vértices que os demais componentes. Por este motivo, a análise prosseguirá sobre o maior componente conexo do Grafo G_3 , que será identificado

como CG_3 . Apresentam-se as características do componente conexo a ser analisado: são 175.010 vértices e 1.166.060 arestas presentes no Componente CG_3 .

4.3.4 Apresentação das Métricas Aplicadas ao Componente Conexo CG_3

A centralidade de grau máximo de um vértice do Componente CG_3 é 25.823, a centralidade de grau mínimo de um vértice do Componente CG_3 é 1 e a centralidade de grau médio dos vértices do Componente CG_3 é 13,3256. Para melhor visualização da distribuição dos valores de centralidade de grau, foi gerado um histograma desses valores normalizados pelo número de vértices presentes no Componente CG_3 (Figura 4.8). Para esse histograma, os valores do eixo x foram normalizados dividindo-se o valor da centralidade de grau do vértice pelo grau máximo possível, dado por $n - 1$, em um grafo não-direcionado obtido a partir de CG_3 , onde n é o número de vértices em CG_3 ; e os valores do eixo y foram ajustados para escala logarítmica.

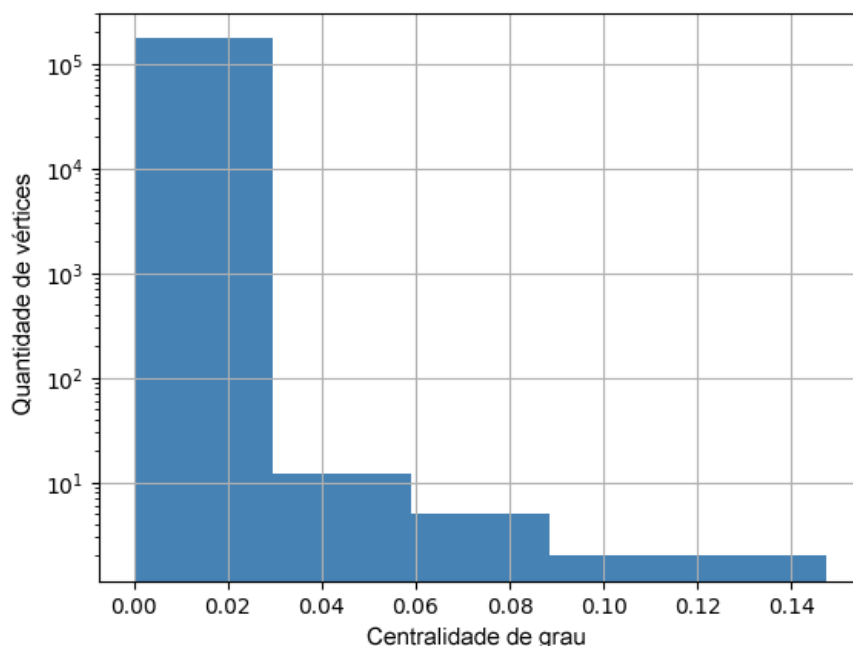


Figura 4.8: Distribuição da centralidade de grau no Componente CG_3 .

Em sequência, é necessário entender quais são os perfis dos vértices com os maiores valores de grau de entrada e dos vértices com os maiores valores de grau de saída. Com esse objetivo, foram geradas duas listas: uma lista apresentando os dez vértices com maiores valores de grau de entrada (Tabela 4.16) no Componente CG_3 e outra lista contendo os dez vértices com maior valor de grau de saída (Tabela 4.17) no Componente CG_3 .

Para entender melhor a relação dos vértices mais centrais apontada anteriormente,

Tabela 4.16: Vértices com os dez maiores valores de grau de entrada no Componente CG₃.

Vértice	Grau de entrada	Grau de saída	Centralidade de grau
jairbolsonaro	25.823	0	25.823
carlazambelli38	22.097	2	22.099
flaviobolsonaro	16.900	7	16.907
bernardopkuster	16.633	0	16.633
eduardomoreira	14.969	19	14.988
jouberth19	13.488	6	13.494
carlosjordy	12.104	2	12.106
taoquei1	11.600	8	11.608
bolsonarosp	10.402	1	10.403
marcelofreixo	9.768	1	9.769

Tabela 4.17: Vértices com os dez maiores valores de grau de saída no Componente CG₃.

Vértice	Grau de saída	Grau de entrada	Centralidade de grau
vivianerecife	1.456	162	1.618
kenpachi_11cap	1.140	8	1.148
joelmadsilva	1.052	5	1.057
monteiro_marize	1.024	32	1.056
leihigor	984	18	1.002
fake_gdo	951	8	959
josrenato14	918	5	923
jucidarcn	890	57	947
bozo_bot	856	8	864
leandrovba	802	7	809

foram calculados as centralidades de autovetor dos vértices presentes no Componente CG_3 . Os dez vértices com os maiores valores de centralidade de autovetor estão expostos na Tabela 4.18.

Tabela 4.18: Vértices com os dez maiores valores de centralidade de autovetor no Componente CG_3 .

Vértice	Centralidade de autovetor	Grau de entrada	Grau de saída	Centralidade de grau
eduardomoreira	0,2231	14.969	19	14.988
jairbolsonaro	0,2177	25.823	0	25.823
marcelofreixo	0,1481	9.768	1	9.769
guilhermeboulos	0,1479	7.068	0	7.068
senadorhumberto	0,1404	5.460	1	5.461
esquerdapt13	0,1318	382	109	491
bacelarbatista	0,1304	338	215	553
gutsesquerda	0,1304	534	126	660
horaora2	0,1241	275	358	633
bolsonarosp	0,1209	10.402	1	10.403

Para avaliar a troca de informações entre os vértices no Componente CG_3 , foi gerada a listagem dos dez maiores valores de centralidade de proximidade no Componente CG_3 que indica quais vértices tiveram suas publicações distribuídas com mais facilidade na rede em estudo. Essa listagem está exibida na Tabela 4.19.

Tabela 4.19: Vértices com os dez maiores valores de centralidade de proximidade no Componente CG_3 .

Vértice	Centralidade de proximidade	Grau de entrada	Grau de saída	Centralidade de grau
vivianerecife	0,0818	162	1.456	1.618
kenpachi_11cap	0,0804	8	1.140	1.148
leandrovba	0,0802	7	802	809
ferrazclaudione	0,0786	3	787	790
fake_gdo	0,0783	8	951	959
solalinancy	0,0781	37	588	625
monteiro_marize	0,0778	32	1.024	1.056
joelmadsilva	0,0778	5	1.052	1.057
leihigor	0,0778	18	984	1.002
josrenato14	0,0778	5	918	923

No Componente CG_3 , o coeficiente de agrupamento médio é 0,1153 e a densidade do componente é 0,00004. Como o Componente CG_3 não é fortemente conexo, não foi possível calcular a distância média e diâmetro.

Para analisar como o Componente CG_3 ficou dividido em comunidades, foi aplicado,

por meio do Gephi, o algoritmo Louvain no Componente CG_3 . Foram obtidas 264 comunidades a partir desse algoritmo. A Figura 4.9 mostra a representação visual do resultado obtido pelo algoritmo executado. Os vértices têm seus identificadores exibidos com tamanho de fonte proporcional às suas centralidades de grau, e quanto mais centrais ao grafo, mais próximos ao centro do círculo são posicionados. As comunidades são posicionadas próximas às comunidades que tiveram mais relações. As principais comunidades detectadas podem ser identificadas na Figura 4.9 a partir das cores com as definições dadas a seguir. As definições foram inferidas a partir da avaliação de amostras dos vértices atribuídos às comunidades identificadas pelo algoritmo Louvain.

1. Vermelho, laranja e amarelo: comunidades do movimento #Somos70PorCento.
2. Verde e azul: comunidades do movimento #Somos57MILHOES.

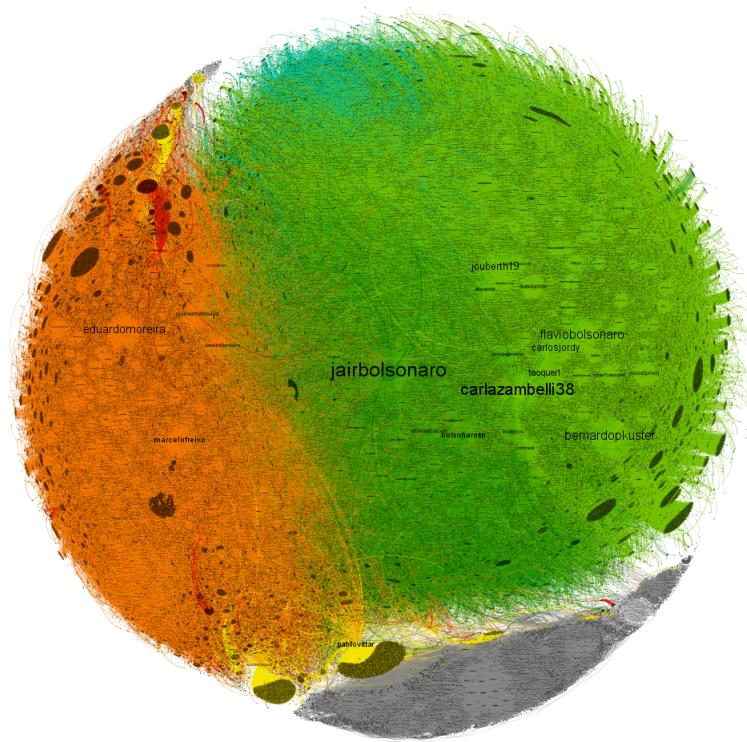


Figura 4.9: Representação visual do Componente CG_3 .

Por fim, para entender quais foram os principais responsáveis pela transferência de informação entre as comunidades do Componente CG_3 , foi gerada a Tabela 4.20 que exhibe os dez vértices com maior valor de centralidade de intermediação.

A Tabela 4.21 mostra um resumo das métricas de grafo para o Grafo G_3 e para o Componente CG_3 .

Tabela 4.20: Vértices com os dez maiores valores de centralidade de intermediação no Componente CG₃

Vértice	Centralidade de intermediação	Grau de entrada	Grau de saída	Centralidade de grau
allantercalivre	0,0264	6.513	8	6.521
eduardomoreira	0,0230	14.969	19	14.988
marialciafigue1	0,0195	380	107	487
carlazambelli38	0,0189	22.097	2	22.099
elsonloula	0,0156	211	93	304
vivianerecife	0,0152	162	1.456	1.618
fora_fascistas	0,0147	868	85	953
flaviobolsonaro	0,0127	16.900	7	16.907
joomikhail	0,0101	1.114	420	1534
taoquei1	0,0087	11.600	8	11.608

Tabela 4.21: Resumo das métricas de grafo para o Grafo G₃ e o Componente CG₃.

	Grafo G₃	Componente CG₃
Vértices	177.572	175.010
Arestas	1.167.922	1.166.060
Centralidade de grau mínima	1	1
Centralidade de grau máxima	25.823	25.823
Centralidade de grau média	13,1543	13,3256
Coeficiente de agrupamento médio	0,0740	0,1153
Densidade	0,00004	0,00004
Comunidades	-	264

4.3.5 Análise dos Resultados

O Componente CG₃ apresenta média de centralidade de grau elevada (13,3256) em comparação às manifestações analisadas anteriormente. Os dados apresentados sobre centralidade de grau somados ao número de arestas (1.166.060 arestas) indicam também que houve um forte debate nas manifestações em análise no Componente CG₃, sendo esse comportamento evidenciado pela quantidade de conexões formadas devido às menções. Além disso, o Componente CG₃ apresenta uma distribuição de centralidade de grau semelhante à cauda-longa, como pode ser observado na Figura 4.8, algo esperado para redes reais.

Observando-se a Tabela 4.16, nota-se que o vértice mais citado é o do presidente Jair Bolsonaro (*jairbolsonaro*), algo esperado pela natureza dos movimentos analisados de oposição e apoio ao Governo Federal. Dos nove vértices restantes para essa métrica, sete representam aliados de Jair Bolsonaro, incluindo os dois filhos dele — Eduardo (*bolsonarosp*) e Flávio Bolsonaro (*flaviobolsonaro*) — e os deputados federais Carla Zambelli (*carlazambelli38*) e Carlos Jordy (*carlosjordy*). Apenas dois vértices são de opositores ao presidente Bolsonaro: um referente ao empresário Eduardo Moreira (*eduardomoreira*) e outro referente ao deputado federal Marcelo Freixo (*marcelofreixo*). Essa informação aponta que a manifestação foi controlada nas comunidades dos apoiadores do Governo Federal pela cúpula dos aliados de Jair Bolsonaro; enquanto a oposição se organizou em torno de poucos vértices, em especial o vértice do organizador do #Somos70PorCento.

