

MEDINDO A INFLUÊNCIA DE USUÁRIOS EM DEBATES POLÍTICOS NA PLATAFORMA *REDDIT*: UMA ANÁLISE MÉTRICA PARA DETECTAR SUBPARTIÇÕES MAIS POLARIZADAS EM UM GRAFO BIPARTIDO

Bryan Lima Cavalcante*
Henrique Viana Oliveira**

RESUMO

As redes sociais vêm recebendo cada vez mais atenção de usuários que frequentemente compartilham suas opiniões por meio delas. Dentro delas, a polarização em debates *online*, em temas como política, questões de gêneros, aborto ou controle de armas surgem como desafios significativos em diversos cenários. Este trabalho se concentra na plataforma da rede social *Reddit* e apresenta um modelo quantitativo para medir a polarização em seus debates *online*. Utilizando uma abordagem baseada em grafos bipartidos representando debates, um algoritmo é utilizado para identificar partições com maior polarização. Além disso, investigamos a detecção de subpartições, levando em consideração a relevância de usuários com base em métricas de influência social dentro do *Reddit*. Neste contexto, a análise das subpartições revela não apenas a polarização em si, mas também aspectos relacionados à influência dos usuários e da compreensão mais abrangente da dinâmica e estrutura social subjacente aos debates *online*.

Palavras-chave: 1. Polarização de Opinião 2. *Reddit* 3. Influência Social.

ABSTRACT

Social networks are receiving more and more attention from users who frequently share their opinions through them. Within them, polarization in online debates, on topics such as politics, gender issues, abortion or gun control emerge as significant challenges in different scenarios. This work focuses on the social network platform *Reddit* and presents a quantitative model to measure polarization in its online debates. Using an approach based on bipartite graphs representing debates, an algorithm identifies partitions with higher polarization. Additionally, we investigated subpartition detection, taking into account the relevance of users based on metrics of social influence within *Reddit*. In this context, the analysis of the subpartitions reveals not only the polarization itself, but also aspects related to the influence of users and the broader understanding of the dynamics and social structure underlying online debates.

Keywords: 1. Opinion Polarization 2. *Reddit* 3. Social Influence.

* Graduando em Bacharelado em Ciência da Computação, Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE), Aracati, Ceará, Brasil. E-mail: bryan.contato.cavalcante@gmail.com

** Doutor em Ciência da Computação, Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE), Aracati, Ceará, Brasil. E-mail: henrique.viana.oliveira@gmail.com

1 INTRODUÇÃO

A quantidade de usuários da internet em todo o mundo tem crescido de maneira significativa ao longo dos anos. Contando atualmente com mais de 5.16 bilhões de pessoas ativas ¹, a internet vem alterando significativamente a forma como as pessoas se comunicam, interagem e compartilham informações. As redes sociais, contribuem para esse fenômeno, pois vieram com a proposta de potencializar a disseminação de opiniões individuais e coletivas.

A redes sociais têm se destacado como locais onde as discussões podem acontecer em uma escala global e envolver uma ampla gama de participantes. No entanto, a dinâmica das redes sociais também levanta questões sobre como as discussões são moldadas, moderadas e influenciadas pela estrutura das plataformas, bem como pelos diferentes tipos de participantes que nela se envolvem. A análise das características desses debates *online*, incluindo a estrutura das interações, os tópicos discutidos e as fontes de informação referenciadas, é crucial para entender como as ideias são disseminadas, como as opiniões são formadas e como a polarização pode ocorrer em ambientes digitais. Hoje em dia, há uma controvérsia crescente sobre o surgimento da polarização nas discussões em redes sociais e na responsabilidade das pessoas neste problema.

Nas Ciências Sociais, a polarização é o processo social pelo qual um grupo social ou político é dividido em dois subgrupos opostos com posições, objetivos e pontos de vista conflitantes e contrastantes, com poucos indivíduos permanecendo neutros ou mantendo uma posição intermediária (ISENBERG, 1986; SUNSTEIN, 1999). Um domínio típico onde a polarização é presente é a Política (DIXIT; WEIBULL, 2007; WAUGH et al., 2009), embora se saiba que uma série de outras questões induzem na sociedade um debate divisivo que muitas vezes cria uma fração de pessoas tem opiniões muito extremas, como aquecimento global (MCCRIGHT; DUNLAP, 2011), controle de armas (MILLER, 2019), casamento entre pessoas do mesmo sexo (FLORES; BARCLAY, 2016) e aborto (MOUW; SOBEL, 2001; EVANS, 2002).