Avaliando-se a Tabela 4.17, percebe-se altos valores de grau de saída em comparação às análises anteriores. Quase todos os vértices são de apoiadores de Jair Bolsonaro, com exceção do *bozo_bot*, que é uma conta automatizada para compartilhar tweets que contenham “Fora Bolsonaro” no texto. Ademais, todos os vértices presentes nessa classificação são de usuários comuns, que não possuem verificação da plataforma ou grande números de seguidores para serem considerados “influenciadores”. Assim, é possível observar que houve um forte engajamento nas manifestações por parte dos apoiadores do presidente. Um engajamento tão significativo que superou, em grau de saída, uma conta automatizada, a *bozo_bot*.

Pode ser observado na Tabela 4.18 que oito vértices mais relevantes, de acordo com os valores de centralidade de autovetor, são relacionados à oposição ao Governo Federal. Entre esses vértices de oposição estão: Eduardo Moreira, Marcelo Freixo, o deputado federal Humberto Costa (*senadorhumberto*) e o candidato à presidência em 2018 e à prefeitura de São Paulo em 2020 pelo Partido Socialismo e Liberdade (PSOL), Guilherme Boulos (*guilhermeboulos*). Dos vértices restantes na lista apresentada na Tabela 4.18,

apenas dois são representantes de apoio ao governo — o próprio presidente Jair Bolsonaro e um dos seus filhos, o deputado federal Eduardo Bolsonaro.

A partir das informações sobre grau de entrada, grau de saída e centralidade de autovetor avaliadas anteriormente, nota-se que os fluxos de informação diferiram entre as comunidades de oposição e as comunidades de apoio ao Governo Federal, do Componente CG₃. A comunidade de apoio ao presidente Bolsonaro mostra ter se concentrado em torno de vértices mais influentes e compartilhado muito o discurso desses vértices, como uma rede descentralizada, e, por consequência, tornando os vértices mais centrais. Por outro lado, na comunidade de oposição ao governo, o debate mostra ter tido mais interação entre todos os vértices, como uma rede distribuída, sem produzir vértices muito mais centrais que os demais. Essas relações são destacadas ao se avaliar a centralidade de autovetor, uma vez que a rede distribuída produzida pelas comunidades de oposição ao Governo Federal dispersou melhor o peso entre os vértices e, dessa forma, acabou agregando mais relevância aos vértices com mais conexões; ao contrário do que aconteceu na rede descentralizada produzida pelas comunidades de apoio ao Governo Federal, em que vértices se concentraram em torno de poucos vértices e a distribuição de peso não aconteceu tão eficientemente.

É muito importante observar na Tabela 4.20 os baixos valores de centralidade de intermediação. Essa informação indica que houve pouca troca de informação entre as comunidades, portanto, os discursos foram mantidos dentro das suas comunidades. Isso reflete a natureza de disputa dos movimentos observados. Pode-se observar também que 80% dos vértices presentes na listagem de grau de saída (Tabela 4.17) se repetem na listagem de centralidade de proximidade (Tabela 4.19).

A partir da Figura 4.9, podem ser observadas as comunidades de oposição ao Governo Federal — em laranja, vermelho e amarelo — e as comunidades de apoio ao Governo Federal — em tons de verde e azul. Nota-se que o vértice do presidente Bolsonaro é o maior e está bem ao centro do gráfico devido a sua centralidade em relação a todo o Componente CG₃. As comunidades de apoio ao governo ocupam a maior parte do gráfico (44,89% do total de vértices) e possuem mais relevância no grafo devido aos vértices de maiores centralidades de grau como, por exemplo, os vértices da deputada bolsonarista Carla Zambelli (*carlazambelli38*) e do senador — filho do presidente — Flávio Bolsonaro (*flaviobolsonaro*). As comunidades de oposição representam uma área menor no gráfico (37,75% do total de vértices) com poucos vértices entre as maiores centralidades de grau, observado anteriormente nas Tabelas 4.16 e 4.17. É possível também ver que as comunidades de apoio e oposição ao Governo Federal não se misturam no gráfico, conforme o que foi evidenciado pelos valores de centralidades de intermediação na Tabela 4.20.

Apesar da pesquisa Datafolha [40], que motivou o movimento #Somos70PorCento, apontar que apenas 33% das pessoas aprovavam o governo Bolsonaro, esta proporção não se refletiu na disputa no Twitter entre os movimentos #Somos70PorCento e #Somos57MILHOES. Importante destacar que essa iniciativa do #Somos70PorCento não partiu de movimentos organizados da oposição a Jair Bolsonaro, mas sim de uma comunidade específica em torno de seguidores do Eduardo Moreira. Tal fato é relevante, pois nem todos os opositores do governo aderiram ao debate no Twitter. Foi um debate que teve suas idiossincrasias, comparado com os demais analisados na dissertação, mais abrangentes. Por outro lado, os apoiadores do governo aderiram em peso ao debate lançando a #Somos57MILHOES como resposta defensiva aos #Somos70PorCento. Consequentemente, as comunidades de apoio ao governo concentraram mais vértices que as comunidades de oposição no Componente CG₃. Entretanto, é importante ressaltar o comportamento extremamente engajado dos vértices de apoio ao presidente Jair Bolsonaro, aparente na Tabela 4.17, que se associa ao de contas automáticas, como o usuário *bozo_bot*. Como não foi realizada para essa pesquisa a avaliação de autenticidade dos vértices nos grafos analisados, não é possível afirmar que houve a presença de contas automatizadas nas manifestações do #Somos57MILHOES, mas estas manifestações apresentam um cenário propício para a busca por inflação de hashtags, criação de tendências artificiais, além da propaganda política de aprovação do governo de Jair Bolsonaro — táticas características para o uso de contas automatizadas em plataformas de redes sociais [43].

4.4 Manifestações Sobre o Auxílio Emergencial

4.4.1 Contexto

Com a chegada da COVID-19 ao Brasil, medidas sanitárias foram estabelecidas pelos governadores para prevenção e combate à doença nos estados brasileiros. A restrição de circulação foi uma das principais decisões tomadas pelos governos estaduais [44]. Devido a essa medida, muitos trabalhadores autônomos e comerciantes ficaram sem possibilidade de trabalhar e gerar renda durante a prevenção ao COVID-19, portanto foi necessário buscar uma solução na política para ajudar essas pessoas.

Buscando minimizar os impactos econômicos e sociais que seriam causados pela pandemia do COVID-19, a Câmara dos Deputados aprovou, em 26/03/2020, o pagamento de um auxílio emergencial no valor de R\$600,00 às pessoas de baixa renda por três meses, sendo prorrogável pelo Poder Executivo Federal enquanto durar a pandemia [45]. O valor aprovado para o auxílio emergencial pela Câmara dos Deputados é três vezes maior que o

inicialmente sugerido pelo Governo Federal (R\$200,00) [46], indicando uma derrota para a presidência de Jair Bolsonaro. Para o recebimento desse auxílio, foi estabelecido que o pagamento seria feito por bancos públicos em contas dos beneficiários.

Após a sanção presidencial, o Governo Federal decidiu realizar o pagamento pelo banco Caixa Econômica por meio do aplicativo “Caixa Tem” para as pessoas que estivessem com o cadastro regularizado com a instituição. Entretanto, essa decisão trouxe dificuldades a quem precisava receber o auxílio, provocou filas nas agências por quem buscava regularizar sua situação cadastral, gerou filas virtuais no aplicativo para quem tentava consultar o pagamento, além atrasar muitos pagamentos provocando insatisfação da população que precisava dessa renda [47]. Esses transtornos causados por uma decisão estratégica do Governo Federal mobilizaram reclamações nas redes sociais e chegaram aos Assuntos do Momento recorrentemente.

De modo a apresentar o teor das manifestações sobre o auxílio emergencial, foi gerada uma lista dos dez tweets com mais interações⁴ a partir dos tweets extraídos nas coletas realizadas para essa análise. Essa lista pode ser visualizada no Apêndice E. Entre os tweets com mais interações, apenas cinco são referentes ao auxílio emergencial. Essa situação acontece devido à “caixa” estar entre os termos de coleta por ser o nome do banco responsável pelo pagamento do auxílio, mas também ser uma palavra comum nos textos das publicações. Entretanto, dois desses tweets com mais interações sobre o auxílio emergencial são críticas em tom de denúncias sobre o Governo Federal e o “Caixa Tem”. Outros dois tweets são informativos da Caixa sobre o pagamento do auxílio.

Ao longo desta seção, serão analisados os dados referentes às manifestações online feitas no Twitter relacionadas ao auxílio emergencial pago pelo Governo Federal.

4.4.2 Coletas de Dados

Os termos e hashtags utilizados para a coleta de tweets relacionados às manifestações sobre o auxílio emergencial, como também suas respectivas datas de coletas, realizadas no período entre 02/06/2020 e 18/08/2020, estão descritos na Tabela 4.22, em que a primeira coluna (#) mostra o número de identificação da coleta, a segunda coluna indica o período selecionado para a coleta, e a terceira coluna relaciona os termos e hashtags observados nas publicações.

Devido ao termo “caixa” presente na coleta número 5, foi retornado um tweet da ativista Luciene Santos (*sapavegana*). Entretanto, a ativista citada não participou das ma-

⁴Para esta listagem, foram consideradas como interações retweets, respostas e comentários.

nifestações em relação ao auxílio emergencial, sua presença nessa lista é devido ao tweet que contém o termo “caixa”. Como “caixa” é um dos termos principais para este assunto, por ser o nome do banco público responsável pelo pagamento do auxílio emergencial, Caixa Econômica Federal, não é possível removê-lo da coleta, portanto, a análise será desenvolvida sem considerar o vértice *sapavegana*.

Tabela 4.22: Descrição das coletas de tweets relacionados às manifestações sobre o auxílio emergencial.

#	Período de publicação	Termos e hashtags observados
1	De 02/06/2020 até 04/06/2020	#AnaliseEterna
2	De 24/06/2020 até 26/06/2020	#PagaLogoBolsonaro
3	De 25/05/2020 até 27/06/2020	#PagaLogoBolsonaro e #auxilioemergencial
4	De 26/06/2020 até 28/06/2020	#caixatem
5	De 28/06/2020 até 30/06/2020	#auxilioemergencial e caixa
6	De 30/06/2020 até 02/07/2020	#caixatem
7	De 01/07/2020 até 03/07/2020	#caixatem e #CaixaTemNAOFUN-CIONA
8	De 02/07/2020 até 04/07/2020	#CaixaTemNAOFUNCIONA, #CAIXATemNADA e #auxilioemergencial
9	De 03/07/2020 até 05/07/2020	#CAIXATemNADA
10	De 16/08/2020 até 18/08/2020	Agências da Caixa

4.4.3 Apresentação do Grafo de Menções G_4

O Grafo de Menções G_4 relacionado às manifestações sobre o auxílio emergencial foi gerado a partir do conjunto de dados obtido pelas coletas exibidas na Tabela 4.22. O Grafo G_4 possui 53.175 vértices e 85.089 arestas. O Grafo G_4 não é fortemente conexo e há 5.644 componentes que não são fortemente conexos nele. Desta maneira, faz-se necessário observar o tamanho dos componentes conexos do Grafo G_4 a fim de entender quais componentes são relevantes para serem analisados mais detalhadamente. Com esse objetivo, um gráfico foi gerado para esclarecer como é a distribuição de tamanho dos componentes conexos do Grafo G_4 e possibilitar a decisão de quais desses componentes devem ser analisados. A Figura 4.10 apresenta esse gráfico.

Na Figura 4.10 é possível observar que há no Grafo G_4 um componente conexo que concentra muito mais vértices do que os demais componentes do Grafo G_4 . Portanto, a análise prosseguirá sobre esse componente conexo do Grafo G_4 , que será identificado como CG_4 . Há 37.442 vértices e 74.241 arestas presentes no Componente CG_4 .

Tabela 4.23: Vértices com os dez maiores valores de grau de entrada no Componente CG₄.

Vértice	Grau de entrada	Grau de saída	Centralidade de grau
caixa	6.168	210	6.378
jairbolsonaro	3.117	2	3.119
sapavegana	1.493	0	1.493
davialcolumbre	1.451	0	1.451
leandrolyrarj	1.291	0	1.291
bobjeffcensored	1.159	0	1.159
motta1302	1.073	0	1.073
mincidadania	1.052	7	1.059
marcelofreixo	838	1	839
govbr	766	15	781

Tabela 4.24: Vértices com os dez maiores valores de grau de saída no Componente CG₄.

Vértice	Grau de saída	Grau de entrada	Centralidade de grau
letciaa02393358	419	23	442
oluap65905953	385	44	429
dricaaprovada	310	23	333
rickbarros15	285	5	290
caixa	210	6.168	6.378
oliver_alves1	207	44	251
araujospzsl	196	3	199
ajuderuan	193	8	201
gustavotelles90	182	57	239
danielecorso9	174	48	222

Tabela 4.25: Vértices com os dez maiores valores de centralidade de autovetor no Componente CG₄.

Vértice	Centralidade de autovetor	Grau de entrada	Grau de saída	Centralidade de grau
caixa	0,6800	6.168	210	6.378
jairbolsonaro	0,1345	3.117	2	3.119
govbr	0,1312	766	15	781
mincidadania	0,1287	1.052	7	1.059
dataprev	0,0954	558	3	561
onyxlorenzoni	0,0935	723	2	725
mineconomia	0,0607	371	3	374
lucianomaaneir1	0,0600	60	13	73
tvbrasilgov	0,0590	485	6	491
gutonds	0,0565	27	8	35

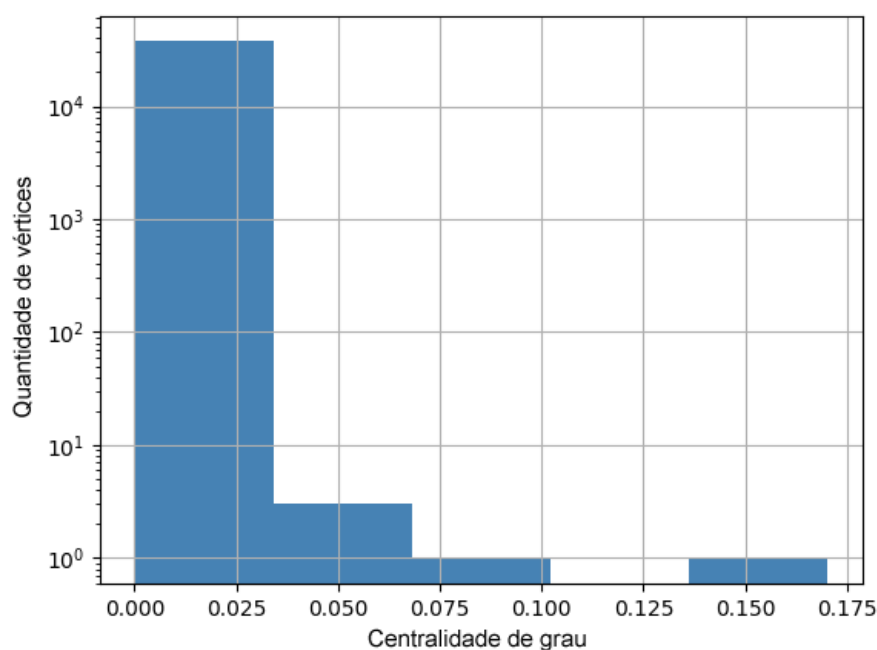


Figura 4.11: Distribuição da centralidade de grau no Componente CG₄.

a densidade é 0,00005. Não foi possível calcular a distância média e diâmetro pois o Componente CG₄ não é fortemente conexo.

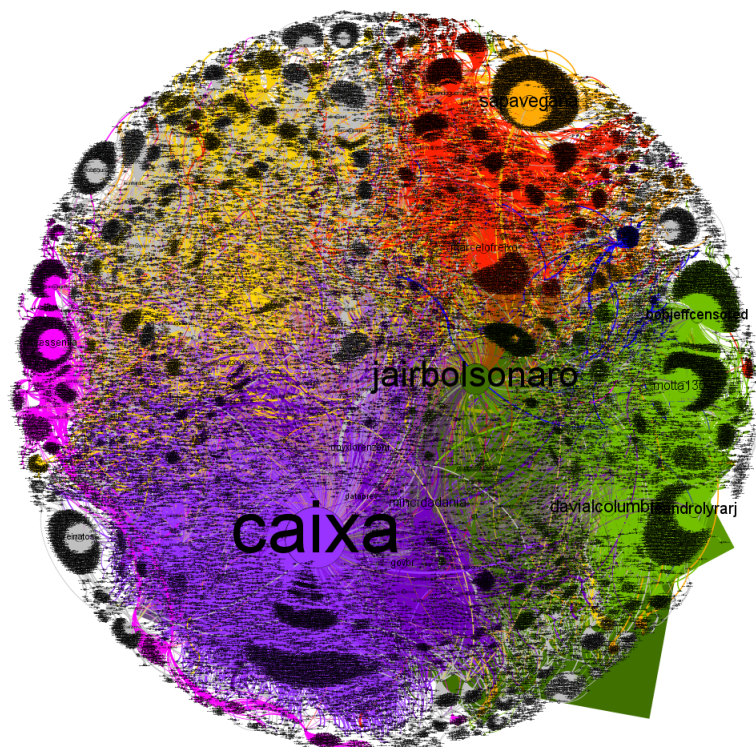
Também foi gerada a listagem dos dez maiores valores de centralidade de proximidade no Componente CG₄. Essa listagem é exibida na Tabela 4.26.

Tabela 4.26: Vértices com os dez maiores valores de centralidade de proximidade no Componente CG₄.

Vértice	Centralidade de proximidade	Grau de entrada	Grau de saída	Centralidade de grau
letciaa02393358	0,0507	23	419	442
rickbarros15	0,0492	5	285	290
araujospzsul	0,0468	3	196	199
lena_patty	0,0460	51	165	216
santos_aurizete	0,0449	47	141	188
oliver_alves1	0,0445	44	207	251
polyanny11	0,0445	5	141	146
castielamadean	0,0444	19	136	155
diegogomestt	0,0444	113	67	180
ironic_feminist	0,0444	1	82	83

Para analisar como o Componente CG₄ ficou dividido em comunidades, foi aplicado, por meio do Gephi, o algoritmo Louvain no Componente CG₄. Foram obtidas 173 comunidades a partir desse algoritmo. A Figura 4.12 mostra a representação visual do resul-

1. Roxo, vermelho, laranja, rosa e amarelo: comunidades neutras ou de oposição ao Governo Federal e em favor do auxílio emergencial e de sua prorrogação.
2. Tons de verde: comunidades referentes a apoiadores do Governo Federal ou órgãos do governo.



Para entender quais foram os principais responsáveis pela transferência de informação entre as comunidades do Componente CG_4 , foi gerada a Tabela 4.27, que exhibe os dez vértices com maior valor de centralidade de intermediação.

77

Tabela 4.27: Vértices com os dez maiores valores de centralidade de intermediação no Componente CG₄

Vértice	Centralidade de intermediação	Grau de entrada	Grau de saída	Centralidade de grau
caixa	0,0411	6.168	210	6.378
biscoitinho93	0,0236	30	100	130
letciaa02393358	0,0110	23	419	442
jairbolsonaro	0,0085	3.117	2	3.119
diegogomestt	0,0044	113	67	180
santos_aurizete	0,0042	47	141	188
dixxvitoriaa	0,0038	9	4	13
rickbarros15	0,0038	5	285	290
antony7717	0,0034	10	80	90
janeb2021	0,0034	55	138	193

Tabela 4.28: Resumo das métricas de grafo para o Grafo G₄ e o Componente CG₄.

	Grafo G₄	Componente CG₄
Vértices	53.175	37.442
Arestas	85.089	74.241
Centralidade de grau mínima	1	1
Centralidade de grau máxima	6.378	6.378
Centralidade de grau média	3,2003	3,9657
Coeficiente de agrupamento médio	0,0627	0,1988
Densidade	0,00003	0,00005
Comunidades	-	173

4.4.5 Análise dos Resultados

Como nos grafos analisados anteriormente, nota-se, a partir da Figura 4.11, que a maior parte dos vértices presentes no Componente CG_4 não são a origem ou o destino da maioria das arestas existentes, formando, nesse caso, uma distribuição próxima à cauda longa.

Na Tabela 4.23, observa-se que os vértices com maiores valores de grau de entrada são relacionados a representantes ou órgãos do Governo Federal e seus apoiadores, como os usuários do Ministério da Cidadania (*mincidadania*) e do ex-deputado federal de extrema-direita, investigado pela Polícia Federal em inquérito sobre milícias digitais, Roberto Jefferson (*bobjeffcensored*) [48, 49]. A oposição fica representada apenas pelo vértice do deputado federal Marcelo Freixo (*marcelofreixo*). É possível afirmar que parte das manifestações se deram em torno de perfis responsáveis pela aprovação, organização e distribuição do pagamento do auxílio emergencial.

A partir da Tabela 4.24, percebe-se que, em geral, o perfil dos vértices que foram origem das conexões no Componente CG_4 é de quem necessitou em algum momento do auxílio emergencial, como *letciaa02393358* e *oluap65905953*. Essa percepção é dada a partir da análise dos perfis dos usuários: 50% dos vértices da Tabela 4.24 são de contas no Twitter criadas entre abril de 2020 e junho de 2020, e seus tweets são referentes ao auxílio emergencial. Portanto, a partir das informações anteriores, é possível concluir que cidadãos que precisavam do auxílio para se sustentar criaram contas no Twitter para cobrá-lo e, muitas vezes, as contas foram usadas com essa única finalidade. Os demais vértices com maiores valores de grau de saída são: a Caixa Econômica Federal (*caixa*) e três concursados da Caixa reivindicando suas convocações (*oluap65905953*, *dricaaprovada* e *danielecorso9*). O vértice da Caixa Econômica provavelmente possui o grau de saída elevado por conta das respostas às pessoas que o cobravam ajuda no Twitter.

Na Tabela 4.25, pode ser observado que o vértice com o maior valor de centralidade de autovetor, um valor muito maior do que o dos demais vértices, é o da Caixa Econômica Federal. Ademais, 89% dos outros vértices são de autoridades ou órgãos do Governo Federal, como os usuários do presidente Jair Bolsonaro (*jairbolsonaro*) e do Ministério da Economia (*mineconomia*), reforçando o caráter de cobrança das manifestações feitas.

Os valores do coeficiente de agrupamento médio e densidade do grafo, 0,1988 e 0,00005 respectivamente, apresentados na Seção 4.4.4, evidenciam que os vértices do Componente CG_4 não estabeleceram conexões estreitas entre si.

Reunindo as informações obtidas sobre grau de entrada, grau de saída e centrali-

dade de autovetor, coeficiente de agrupamento médio e densidade no Componente CG_4 , conclui-se que os vértices representando pessoas que precisaram do auxílio emergencial nesse componente mencionaram algumas autoridades e órgãos relacionados ao Governo Federal em cobrança pelo pagamento do auxílio emergencial e interagiram pouco entre eles mesmos. A oposição e apoiadores do Governo Federal tiveram pouca participação nessas manifestações.

Os vértices presentes na listagem da Tabela 4.26 possuem valores baixos de centralidade de proximidade. Tal situação pode ter ocorrido pela pouca interação devido à baixa conexão entre os vértices apontada no Componente CG_4 .

Na Tabela 4.27, observa-se, mais uma vez, a relevância do vértice da Caixa Econômica Federal no Componente CG_4 . Os demais vértices possuem os valores de centralidade de intermediação muito menores do que o da Caixa. Essa informação demonstra que as conexões se deram em torno do vértice da Caixa Econômica Federal, dado que é o banco responsável pelo pagamento do auxílio emergencial que parte da população, representado pelos outros vértices, estava precisando.

A partir da Figura 4.12, pode ser observado que as comunidades no Componente CG_4 se organizaram mais em torno dos vértices referentes aos órgãos públicos e autoridades federais. A maior comunidade, em roxo, relaciona-se com o vértice da Caixa Econômica Federal. As comunidades próximas aos apoiadores e oposição ao Governo Federal ocupam pouco espaço e ficam mais afastadas do centro do gráfico. Além das comunidades se misturarem pouco entre elas, o que demonstra os baixos valores de centralidades de proximidade e intermediação vistos para esse componente conexo.

4.5 Manifestações Sobre a Realização do ENEM 2020

4.5.1 Contexto

O Exame Nacional do Ensino Médio (ENEM) é um exame aplicado anualmente no Brasil desde 1998. Inicialmente com o objetivo de avaliar o desempenho escolar dos estudantes da educação básica, o ENEM passou a ser utilizado, a partir de 2009, como mecanismo de ingresso à educação superior [50]. Atualmente, o resultado obtido nesse exame permite ao estudante concorrer a uma vaga em cursos de graduação em instituições de ensino superior por meio do Sistema de Seleção Unificada (SISU), do Programa Universidade para Todos (PROUNI) e de convênios com instituições portuguesas, como também solicitar financiamento estudantil pelo Fundo de Financiamento Estudantil (FIES).

O ENEM, realizado durante dois dias a cada ano, é a maior avaliação educacional do Brasil [51].

Contudo, com a pandemia da COVID-19, preocupações surgiram acerca da segurança sanitária para realização da prova, além da desigualdade ainda maior entre estudantes de escolas públicas e privadas devido às condições de estudo durante o isolamento social. Diante de muitas incertezas sobre o ENEM 2020, manifestações foram feitas nas redes sociais cobrando uma definição do Ministério da Educação e da instituição responsável pela realização da prova, Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), sobre o ENEM 2020.

Depois de alguns meses de indefinição sobre o ENEM 2020, o Ministério da Educação com o INEP definiram as novas datas para janeiro e fevereiro de 2021 e regras da prova [52].

Ao longo dessa seção, serão analisadas as manifestações ocorridas no Twitter relacionadas ao ENEM 2020. Com a finalidade de contextualizar a análise dos dados, foi gerada uma lista dos dez tweets das manifestações sobre o ENEM 2020 com mais interações⁵. Os tweets listados fazem parte do conjunto de dados extraídos para essa análise. Essa lista pode ser visualizada no Apêndice F. Metade dos tweets com mais interações são notícias sobre o ENEM 2020, informando aos leitores atualizações sobre a prova a ser realizada na época. Entre esses tweets, dois deles foram publicados por veículos de notícias: o portal G1 e o jornal Estado de SP. Dos tweets restantes com mais interações, quatro são críticas ao Governo Federal e ao Ministério da Educação sobre a condução e divulgação do ENEM 2020.