Como esperado, a ascensão dos sistemas de mídias sociais rapidamente transformou a *web* em uma plataforma de debates e discussões animadas (WEINBERGER, 2011); agora as batalhas *online* são travadas sobre questões polarizadoras e polêmicas, especialmente quando novas evidências que apoiam um lado da discussão surgem.

Neste contexto, é relevante fornecer um modelo mais claro e quantitativo para medir a polarização em debates *online*, de modo que esse comportamento possa ser monitorado, gerando um sinal de alerta sempre que necessário. Falando de outro modo, podemos detectar padrões de comunicação onde os usuários parecem interagir positivamente apenas com um grupo fixo de usuários e negativamente com o resto. Outro importante tópico de pesquisa em debates é entender como diferentes usuários interagem entre si e tentar quantificar sua relevância em uma rede social. Por exemplo, a influência de um usuário pode ser medida através do número de seus seguidores, o número de mensagens e o número de menções de outros usuários a este usuário específico (CHA et al., 2010). De maneira semelhante, com o objetivo de distinguir entre líderes

¹ <https://datareportal.com/reports/digital-2023-global-overview-report>

e seguidores nas redes sociais, diferentes características são estudadas para tentar caracterizar tais perfis de usuários (SHAFIQ et al., 2013).

Visto isso, este trabalho consiste em duas principais tarefas. A primeira visa trabalhar com uma estrutura de análise de polarização na rede social *Reddit*. Esta estrutura modela um debate no *Reddit* como um grafo ponderado e com arestas rotuladas, onde os pesos dos nós representam o lado da opinião dos usuários no debate e os rótulos das arestas representam os sentimentos gerais entre as opiniões dos usuários. O grafo é dividido em duas partições, no qual é quantificado o grau de polarização medindo quão homogênea é cada partição e quão negativas são as interações entre ambas partições. Então, o modelo quantitativo é baseado na polarização máxima e é calculado com um algoritmo de busca local gulosa. A segunda tarefa consiste em detectar possíveis subpartições em que a polarização seja maior que as duas partições originais, considerando os usuários mais influentes do debate para essa tarefa. Para detectar quais os usuários mais influentes, serão utilizadas medidas de acordo com o *Reddit*, como *karma*, atividade, atratividade e grau de centralidade dos usuários.

O presente trabalho está estruturado em sequência. A Seção 2 apresenta o referencial teórico, onde são explicados alguns conceitos e teorias sobre modelos de debates, que servem como base para este estudo. A Seção 3 detalha os trabalhos relacionados, envolvendo a polarização em debates, que se aparentam com a ideia proposta para este trabalho. Na Seção 4, é apresentada a metodologia do trabalho, onde é explicado todo o processo da construção do modelo. A Seção 5 apresenta os resultados alcançados do trabalho. Finalmente, a Seção 6 conclui o trabalho com considerações finais e seus trabalhos futuros.

2 REFERENCIAL TEÓRICO

Nesta seção é apresentada a fundamentação teórica necessária para um aprofundamento melhor do entendimento da proposta do trabalho. Ela compreende conceitos gerais sobre a plataforma da rede social *Reddit* e de Grafo de Debate de Usuário, o modelo matemático utilizado para expressar debates no *Reddit* e para calcular o grau de polarização desses debates.