4.5.2 Coletas de Dados

Os termos e hashtags utilizados para a coleta de tweets relacionados a esse debate, como também suas respectivas datas de coletas, realizadas no período entre 05/06/2020 e 02/08/2020, são apresentados na Tabela 4.29, em que a primeira coluna (#) mostra o número de identificação da coleta, a segunda coluna indica o período selecionado para a coleta, e a terceira coluna relaciona os termos e hashtags observados nas publicações.

Excepcionalmente, foi aplicada na coleta número 3 uma restrição para apenas tweets em língua portuguesa, pois o termo MEC, que se refere ao Ministério da Educação, é também uma expressão francesa que significa “cara” (no sentido de “rapaz”), portanto, tweets não relacionados ao ENEM estavam sendo retornados para essa coleta quando não

⁵Para esta listagem, foram consideradas como interações retweets, respostas e comentários.

aplicada tal restrição.

Tabela 4.29: Descrição das coletas de tweets relacionados às manifestações sobre o ENEM 2020.

#	Período de publicação	Termos e hashtags observados
1	De 05/06/2020 até 07/06/2020	Enem
2	De 30/06/2020 até 02/07/2020	Enem, Inep e Sisu
3	De 07/07/2020 até 09/07/2020	Enem, Inep e MEC
4	De 08/07/2020 até 10/07/2020	Enem e Inep
5	De 30/07/2020 até 01/08/2020	#CANCELAENEM
6	De 31/07/2020 até 02/08/2020	#CANCELAENEM

4.5.3 Apresentação do Grafo de Menções G_5

A partir do conjunto de dados obtido pelas coletas relacionadas às manifestações sobre o ENEM 2020 exibidos na Tabela 4.29, um Grafo de Menções G_5 foi modelado. O Grafo G_5 possui 135.695 vértices e 194.759 arestas únicas. O Grafo G_5 não é fortemente conexo e possui 8.204 componentes que não são fortemente conexos. De tal maneira, é necessário avaliar os componentes conexos do Grafo G_5 a fim de se entender quais componentes são relevantes para serem analisados mais a fundo. Com esse objetivo, um gráfico foi gerado para esclarecer como é a distribuição de tamanho dos componentes conexos do Grafo G_5 e possibilitar definir quais desses componentes deverão ser analisados. A Figura 4.13 exibe esse gráfico.

Dada a Figura 4.13 é possível observar que há no Grafo G_5 um componente conexo com número superior de vértices quando comparado com os demais componentes do Grafo G_5 . Portanto, a análise prosseguirá sobre esse componente conexo do Grafo G_5 , que será identificado como CG_5 . O Componente CG_5 possui 111.889 vértices e 177.676 arestas.

4.5.4 Apresentação das Métricas Aplicadas ao Componente Conexos CG_5

Em relação à centralidade dos vértices no Componente CG_5 , a centralidade de grau máximo de um vértice é 11.446, a centralidade de grau mínimo de um vértice é 1 e a centralidade de grau médio dos vértices é 3,1759. Para avaliar melhor a distribuição de centralidade de grau no Componente CG_5 , foi gerado um histograma dos valores normalizados pelo número de vértices presentes no Componente CG_5 que está exibido na Figura 4.14. Para esse histograma, os valores do eixo x foram normalizados dividindo-se o valor da centralidade de grau do vértice pelo grau máximo possível, dado por $n - 1$, em um grafo não-direcionado obtido a partir de CG_5 , onde n é o número de vértices em

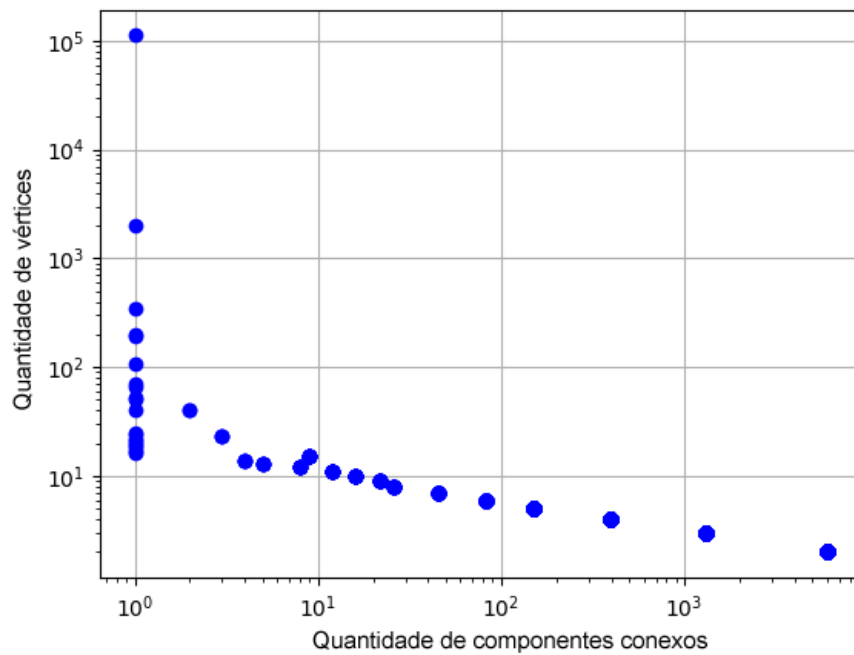


Figura 4.13: Distribuição do tamanho dos componentes no Grafo G_5 (eixos ajustados para escala logarítmica).

CG_5 ; e os valores do eixo y foram ajustados para escala logarítmica.

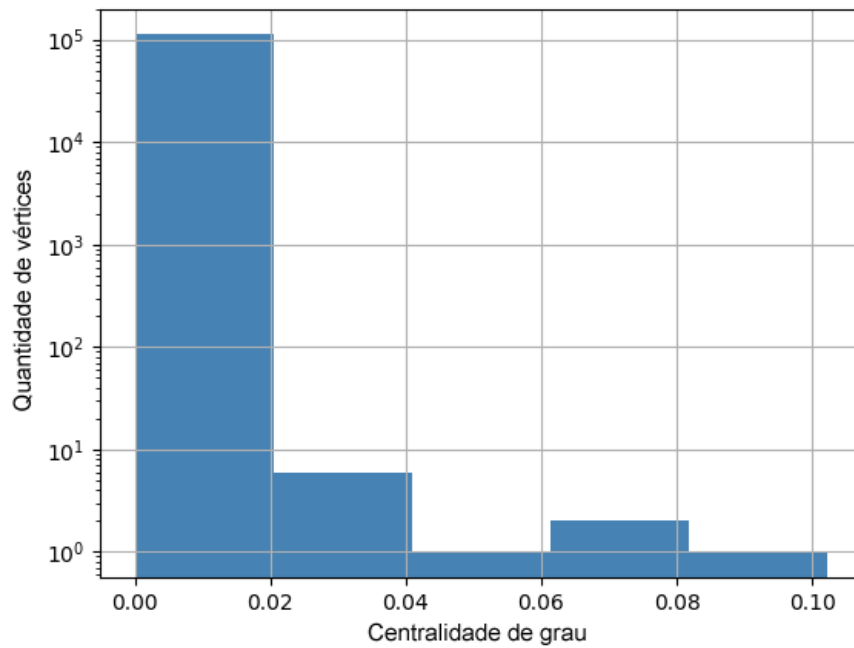


Figura 4.14: Distribuição da centralidade de grau no Componente CG_5 .

Para entender quais são os perfis dos vértices que tiveram mais conexões no Compo-

nente CG₅, foram geradas duas listas: uma lista apresentando os dez vértices com maiores valores de grau de entrada no Componente CG₅, apresentada na Tabela 4.30, e outra lista com os dez vértices com maior valor de grau de saída no Componente CG₅, apresentada na Tabela 4.31.

Tabela 4.30: Vértices com os dez maiores valores de grau de entrada no Componente CG₅.

Vértice	Grau de entrada	Grau de saída	Centralidade de grau
beawtifuleyes	11.446	0	11.446
itspedrito	8.864	1	8.865
lucasfelpi	7.316	4	7.320
augustopbarba	4.794	3	4.797
kkauan_	4.573	0	4.573
g1	4.534	2	4.536
inep_oficial	4.517	24	4.541
luscafp	3.435	0	3.435
jairbolsonaro	3.254	0	3.254
updatechartuc	2.300	0	2.300

Tabela 4.31: Vértices com os dez maiores valores de grau de saída no Componente CG₅.

Vértice	Grau de saída	Grau de entrada	Centralidade de grau
oaba_	177	261	438
botcefet	153	0	153
thenewest14	144	0	144
jssica20028683	78	0	78
lunaticalunna	76	1	77
angelux26	57	10	67
filipinazzz	55	5	60
lehmon4de	54	0	54
josuelrony	50	0	50
assuntosbrasil	49	2	51

Em seguida, foram calculadas as centralidades de autovetor dos vértices presentes no Componente CG₅. Os dez vértices com os maiores valores de centralidade de autovetor estão listados na Tabela 4.32.

Para avaliar o fluxo de informação foram calculados o coeficiente de agrupamento médio e a densidade do Componente CG₅. O coeficiente de agrupamento médio calculado é 0,1254 e a densidade do componente é 0,00001. Como o Componente CG₅ não é fortemente conexo, não foi possível calcular a distância média e diâmetro.

Ainda em relação à troca de informações entre os vértices, foi gerada a listagem dos

Tabela 4.32: Vértices com os dez maiores valores de centralidade de autovetor no Componente CG₅.

Vértice	Centralidade de autovetor	Grau de entrada	Grau de saída	Centralidade de grau
pretosnoenem	0,4709	1.986	37	2.023
inep_oficial	0,3605	4.517	24	4.541
itspedrito	0,1652	8.864	1	8.865
friviao	0,1297	41	0	41
karevalwx	0,1217	17	0	17
vascodagama	0,1106	3	0	3
nayarafelizardo	0,1104	186	2	188
aniellefranco	0,1052	941	1	942
kaiquebritor	0,1051	8	0	8
brunogagliasso	0,1051	3	0	3

dez maiores valores de centralidade de proximidade no Componente CG₅. Essa listagem está exibida na Tabela 4.33.

Tabela 4.33: Vértices com os dez maiores valores de centralidade de proximidade no Componente CG₅.

Vértice	Centralidade de proximidade	Grau de entrada	Grau de saída	Centralidade de grau
oaba_	0,0027	261	177	438
wtfrb	0,0023	5	25	30
noitaripsniaton	0,0023	0	11	11
lucasgaels	0,0022	0	22	22
universo_flash	0,0022	0	23	23
nah_kath	0,0022	0	21	21
cigarretsaft	0,0022	0	4	4
tealisa_	0,0022	2	17	19
orbitstardust	0,0022	0	15	15
samuraiantifa	0,0022	1	11	12

Para entender como os vértices se organizaram em comunidades nas manifestações sobre o ENEM 2020 foi executado, por meio do Gephi, o algoritmo Louvain no Componente CG₅. Foram obtidas 66 comunidades a partir desse algoritmo. A Figura 4.15 mostra a representação visual do resultado obtido pelo algoritmo executado. Os vértices têm seus identificadores exibidos com tamanho de fonte proporcional às suas centralidades de grau, e quanto mais centrais ao grafo, mais próximos ao centro do círculo são posicionados. As comunidades são posicionadas próximas às comunidades que tiveram mais relações. As principais comunidades detectadas podem ser identificadas na Figura 4.15 a partir das cores com as definições dadas a seguir. As definições foram inferidas a partir da avaliação de amostras dos vértices atribuídos às comunidades identificadas pelo algoritmo Louvain.

1. Vermelho, laranja, rosa e amarelo: comunidades em favor do cancelamento ou prorrogação do ENEM 2020.
2. Roxo: comunidade referente a veículos de imprensa.
3. Tons de verde: comunidades referentes a apoiadores do Governo Federal ou órgãos do governo.

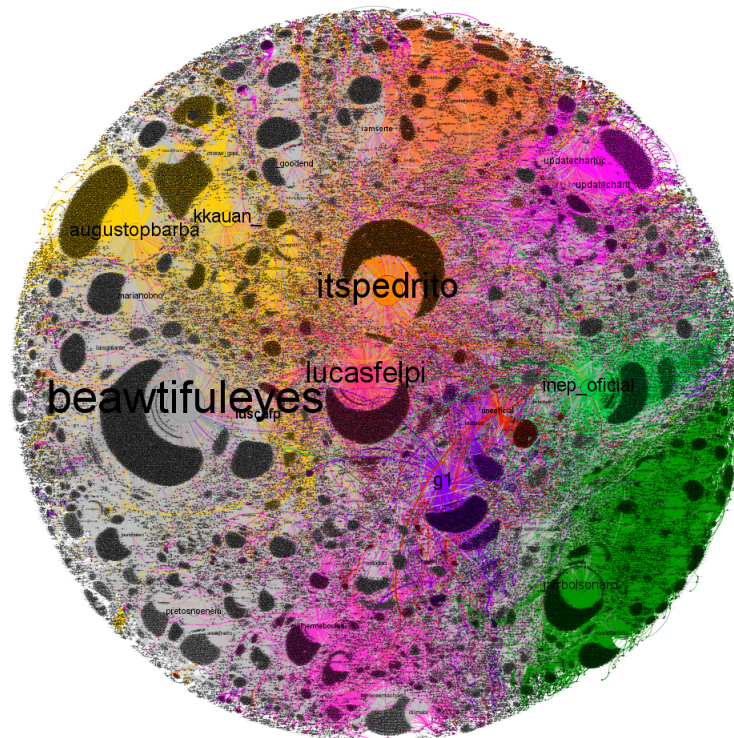


Figura 4.15: Representação visual do Componente CG_5 .

Por fim, foram listados os maiores valores de centralidade de intermediação, como exposto na Tabela 4.34, para análise.

A Tabela 4.35 mostra um resumo das métricas de grafo para o Grafo G_5 e para o Componente CG_5 .

4.5.5 Análise dos Resultados

Observa-se, na Tabela 4.30, que figuram entre os vértices com os maiores valores de grau de entrada apenas dois vértices relacionados ao Governo Federal: INEP (*inep_oficial*) e presidente Jair Bolsonaro (*jairbolsonaro*). São dois vértices que possuem responsabilidade nas decisões a serem tomadas sobre o ENEM 2020. Além destes, há vértices de

Tabela 4.34: Vértices com os dez maiores valores de centralidade de intermediação no Componente CG₅.

Vértice	Centralidade de intermediação	Grau de entrada	Grau de saída	Centralidade de grau
inep_oficial	0,0019	4.517	24	4.541
will0101010	0,0018	17	20	37
fraulein_erika	0,0009	7	9	16
muzzialves	0,0008	245	4	249
bandeirinha007	0,0008	1	7	8
lucasfelpi	0,0008	7.316	4	7.320
hwneymoon_	0,0008	179	27	206
gabriel25164967	0,0008	30	15	45
drikkjk	0,0008	7	6	13
otsulasimperian	0,0007	16	13	29

Tabela 4.35: Resumo das métricas de grafo para o Grafo G₅ e para o Componente CG₅.

	Grafo G₅	Componente CG₅
Vértices	135.695	111.889
Arestas	194.759	177.676
Centralidade de grau mínima	1	1
Centralidade de grau máxima	11.446	11.446
Centralidade de grau média	2,8705	3,1759
Coeficiente de agrupamento médio	0,0375	0,1254
Densidade	0,00001	0,00001
Comunidades	-	66

influenciadores que costumam publicar sobre educação e assuntos de interesse dos jovens: Lucas Felpi (*lucasfelpi*), Pedrão (*itspedrito*) e UpdateChart (*updatechartuc*). Esses vértices podem ter estimulado as cobranças sobre as autoridades em relação ao ENEM 2020 no Twitter.

Em relação aos vértices com os maiores valores de grau de saída (Tabela 4.31), há uma situação adversa: os quatro primeiros vértices em valor de grau de saída — *oaba*, *botcefet*, *thenewest14* e *jssica20028683* — aparentam ser contas automatizadas ou inautênticas. Os demais vértices são usuários com baixo número de seguidores e sem verificação do Twitter, indicando que as manifestações ocorreram de maneira informal, sem uma organização a partir de um grupo. Além disso, observando-se as publicações coletadas, nota-se que os usuários não demonstram posicionamento político explícito e aparentam ser jovens. Essa observação encontra relação com o público atingido pela incerteza sobre o ENEM 2020: jovens estudantes preocupados com a situação da prova, que é tão importante ao final do ensino médio.

A partir dos dados sobre a distribuição do centralidade de grau (Figura 4.14), os graus de entrada (Tabela 4.30) e os graus de saída (Tabela 4.31), é possível considerar que o fluxo de informação das manifestações sobre o ENEM 2020 é iniciado a partir de jovens preocupados com a realização da prova e se direcionaram às autoridades responsáveis pelo ENEM 2020. As manifestações contaram com o apoio de usuários mais influentes, como *itspedrito* e *lucasfelpi*, para a visibilidade da causa.

Pode ser observado na Tabela 4.32 que há apenas um vértice relacionado ao Governo Federal entre os dez vértices com maiores valores de centralidade de autovetor, o INEP. Os demais vértices, com exceção do vértice do Vasco da Gama (*vascodagama*), são de ativistas de diversas causas, principalmente do movimento negro, como *pretosnoenem* e *aniellefranco*, que se manifestaram sobre o ENEM 2020. Com isso, a observação anterior sobre o fluxo de informação é reforçada e nota-se que a autoridade em evidência, ou mais cobrada, nas manifestações foi o INEP.

Os valores de densidade e coeficiente de agrupamento médio, apresentados na Seção 4.5.4, mostram que os vértices do Componente CG₅ não estabeleceram conexões estreitas entre si. Pode-se observar o impacto desse comportamento a partir dos valores de centralidade de proximidade observados na Tabela 4.33. Foram valores muito baixos para todos os dez vértices listados, o que indica que as informações não percorreram o grafo com facilidade. Dessa maneira, pode-se supor, mais uma vez, que as manifestações foram individuais e sem uma organização para engajar.

É possível observar a descentralização da rede do Componente CG₅ na Figura 4.15. Há duas comunidades formadas em torno dos vértices dos influenciadores *itspedrito* e *lucasfelpi*. Além disso, os vértices *itspedrito* e *lucasfelpi*, este último faz publicações sobre educação, ficaram bem centralizados no gráfico, indicando suas relevâncias nestas manifestações. Também há comunidades relacionadas aos responsáveis pela realização do ENEM 2020, em verde no gráfico: INEP e presidente Jair Bolsonaro. Entretanto, essas comunidades possuem apenas 12,51% dos vértices do grafo, sendo assim, uma pequena parte da rede de informação formada. Esta composição mostra a ausência do Governo Federal no debate sobre o ENEM no Twitter, seja pela cobrança direta dos estudantes ou pela comunicação do próprio governo sobre as mudanças no ENEM. Dessa forma, percebe-se que os estudantes buscaram produzir uma mobilização sobre o cancelamento ou adiamento do ENEM por meio de hashtags e com o apoio de influenciadores jovens ao invés de direcionar as críticas e cobranças aos responsáveis pelas decisões sobre a prova — o Governo Federal, o MEC ou o INEP.

5. Discussão Sobre os Resultados

Após as análises individuais dos assuntos selecionados, os resultados obtidos serão reunidos e discutidos nesta seção a fim de se obter as informações que este estudo tem como objetivo.

Em três das cinco manifestações analisadas, houve pouca presença dos apoiadores do presidente Jair Bolsonaro, o que dificultou comparar as redes dos grupos de apoiadores com as de oposição ao presidente. Nas manifestações sobre os movimentos #Somos70PorCento e #Somos57MILHOES (Seção 4.3), em que houve forte presença de apoiadores do governo, pode-se notar que as comunidades de apoio a Jair Bolsonaro produziram uma rede com característica descentralizada, enquanto as comunidades de oposição ao presidente produziram uma rede com característica distribuída. Entretanto, ao analisar as demais manifestações, onde há maioria de comunidades de grupos contrários ou neutros ao Governo Federal, pode-se notar, pelos gráficos e valores de centralidade de grau, que essas comunidades também produziram redes com características descentralizadas. Dessa forma, há um indicativo de que o padrão de topologia produzido pelas redes de manifestações é a descentralizada e ocorreu um comportamento diferente para as manifestações dos movimentos #Somos70PorCento e #Somos57MILHOES. Como o movimento #Somos70PorCento surgiu em outra rede social, o YouTube, e ganhou apoiadores famosos, como a cantora Pabllo Vittar e a apresentadora Xuxa, é possível relacionar essa diferença de topologia pela origem e espontaneidade da manifestação. Ao contrário das outras manifestações que surgem e se expandem no Twitter, o movimento #Somos70PorCento surge no YouTube e se espalha para outras redes sociais. Já o movimento #Somos57MILHOES surge como resposta, no Twitter, para o #Somos70PorCento e, de tal maneira, mantém a topologia encontrada nas demais manifestações analisadas.

A estrutura global encontrada nas redes online das manifestações analisadas nesta pesquisa diverge das pesquisas apontadas por Gonzalez-Bailón e Wang [15] que sugerem que as redes online estão longe das estruturas descentralizadas. No entanto, Park [17]

explica que o Twitter forma estruturas hub-and-spoke, isto é, que as informações são provavelmente organizadas por vários usuários com posições centrais nas redes. "Eles não apenas desempenham papéis essenciais como líderes de opinião na disseminação de informações, mas também criam pontes para outras pequenas comunidades na rede" [17]. Dessa maneira, pode haver limitações nos fluxos de informação sem alguns atores-chave que unem comunidades diferentes, e é provável que as informações sejam controladas por esses atores. Gonzalez-Bailón e Wang [15] afirmam que o controle do fluxo de informação concede poder e que esta é a ideia no cerne dos relatos teóricos recentes de como as tecnologias digitais estão transformando os movimentos sociais e a ação coletiva.

Observaram-se também as interações entre comunidades distintas nas manifestações analisadas. Nas comunidades que representam grupos de oposição ou neutralidade ao governo Jair Bolsonaro, é possível observar integração entre elas provocada pelas trocas de mensagens entre seus vértices. Por outro lado, há pouca integração entre as comunidades de grupos de oposição ou neutralidade ao governo Jair Bolsonaro e as comunidades de grupos de apoio ao governo Jair Bolsonaro. Agathangelou et al. [19] afirmam que a ideologia é a força motriz na formação das câmaras de eco nas redes sociais. Como tal, é visível o isolamento das comunidades de grupos de apoio ao presidente em todas as redes de manifestações analisadas. Este comportamento sugere que comunidades com posicionamentos contrários não trocam mensagens entre si, ainda que para dissenso, portanto os fluxos de informação em maior parte são mantidos entre comunidades de posicionamentos semelhantes. Entretanto, como há mais diversidade política nas comunidades de grupos de oposição ou neutralidade ao governo Jair Bolsonaro, há também um debate de ideias mais forte, algo importante e essencial para uma democracia.

Um dos principais atores identificados nas manifestações analisadas, seja por centralidade de grau ou por centralidade de autovetor, é o deputado federal Marcelo Freixo. O deputado de oposição Freixo apareceu em quase todas as manifestações com valores altos para essas centralidades citadas. Outros atores de destaque na oposição são políticos do PCdoB, PSOL e PT, todos partidos de esquerda, como Jandira Feghali, Manuela D'Ávila, Sâmia Bomfim, Guilherme Boulos, Fernando Haddad e Humberto Costa. Esses partidos tiveram seus representantes presentes com destaque nas manifestações analisadas, seja por centralidade de grau, intermediação, proximidade ou autovetor, tendo, dessa maneira, participações centrais de diversas formas nas manifestações. Pode ser notado também uma forte participação de atores relacionados ao movimento negro no Twitter, como, por exemplo, Andreza Delgado e Anielle Franco, principalmente em manifestações de caráter mais popular como Breque dos Apps (Seção 4.2), auxílio emergencial (Seção 4.4) e ENEM 2020 (Seção 4.5).

Pelo lado do governo, os principais atores identificados nas manifestações analisadas são da família Bolsonaro — o presidente Jair Bolsonaro, o senador Flávio Bolsonaro e o deputado federal Eduardo Bolsonaro — que aparecem em quase todas as manifestações (exceto a Breque dos Apps) com centralidades de grau altas. Entre seus apoiadores, políticos da base governista — *e.g.* Bia Kicis e Carla Zambelli — e jornalistas bolsonaristas — *e.g.* Bernardo Küster e Guilherme Fiuza — aparecem pontualmente nas manifestações analisadas, mas, esses apoiadores não se repetem a cada manifestação. O apoiador do governo que aparece em mais de uma manifestação é o blogueiro de extrema-direita Allan dos Santos. Como Jair Bolsonaro é o atual presidente do Brasil, a centralidade dele nas manifestações é esperada e se explica pelo seu cargo atual. Ainda assim, os principais atores que apoiam o governo nas manifestações são mantidos dentro da família Bolsonaro, principalmente o presidente Jair Bolsonaro, apresentando uma característica bastante personalista por parte dos grupos de seus apoiadores no Twitter.

Ao se observar o perfil dos atores que participaram das manifestações analisadas, é possível notar que há diferenças conforme a manifestação. As manifestações sobre o novo FUNDEB (Seção 4.1), que é relacionada a um assunto legislativo, foi composta por políticos, partidários, militantes políticos, movimentos sociais, sindicatos e ativistas pela educação brasileira. Esta manifestação, portanto, apresentou um caráter mais político.

Examinando-se os atores que compuseram as manifestações sobre Breque dos Apps, auxílio emergencial e ENEM 2020, identifica-se que os atores são, em geral, pessoas sem identificação político-partidária que reivindicam seus direitos ou defendem uma pauta. Com exceção das manifestações sobre a Breque dos Apps, poucos agentes políticos, além do presidente e ministros, foram observados com relevância nessas manifestações. Dessa maneira, as manifestações sobre Breque dos Apps, auxílio emergencial e ENEM 2020 demonstraram um caráter mais social.

Na análise das manifestações sobre os movimentos virtuais #Somos70PorCento e #Somos57MILHOES (Seção 4.3), é possível observar que há a presença tanto de militantes políticos quanto de pessoas sem identificação político-partidária, além de uma grande participação de todos os atores envolvidos. Estas manifestações produziram o grafo com o maior número de arestas deste estudo e também apresentam a maior média de centralidade de grau entre as manifestações analisadas. Com um caráter mais partidário, essas manifestações foram as que mais engajaram participantes para demonstrar apoio a um dos movimentos na disputa.

De forma geral, foi observado que a oposição ao governo Bolsonaro teve mais destaque nas manifestações sobre o FUNDEB (Seção 4.1), uma matéria legislativa. Apesar

de o FUNDEB ser relacionado à educação brasileira, não é possível considerar que esse tema foi o maior motivador das manifestações visto que o ENEM 2020 não apresentou o mesmo engajamento pela oposição ainda que também seja um assunto referente à educação brasileira. No entanto, pode-se considerar a urgência do FUNDEB, por conta do prazo de validade, como fator motivador para esta manifestação.

Os apoiadores do governo Bolsonaro apareceram com mais relevância nas manifestações de confronto entre apoio e oposição ao governo como as manifestações sobre os movimentos virtuais #Somos70PorCento e #Somos57MILHOES. Nas demais manifestações, os grupos de apoiadores tiveram pouca relevância ou sequer foram possíveis de detectar. Esse aspecto demonstra que a motivação é entre eles é de caráter personalista ou partidária.

Observando os dados sobre distribuição de centralidade de grau de todas as manifestações, nota-se que todas as distribuições muito se assemelham à cauda-longa, distribuição esperada para redes reais. A distribuição de centralidade de grau apresentada pelas manifestações apresentaram concentração em valores baixos de centralidade de grau seguida de poucos vértices com valores altos de centralidade de grau. A presença majoritária de conexões em poucos vértices é uma propriedade do modelo de redes sem escala. Este modelo de redes complexas apresenta-se então como padrão de modelo de redes para manifestações populares no Twitter. Uma investigação mais profunda sobre os modelos de redes complexas em outros contextos pode trazer orientações de como gerar uma democracia mais participativa por meio de manifestações online e como os políticos podem aproveitar para se comunicar com os cidadãos nesses momentos.

6. Conclusão

As redes sociais tornaram-se “protagonistas” em relação à mobilização e ao engajamento da sociedade nos debates políticos e sociais. Essa politização produzida por meio da Internet tem se intensificado a cada ano em todo o Mundo. As manifestações populares online reforçam a importância das redes sociais como plataformas políticas e ambiente propício para a troca de opiniões entre as pessoas no Brasil, a exemplo do que acontece em outros países.

Este estudo analisou algumas manifestações populares realizadas no Twitter que ocorreram ao longo do ano de 2020, durante a pandemia da COVID-19, a fim de entender como são as manifestações políticas, partidárias e sociais em ambientes virtuais no Brasil, esclarecendo seus fluxos de informação, analisando a estrutura global das redes online e apontando seus principais atores. Para alcançar o objetivo principal desta pesquisa, uma metodologia foi planejada e realizada, cujos resultados foram discutidos considerando seus objetivos determinados. Para isso, foi necessário identificar as manifestações em destaque no Twitter, para então coletar suas publicações, modelá-las em grafos e analisar os dados obtidos por meio da Teoria de Grafos. Além disso, foi feito o acompanhamento das manifestações selecionadas no Twitter para gerar observações qualitativas sobre elas, o que ajudou no desenvolvimento das análises por realizar uma aproximação ao objeto de estudo.

A fim de compreender os fluxos de informação em manifestações populares realizadas no Twitter, foi possível verificar que as redes das manifestações analisadas no Twitter tinham algumas semelhanças entre si. Entre essas semelhanças, pode ser citada a distribuição de grau em cauda-longa, o que é esperado para redes reais. Outra semelhança entre as redes observadas é como estas se organizaram muitas vezes em topologia descentralizada e com característica de redes sem escala, dando a alguns atores mais relevância que os demais nas manifestações e permitindo que suas mensagens fossem espalhadas mais facilmente para a rede na totalidade.

Em seguida, foram identificados os grupos presentes nas manifestações online por meio de uma técnica que aplica o uso do algoritmo de Louvain, uma heurística de detecção de comunidades, em associação com visualizações dos grafos geradas pela ferramenta Gephi. Dessa maneira, foi possível identificar a presença de grupos de apoiadores do presidente Jair Bolsonaro com relevância em algumas manifestações. Esses grupos sempre estão isolados nos grafos dos demais grupos que oferecem oposição ao governo Bolsonaro. Foram também identificados grupos relacionados à oposição ao governo de Jair Bolsonaro, com mais destaque aos partidos PT, PSOL e PCdoB. Esses grupos também tiveram presença em quase todas as manifestações. Em manifestações com caráter mais popular (como as da Breque dos Apps, do auxílio emergencial e da realização do Enem 2020), foi observada a presença de grupos relacionados ao movimento negro no Twitter. Todos esses grupos foram identificados também nas observações qualitativas das manifestações no Twitter, o que assegura, até certo ponto, a reflexão da realidade às comunidades que representam tais grupos, identificadas pelo algoritmo.

Por fim, foram identificados os atores mais importantes nas manifestações online por meio de uma técnica que aplica um conjunto de métricas de Análise de Redes Sociais. A partir dessa técnica, foi possível identificar que o ator mais relevante no Twitter nas manifestações populares analisadas não foi o presidente Jair Bolsonaro, algo que era esperado devido ao seu cargo, mas o deputado federal da oposição Marcelo Freixo. Esse cenário reflete positivamente com o que foi encontrado nas observações qualitativas feitas no Twitter. Outros políticos da oposição como, por exemplo, Guilherme Boulos e Jandira Feghali, também foram atores de destaque nas manifestações. Os deputados aliados do governo Bolsonaro, além do próprio presidente e seus filhos, estiveram presentes em todas as manifestações analisadas, entretanto sem tanta relevância em três delas - as manifestações sobre o novo Fundeb, a manifestação da Breque dos Apps e a manifestação sobre o ENEM 2020. Esta situação demonstra que há no mínimo um equilíbrio de forças nessa rede social. Também pode-se apontar como os apoiadores do governo Bolsonaro têm engajamento partidário enquanto os opositores do governo Bolsonaro se destacam pelo engajamento em pautas sociais e legislativas.

Em conclusão, pode-se reforçar que o uso de plataformas de redes sociais para manifestação social e militância política está crescendo e afetando como a população debate e se mobiliza. Como tal, torna-se difícil compreender todos os cenários e impactos possíveis na sociedade, visto que várias plataformas tecnológicas estão sendo exploradas para esses fins atualmente. Como George e Leidner [16] citam: o nível de ativismo digital que vemos hoje é provavelmente um precursor de grandes ações empreendidas pelos cidadãos do futuro e, conforme a tecnologia se desenvolve com as habilidades em sistemas

de informação dos cidadãos, a oportunidade de agir diretamente no meio virtual aumenta. Portanto, métodos maliciosos, como *bots*, *fake news* e *trolls*, que podem ser cada vez mais explorados e já têm sido usados para simular apoio, realizar ataques e afetar os debates e as escolhas dos cidadãos [43], devem ser colocado em perspectiva para trabalhos futuros. Apesar de a pandemia do COVID-19 não ter sido o tema central desta pesquisa, também é relevante identificar o impacto da pandemia nas plataformas de redes sociais online, o que torna bastante interessante um estudo futuro sobre o impacto desse evento na estrutura e características das redes. Além disso, seria importante viabilizar para outros pesquisadores a realização de análises semelhantes às apresentadas nesta pesquisa. Portanto, o trabalho de elaboração de um suporte computacional para apoiar observadores de redes sociais a fazerem pesquisas sobre publicações em plataformas de redes sociais online, de tal maneira a facilitar o trabalho de reportar o que está acontecendo nas redes sociais e, assim, diminuir o tempo e o conhecimento técnico necessário para fazer análises como as apresentadas nesta dissertação, é mais uma oportunidade de trabalho futuro nesta área de pesquisa.

Referências Bibliográficas

- [1] BARAN, P. *On Distributed Communications: I. Introduction to Distributed Communications Networks*. Santa Monica: RAND Corporation, 1964.
- [2] ELIXIR: A distributed infrastructure for life-science information. 2015. Disponível em: <<http://160.80.34.9/elixir2015/>>. Acesso em: 09/02/2021.
- [3] SOCIAL Networking Popular Across Globe. Pew Research Center, 2012. Disponível em: <<https://www.pewresearch.org/global/2012/12/12/social-networking-popular-across-globe/>>. Acesso em: 21/09/2020.
- [4] LEE, S.; XENOS, M. Social distraction? social media use and political knowledge in two u.s. presidential elections. *Computers in Human Behavior*, v. 90, p. 18–25, 2019.
- [5] KEMP, S. *Digital 2019: Brazil*. 2019. 77 p. Disponível em: <<https://datareportal.com/reports/digital-2019-brazil>>. Acesso em: 21/09/2020.
- [6] COHEN, R.; HAVLIN, S. *Complex Networks: Structure, Robustness and Function*. Nova York: Cambridge University Press, 2010.
- [7] TSVETOVAT, M.; KOUZNETSOV, A. *Social Network Analysis for Startups*. Sebastopol: O'Reilly Media, Inc., 2011.
- [8] DEKKER, A. Conceptual distance in social network analysis. *Journal of Social Structure*, v. 6, 2005.
- [9] ROCHAT, Y. *Closeness Centrality Extended To Unconnected Graphs: The Harmonic Centrality Index*. Zürich: Institute of Applied Mathematics, 2009.
- [10] SATTAR, N. S.; ARIFUZZAMAN, S. Parallelizing louvain algorithm: Distributed memory challenges. In: *2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl*

- [11] MITZLAFF, F. et al. The social distributional hypothesis: a pragmatic proxy for homophily in online social networks. *Social Network Analysis and Mining*, Springer Vienna, v. 4, 2014.
- [12] STROGATZ, S. Exploring complex networks. *Nature*, v. 410, p. 268–276, 2001.
- [13] WATTS, D. J.; STROGATZ, S. H. Collective dynamics of 'small-world' networks. *Nature*, Nature Publishing Group, v. 393, n. 6684, p. 440–442, 1998.
- [14] NEWMAN, M. E. J. Power laws, pareto distributions and zipf's law. *Contemporary Physics*, Taylor and Francis, v. 46, n. 5, p. 323–351, 2005.
- [15] GONZÁLEZ-BAILÓN, S.; WANG, N. Networked discontent: The anatomy of protest campaigns in social media. *Social Networks*, v. 44, p. 95–104, 2016.
- [16] GEORGE, J.; LEIDNER, D. Digital activism: a hierarchy of political commitment. In: *2018 51st Hawaii International Conference on System Sciences*. Havaí: AIS, 2018.
- [17] PARK, S. J.; LIM, Y. S.; PARK, H. W. Comparing Twitter and YouTube networks in information diffusion: The case of the "Occupy Wall Street" movement. *Technological Forecasting and Social Change*, v. 95, p. 208–217, 2015.
- [18] NAM, Y.; LEE, Y.; PARK, H. Can web ecology provide a clearer understanding of people's information behavior during election campaigns? *Social Science Information*, v. 52, p. 91–109, 2013.
- [19] AGATHANGELOU, P. et al. Understanding online political networks: The case of the far-right and far-left in Greece. In: *9th International Conference SocInfo 2017*. Oxford: Springer International Publishing, 2017. p. 162–177.
- [20] SILVA, E. L. da; MENEZES, E. M. *Metodologia da Pesquisa e Elaboração de Dissertação*. 3. ed. Florianópolis, 2001.
- [21] FACHIN, O. *Fundamentos de Metodologia*. 5. ed. São Paulo, 2005. 209 p.
- [22] MARTINS, E.; GONÇALVES, K.; FILHO, R. M. Caracterizando a campanha presidencial brasileira em 2018 usando dados do Twitter. In: *Anais do VIII Brazilian Workshop on Social Network Analysis and Mining*. Porto Alegre: SBC, 2019. p. 131–142.

- [23] Jayawardena, L. Impact of internet and social networks on the final result of 2016 presidential election of the United States. In: *2017 National Information Technology Conference (NITC)*. Colombo: IEEE, 2017. p. 80–85.
- [24] Rofrío, D. et al. Presidential elections in Ecuador: Bot presence in Twitter. In: *2019 Sixth International Conference on eDemocracy eGovernment (ICEDEG)*. Quito: IEEE, 2019. p. 218–223.
- [25] TWITTER, I. *Getting started*. 2021. Disponível em: <<https://developer.twitter.com/en/docs/twitter-api/getting-started/about-twitter-api>>. Acesso em: 30/01/2022.
- [26] GETDAYTRENDS. 2020. Disponível em: <<https://getdaytrends.com/brazil/>>. Acesso em: 21/09/2020.
- [27] MONTENEGRO, C.; DAMASCENO, N. Imprensa, internet e redes sociais: Uma análise sobre o posicionamento dos jornais no Twitter. In: *Congresso Brasileiro de Ciências da Comunicação*. Rio de Janeiro: Intercom, 2019.
- [28] NEWMAN, N. et al. *Reuters Institute Digital News Report 2020*. 2020.
- [29] NETWORKX. 2020. Disponível em: <<https://networkx.org/>>. Acesso em: 10/01/2021.
- [30] BASTIAN, M.; HEYMANN, S.; JACOMY, M. Gephi: An open source software for exploring and manipulating networks. *Proceedings of the International AAAI Conference on Web and Social Media*, v. 3, n. 1, p. 361–362, Mar. 2009.
- [31] BLONDEL, V. D. et al. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, IOP Publishing, v. 2008, n. 10, 2008.
- [32] MORALES, J. Entenda o novo Fundeb, aprovado na Câmara dos Deputados. *Guia do Estudante*, 20 jul. 2020. Disponível em: <<https://guiadoestudante.abril.com.br/atualidades/entenda-o-novo-fundeb-que-sera-votado-na-camara-dos-deputados/>>. Acesso em: 23/09/2020.
- [33] IDOETA, P. A. O que é o Fundeb, fundo essencial para a educação que terá votação crucial no Congresso. *BBC News Brasil*, São Paulo, 15 jul. 2020. Disponível em: <<https://www.bbc.com/portuguese/brasil-53398303>>. Acesso em: 23/09/2020.
- [34] MENDES, G. *Alexandre de Moraes determina extradição e prisão de Allan dos Santos*. 2021. Disponível em:

<<https://congressoemfoco.uol.com.br/area/justica/alexandre-de-moraes-prisao-allan-dos-santos/>>. Acesso em: 03/12/2021.

- [35] ALVARENGA, D.; SILVEIRA, D. Desemprego no Brasil atinge recorde de 14,4% no trimestre encerrado em agosto, diz IBGE. *GI*, 30 out. 2020. Disponível em: <<https://g1.globo.com/economia/noticia/2020/10/30/desemprego-no-brasil-sobe-para-144percent-em-agosto-diz-ibge.ghtml>>. Acesso em: 02/11/2020.
- [36] GARCIA, D. Informalidade supera 50% em 11 estados do país, diz IBGE. *Folha de São Paulo*, Rio de Janeiro, 14 fev. 2020. Disponível em: <<https://www1.folha.uol.com.br/mercado/2020/02/informalidade-atinge-recorde-em-19-estados-e-no-df-diz-ibge.shtml>>. Acesso em: 02/11/2020.
- [37] HENDERSON, A.; PIRAJIBE, P. Para sobreviver à crise, bares e restaurantes do Rio investem no atendimento delivery. *GI*, Rio de Janeiro, 14 ago. 2020. Disponível em: <<https://g1.globo.com/rj/rio-de-janeiro/noticia/2020/08/14/para-sobreviver-a-crise-bares-e-restaurantes-do-rio-investem-no-atendimento-delivery.ghtml>>. Acesso em: 02/11/2020.
- [38] MACHADO, L. Greve dos entregadores: o que querem os profissionais que fazem paralisação inédita. *BBC News Brasil*, São Paulo, 22 jun. 2020. Disponível em: <<https://www.bbc.com/portuguese/brasil-53124543>>. Acesso em: 02/11/2020.
- [39] NOGUEIRA, P.; BENITES, A. “Nossa vida vale mais do que levar um prato de comida para as pessoas”. *El País Brasil*, 25 jul. 2020. Disponível em: <<https://brasil.elpais.com/brasil/2020-07-25/nossa-vida-vale-mais-do-que-levar-um-prato-de-comida-para-as-pessoas.html>>. Acesso em: 02/11/2020.
- [40] Bolsonaro tem reprovação de 43% e aprovação de 33%, diz Datafolha. *GI*, São Paulo, 28 mai. 2020. Disponível em: <<https://g1.globo.com/politica/noticia/2020/05/28/bolsonaro-tem-reprovacao-de-43percent-e-aprovacao-de-33percent-diz-datafolha.ghtml>>. Acesso em: 21/11/2020.
- [41] MOVIMENTOS anti-Bolsonaro: o que é o ‘Somos 70%’ e outras iniciativas da sociedade civil contra o governo? *BBC*, 2 jun. 2020. Disponível em: <<https://www.bbc.com/portuguese/brasil-52898476>>. Acesso em: 21/11/2020.
- [42] G1. *Jair Bolsonaro é eleito presidente com 57,8 milhões de votos*. 2018. Disponível em: <<https://g1.globo.com/politica/eleicoes/2018/apuracao/presidente.ghtml>>. Acesso em: 21/11/2020.

- [43] ALMEIDA, Y. et al. O uso das redes sociais para interferir nas democracias: Um mapeamento sistemático da literatura. In: *Anais do IX Brazilian Workshop on Social Network Analysis and Mining*. Porto Alegre: SBC, 2020. p. 178–183.
- [44] Coronavírus: veja ações que estados e prefeituras estão adotando. *Jornal Hoje*, 14 mar. 2020. Disponível em: <<https://g1.globo.com/bemestar/coronavirus/noticia/2020/03/14/coronavirus-veja-medidas-que-estados-e-municipios-estao-tomando-para-evitar-surto.ghtml>>. Acesso em: 14/12/2020.
- [45] PIOVESAN, E.; MORAES, G. Câmara aprova auxílio de R\$600 para pessoas de baixa renda durante epidemia. *Agência Câmara Notícias*, Brasília, 26 mar. 2020. Disponível em: <<https://www.camara.leg.br/noticias/648863-camara-aprova-auxilio-de-r-600-para-pessoas-de-baixa-renda-durante-epidemia/>>. Acesso em: 14/12/2020.
- [46] PIOVESAN, E.; MORAES, G. *Câmara aprova auxílio de R\$ 600 para pessoas de baixa renda durante epidemia - Notícias*. Câmara dos Deputados, 2020. Disponível em: <<https://www.camara.leg.br/noticias/648863-camara-aprova-auxilio-de-r-600-para-pessoas-de-baixa-renda-durante-epidemia/>>. Acesso em: 03/12/2021.
- [47] BRANCO, A. P. Dificuldades e filas marcam os 111 dias de auxílio emergencial. *Agora*, São Paulo, 26 jul. 2020. Disponível em: <<https://agora.folha.uol.com.br/grana/2020/07/dificuldades-e-filas-marcam-os-111-dias-de-auxilio-emergencial.shtml>>. Acesso em: 14/12/2020.
- [48] TEÓFILO, S.; SOUZA, R.; GRECCHI, F. *Roberto Jefferson saiu de pivô do mensalão à extrema direita*. 2021. Disponível em: <<https://www.correiobraziliense.com.br/politica/2021/08/4943648-roberto-jefferson-saiu-de-pivo-do-mensalao-a-extrema-direita.html>>. Acesso em: 03/12/2021.
- [49] BOMFIM, C.; FALCÃO, M. *Ex-deputado Roberto Jefferson é preso pela Polícia Federal no inquérito das milícias digitais*. 2021. Disponível em: <<https://g1.globo.com/politica/noticia/2021/08/13/policia-federal-faz-operacao-para-prender-roberto-jefferson.ghtml>>. Acesso em: 03/12/2021.
- [50] EXAME Nacional do Ensino Médio (Enem). 2020. Disponível em: <<https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/enem>>. Acesso em: 26/12/2020.
- [51] AOS 20 anos, Enem Se consolida como a maior prova do país, mas deve mudar. *Revista Exame*, 28 out. 2018. Disponível em: <<https://exame.com/brasil/aos-20>>.

anos-enem-se-consolida-como-a-maior-prova-do-pais-mas-deve-mudar>. Acesso em: 26/12/2020.

[52] PROVAS do Enem 2020 serão em janeiro e fevereiro de 2021. *GI*, 8 jul. 2020. Disponível em: <<https://g1.globo.com/educacao/noticia/2020/07/08/prova-do-enem-2020-sera-em-janeiro-de-2021.ghtml>>. Acesso em: 26/12/2020.

A. Estrutura das Tabelas para Armazenamento dos Resultados das Coletas no SGBD MySQL

As Tabelas A.1, A.2 e A.3 apresentam a estrutura das tabelas utilizadas para armazenamento das coletas no SGBDR MySQL e utilizadas no Capítulo 4.

Tabela A.1: Estrutura das tabelas de tweets coletados (separadas por assunto selecionado)

Nome da coluna	Descrição da coluna	Tipo de dado
id	Número identificador do registro	Texto
tweet_id	Número identificador do tweet	Texto
author_id	Número identificador da conta do(a) autor(a) do tweet	Texto
tweet_text	Texto do tweet	Texto
tweet_created_at	Data e hora da publicação do tweet	Texto
tweet_hashtags	Hashtags presentes no texto do tweet	Texto
tweet_retweets	Número de retweets feitos do tweet	Inteiro
tweet_replies	Número de respostas recebidas no tweet	Inteiro
tweet_likes	Número de curtidas recebidas pelo tweet	Inteiro
tweet_quotes	Número de citações feitas do tweet	Inteiro
tweet_mentions	Números identificadores dos usuários mencionados no tweet	Texto
tweet_referenced_type	Tipo de interação feita pelo tweet (retweet, citação ou resposta)	Texto
tweet_referenced_tweet	Números identificadores dos tweets retweetados, citados ou respondidos pelo tweet	Texto

Tabela A.2: Estrutura da tabela de tweets referenciados (retweetados, citados ou respondidos)

Nome da coluna	Descrição da coluna	Tipo de dado
id	Número identificador do registro	Texto
tweet_id	Número identificador do tweet referenciado	Texto
author_id	Número identificador da conta do(a) autor(a) do tweet referenciado	Texto
tweet_text	Texto do tweet referenciado	Texto
tweet_created_at	Data e hora da publicação do tweet referenciado	Texto
tweet_hashtags	Hashtags presentes no texto do tweet referenciado	Texto
tweet_retweets	Número de retweets feitos do tweet referenciado	Inteiro
tweet_replies	Número de respostas recebidas no tweet referenciado	Inteiro
tweet_likes	Número de curtidas recebidas pelo tweet referenciado	Inteiro
tweet_quotes	Número de citações feitas do tweet referenciado	Inteiro

Tabela A.3: Estrutura da tabela de usuários

Nome da coluna	Descrição da coluna	Tipo de dado
id	Número identificador do registro	Texto
user_id	Número identificador da conta do(a) usuário(a) do Twitter	Texto
user_name	Nome do(a) usuário(a) do Twitter	Texto
user_screen_name	Identificador do(a) usuário(a) do Twitter	Texto
user_description	Biografia do(a) usuário(a) do Twitter	Texto
user_verified	Verificação da conta do(a) usuário(a) do Twitter	Booleano

B. Dados Complementares da Seção 4.1

Este apêndice é um complemento aos experimentos realizados na Seção 4.1. A Tabela B.1 apresenta os tweets coletados sobre o FUNDEB com mais interações.

Tabela B.1: Tweets coletados sobre o FUNDEB com mais interações

Data de publicação	Autor	Texto do tweet	Total de interações	# de retweets	# de respostas	# de citações
22/07/2020 14:31:10	GFiuzza_Oficial	A frente pela educação, que reúne Tábata Amaral, Alexandre Frota, FHC, Luciano Huck, Rodrigo Maia, Dilma Rousseff e outros educados só será verdadeiramente homogênea quando confirmar a adesão do Tiririca.	7981	6339	1250	392
20/07/2020 00:10:35	MarceloFreixo	ESCÂNDALO! O desgoverno Bolsonaro está propondo a deputados aliados que o Fundeb, principal fonte de financiamento da educação pública, fique suspenso em 2021. Isso significará o fechamento de escolas e a destruição da educação básica. É um c...	4170	3458	424	288
22/07/2020 00:48:30	GuilhermeBoulos	O Novo tentou ZERAR o aumento do piso salarial dos professores do Fundeb, que hoje é menos de R\$3 mil por 40 horas semanais. Foi derrotado. Querem que professor siga ganhando pouco e banqueiro ganhando muito. Isso é novo onde?	3519	3022	291	206
22/07/2020 02:42:15	RodrigoMaia	Aprovamos com 499 votos no primeiro turno e 492 no segundo, a renovação do Fundeb, fundo que financia a educação básica, com ampliação da participação de recursos da União em até 23%, e diminui desigualdades. Foi uma votação histórica, com ...	3291	867	2206	218
22/07/2020 11:54:19	Haddad_Fernando	A transformação do Fundeb em política pública permanente é a consagração da atuação do governo @LulaOficial na educação básica. Orgulho de ter participado dessa pequena revolução!	3287	2118	1001	168
21/07/2020 17:39:57	DomDasThreads	Aqui na minha casa esse moleque irresponsável não entra, tenho orgulho da educação que dei aos meus filhos. Alguém mais concorda comigo??? #FamiliasContraFelipeNeto	3235	2313	714	208
21/07/2020 00:48:00	MarceloFreixo	O PSL Personnalité, também conhecido como Partido Novo, não surpreende. Orientou a bancada a votar CONTRA a PEC do Fundeb e ainda diz que está defendendo os interesses nacionais. O investimento por aluno no Brasil é apenas 1/3 da OCDE, mas o Novo ac...	3165	2425	473	267
21/07/2020 23:46:09	GuilhermeBoulos	Fundeb aprovado! Vitória da educação e da mobilização social. Derrota de Bolsonaro.	3106	2377	532	197
22/07/2020 12:28:50	MinLuizRamos	Aprovamos o Fundeb na Câmara dos Deputados!!! O que os governos anteriores não fizeram o Pres. @jaírbolsonaro deu o 1º passo em menos de 2 anos. Esse Governo se preocupa com a Educação, afinal de contas, ela é um dos pilares para o desenvolvimento ...	3046	566	2262	218
22/07/2020 02:19:01	allantercalivre	@allantercalivre's account has been withheld in Brazil, Worldwide in response to a legal demand. Learn more.	2879	2237	525	117

C. Dados Complementares da Seção 4.2

Este apêndice é um complemento aos experimentos realizados na Seção 4.2. A Tabela C.1 apresenta os tweets coletados sobre o Breque dos Apps com mais interações.

Tabela C.1: Tweets coletados sobre o Breque dos Apps com mais interações

Data de publicação	Autor	Texto do tweet	Total de interações	# de retweets	# de respostas	# de citações
01/07/2020 13:05:10	nathfinancas	Não realizar pedidos hoje você: - Apoia os entregadores que estão buscando seus direitos - Economiza seu dinheiro - Faz aquela comida caseira deliciosa Entregador é da classe trabalhadora, apoie #BrequeDosApps	7527	7178	129	220
01/07/2020 12:58:32	YuriMarcal	quem pedir coisa nos app hoje é nazista #BrequeDosApps	6086	5638	197	251
01/07/2020 17:55:15	gduvievier	#BrequeDosApps gigante! FODA! https://t.co/mpNoYVmj4d	5723	4543	559	621
01/07/2020 01:57:05	galodeluta	Aqui nois não é empreendedor de porra nenhuma, nois é força de trabalho! Certo? #BrequeDosApps	5565	5029	164	372
01/07/2020 13:50:03	galodeluta	GREVE VEJA COMO AJUDAR: Não faça pedidos por App hoje Use a #BrequeDosApps Se precisar pedir comida, ligue no restaurante do bairro Dê nota baixa nas lojas dos aplicativos A LUTA É DE TODOS! #BrequeDosApps Il...	5144	4541	109	494
01/07/2020 16:37:18	galodeluta	RESPOSTA AO IFOOD - A THREAD #BrequeDosApps https://t.co/00QZWM5Tik	5102	4594	106	402
01/07/2020 10:48:04	ariaritamusic	galera vamo lembrar que HOJE não é pra NINGUÉM pedir entrega em ifood e apps semelhantes!! os entregadores estarão em paralisação nacional exigindo seus DIREITOS TRABALHISTAS e a nossa parte é divulgar a ideia e fazer nossa comidinha em casa #B...	4918	4702	64	152
01/07/2020 17:38:40	tretanotrampo	Gigante o #BrequeDosApps https://t.co/zCiOoTAow3	4086	3368	93	625
01/07/2020 11:00:51	gabrielaz	A Rappi vale \$3bi de dólares no mercado!!! Enquanto isso entregadores pedalam 10km por 2 reais. O corre de hoje é por EPIs, aumento do valor do km, valor mínimo de entrega, seguro contra roubos e acidentes e fim dos bloqueios! #ApoioBrequeDosApps #Bre...	3832	3670	23	139
30/06/2020 17:03:05	galodeluta	E ai pessoal vamos levantar a hastag #BrequeDosApps	3424	2597	187	640

D. Dados Complementares da Seção 4.3

Este apêndice é um complemento aos experimentos realizados na Seção 4.3. A Tabela D.1 apresenta os tweets coletados sobre os movimentos #Somos70porcento e #Somos57MILHOES com mais interações.

Tabela D.1: Tweets coletados sobre os movimentos #Somos70porcento e #Somos57MILHOES com mais interações

Data de publicação	Autor	Texto do tweet	Total de interações	# de retweets	# de respostas	# de citações
01/06/2020 11:57:00	pablovittar	NÃO SE CALE #AntiFascista #Antiracista #VidasNegrasImportam #Somos70porcento https://t.co/uMYkkRYcjM	15919	14726	689	504
31/05/2020 13:24:39	bernardopkuster	@bernardopkuster's account has been withheld in Brazil, Worldwide in response to a legal demand. Learn more.	12473	7225	3797	1451
31/05/2020 01:46:22	CarlaZambelli38	Neste domingo todos que apoiam a Democracia e Bolsonaro devem demonstrar indo para as ruas! Seja onde for! Apoiaremos nosso PR @jairbolssonaro sempre! #Somos57MILHOES	11225	5850	4558	817
30/05/2020 16:34:07	Joubert19	Você continua apoiando o Presidente Jair Bolsonaro? Sim/Não Responda usando a tag: #Somos57MILHOES	9517	2625	5926	966
30/05/2020 14:18:16	eduardomoreira	70% rejeitam aproximação ao Centrão (Datafolha) 70% acham Bolsonaro Pessimo/Ruim/Regular 70% apoiam medidas de isolamento Mais de 70% sabem que a terra é redonda #Somos70porcento	9154	6477	1296	1381
30/05/2020 16:02:52	randolfeap	Somos os 70% que rejeitam Bolsonaro! Fazemos parte dos 70% que apoiam o isolamento e a vida! Somos 70% de pessoas, cidadãos que querem ver o Brasil crescendo e não indo pro buraco c/ Bolsonaro no poder! Não precisamos de robôs, somos 70% do Brasil de...	9103	3282	5169	652
31/05/2020 20:16:41	MarceloFreixo	A escolha é entre bandeiras neonazistas e democracia. De que lado você está? #Somos70porcento	8893	4571	3780	542
30/05/2020 18:59:08	carlosjordy	A #Somos57MILHOES passou a tag dos derrotados que querem 3o turno. Entendam, vocês perderam nas urnas e vão continuar perdendo em tudo!	8787	5558	2444	785
30/05/2020 21:37:13	FlavioBolsonaro	Tem que aturar!!! #Somos57MILHOES https://t.co/8LD4GZIwGv	8733	4505	3355	873
30/05/2020 22:22:11	CarlaZambelli38	A esquerda pode até tentar. Mas não adianta, estamos em maior número e somos apaixonados pelo nosso Brasil. #Somos57MILHOES Beep-beep https://t.co/Zj0SFE2JhV	8642	3765	4130	747

E. Dados Complementares da Seção 4.4

Este apêndice é um complemento aos experimentos realizados na Seção 4.4. A Tabela E.1 apresenta os tweets coletados sobre o auxílio emergencial com mais interações.

Tabela E.1: Tweets coletados sobre o auxílio emergencial com mais interações

Data de publicação	Autor	Texto do tweet	Total de interações	# de retweets	# de respostas	# de citações
25/06/2020 23:29:24	MarceloFreixo	É DESUMANO @jairbolsonaro estender por + 3 meses apenas o auxílio emergencial. E ainda pagando parcelas de R\$ 500, R\$ 400 e R\$ 300. Ele está abandonando milhões de famílias na pior crise da história! Nossa proposta é pagar R\$ 600 pelo menos até d...	3518	885	2478	155
28/06/2020 15:10:59	sapavegana	Precisamos falar sobre a pressão social que é fazer compras sozinha, passar no caixa e ter que colocar as coisas na esteira, guardar na sacola e pagar rápido o suficiente para não sentir culpa por ter gente esperando na fila	3413	2648	135	630
30/06/2020 09:14:09	motta1302	PF finaliza inquérito e atribui a Maia corrupção, lavagem de dinheiro e 'caixa três' https://t.co/JOXKt1fRYh	1939	1450	238	251
29/06/2020 18:17:16	Reinatos	eu tô só essa caixa de morango ultimamente https://t.co/eB2zk93hAG	1471	1345	15	111
29/06/2020 16:33:58	bobjeffcensored	@blogdojefferson's account has been withheld in Brazil, Worldwide in response to a legal demand. Learn more.	1400	1249	107	44
30/06/2020 09:15:03	TechTudo	iPhone 12 deve vir sem carregador e sem fones de ouvido na caixa https://t.co/s2s30YyX1o https://t.co/c0MbAPrnwK	1238	58	182	998
04/07/2020 01:50:46	Caixa	Já são 65,1 milhões de brasileiros beneficiados pelo #AuxílioEmergencial do @govbr . Confira todos os resultados do nosso atendimento nesta thread. Conte com a CAIXA e lembre-se: não precisa ir até uma agência, faça tudo através dos nossos canais...	984	19	934	31
02/07/2020 10:20:13	Caixa	Fique atento! Hoje tem o pagamento do Auxílio Emergencial para os beneficiários nascidos em julho e agosto. #auxilioemergencial #caixa @govbr https://t.co/j61vLJHlzn	871	50	785	36
29/06/2020 23:26:46	ifobscuro	Caixa de email: você tem um novo trabalho Eu: https://t.co/EhhmB0PnVK	727	688	9	30
29/06/2020 18:19:34	zecajojo	Gosto muito desse conceito da @Caixa de colocar uma fila no aplicativo Caixa Tem. É uma experiência de usuário muito unica. Ele realmente faz vc se sentir numa agencia física esperando horas e horas. Muito realístico. https://t.co/RM3evbmUef	706	549	87	70

F. Dados Complementares da Seção 4.5

Este apêndice é um complemento aos experimentos realizados na Seção 4.5. A Tabela F.1 apresenta os tweets coletados sobre o ENEM 2020 com mais interações.

Tabela F.1: Tweets Coletados sobre o ENEM 2020 com mais interações

Data de publicação	Autor	Texto do tweet	Total de interações	# de retweets	# de respostas	# de citações
08/07/2020 20:19:46	Itspedrito	O INEP fez uma enquete pra estudantes votarem pra quando prefeririam a data do ENEM... estudantes: MAIO INEP: blz o ENEM vai ser em janeiro	12175	10836	161	1178
05/06/2020 23:05:37	augustopbarba	Enem 2021 pode ser cancelado Estudantes do 2º ano: https://t.co/4Rn2sg0prD	9582	7636	288	1658
08/07/2020 20:17:53	gl	#URGENTE - Prova do Enem 2020 será em janeiro de 2021 #G1 https://t.co/tFUF9n7sPr	9174	2474	747	5953
05/06/2020 23:27:52	kkauan_	Estudantes: fazem protesto pro Enem 2020 ser adiado* Ministério da educação: adia* Ministério da educação: agr vou cancelar ENEM 2021 Estudantes: https://t.co/tCBFeapZRg	8209	7828	60	321
09/07/2020 18:50:08	bylucasmachado	Esse é a última peça publicitária do MEC na gestão @dilmabr. Olhem a diferença de abordagem, de mensagem e sobretudo, de conexão com a realidade de muitos brasileiros. Saudade de quando o @MEC_Comunicacao era Ministério da Educação. https://t.co/...	7903	4672	504	2727
05/06/2020 20:37:20	Estadao	EXCLUSIVO: Ministério da Educação alerta que Enem pode ser suspenso em 2021 devido à falta de recursos https://t.co/VNqV2wou7G https://t.co/rxJL7GrjBH	7701	1184	755	5762
02/07/2020 20:39:43	jairbolsonaro	Parecer AGU JL-05 (jul/20) esclarece que escolas militares (EN, EsPCEEx, EPCAR e CMs) são públicas p/ todos os fins, inclusive p/ acesso às universidades pelo SISU. Por analogia, as escolas militares de outros entes (PMs e CBMs) terão o mesmo tratamen...	6769	5006	1632	131
10/07/2020 17:42:16	_goodend	Débora Aladim fez um curso com valor simbólico de 1,99 pra ajudar estudantes a estudarem pro ENEM e ofereceu ajuda pros que não podiam pagar. Além disso, tem sido mto ativa nesse período de incertezas da educação no Brasil. Mas o que importa msm s...	6362	5746	143	473
08/07/2020 20:14:27	lucasfelpi	O INEP ANUNCIOU A NOVA DATA DO ENEM: 17 e 24 de janeiro de 2021	6023	3665	296	2062
05/06/2020 23:49:28	luscafp	"Enem 2021 pode ser cancelado"Propaganda enem 2020: https://t.co/iaqiDEP0WV	5747	5452	53	242