2.1 *Reddit*

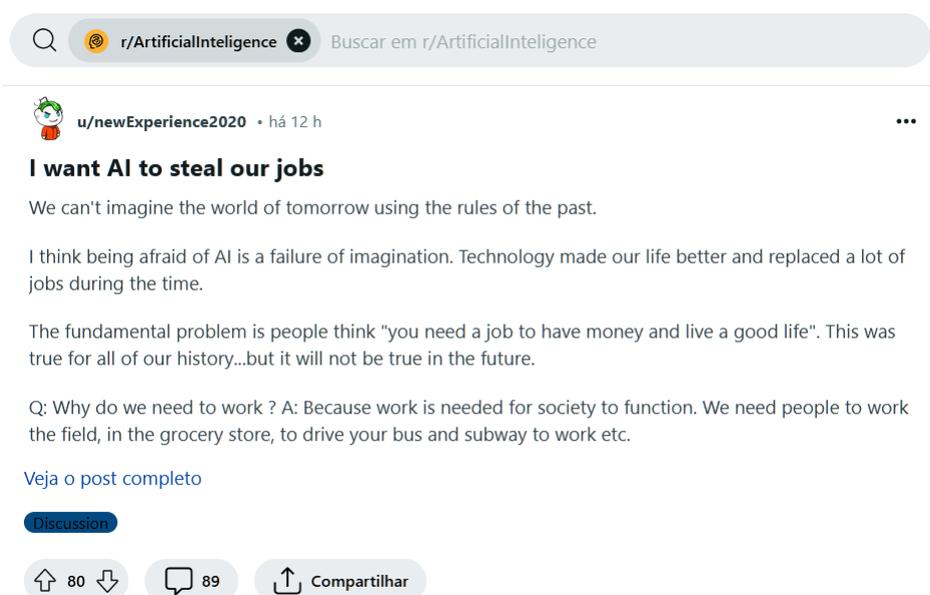
O *Reddit* é uma rede de comunidades *online* que permite aos usuários participar de discussões e avaliar conteúdos postados por outras pessoas. Fundado em 2005 por Steve Huffman e Alexis Ohanian, o site hospeda muitas comunidades interativas que impulsionam conversas na *web*². As comunidades do site são chamadas de *subreddits* e cobrem um número aparentemente infinito de tópicos. Esses tópicos podem variar de assuntos cotidianos, incluindo política, esportes e estilos de vida, até coisas que só podem ser descritas como bizarras. Os usuários podem compartilhar postagens de texto, *links* para notícias e outros conteúdos, vídeos e imagens relacionados ao *subreddit* em que estão.

² <https://medium.com/digital-marketing-lab/reddit-overview-f475ef13aae7>

Todos os *subreddits* começam com “r” – a estrutura de URL que o *Reddit* usa para denotá-los. Alguns *subreddits* populares incluem “r/askreddit”, no qual os usuários fazem e respondem perguntas instigantes; “r/todayilearned”, onde os usuários enviam fatos interessantes que descobriram recentemente; e “r/aww”, onde os usuários postam fotos e vídeos de coisas fofas, principalmente animais. Os tópicos variam muito, desde o simples “r/music” até o maior depósito de humor “r/funny”.

Cada *subreddit* tem moderadores responsáveis por estabelecer regras e garantir que a comunidade funcione adequadamente. Os usuários, conhecidos como “*Redditors*”, podem participar de um *subreddit* para realizar postagens e comentários nas postagens de outras pessoas. Cada postagem de um usuário pode ser votada com uma seta para cima (*upvote*) ou com uma seta para baixo (*downvote*) pelos membros da comunidade (ver Figura 1), influenciando a visibilidade do conteúdo. Cada postagem pode ter uma seção de comentários onde os usuários podem discutir o conteúdo, e os comentários também podem ser votados com seta para cima ou para baixo.

Figura 1 – Postagem no *Reddit*.



Fonte: Elaborado pelos autores.

Além disso, o *Reddit* usa um sistema de pontuação chamado *karma*³, que os usuários acumulam através dos *upvotes* em suas postagens e comentários (ver Figura 2). Embora o *karma* não tenha um valor prático, muitos usuários o veem como uma medida de sua contribuição para a comunidade. Além disso, existe um conjunto informal de regras de etiqueta chamado “*Reddiquette*”⁴ que os usuários são incentivados a seguir para manter um ambiente saudável e respeitoso.

³ <https://support.reddithelp.com/hc/en-us/articles/204511829-What-is-karma->

⁴ <https://support.reddithelp.com/hc/en-us/articles/205926439-Reddiquette>

Figura 2 – Karma de um usuário no *Reddit*.

Fonte: Elaborado pelos autores.

2.2 Modelo de Medida de Polarização em Debates no *Reddit*

Para examinar a concordância e discordância entre os comentários dos usuários em debates no *Reddit*, no trabalho de (ALSINET et al., 2021) foi desenvolvido um sistema de análise que representa um debate no *Reddit* como um grafo de debate bilateral. Nesse contexto, os comentários são categorizados em dois grupos: aqueles que expressam concordância com o comentário raiz do debate e aqueles que discordam. As arestas do grafo indicam a presença de concordância e discordância entre os comentários dos dois grupos. Desta maneira, neste estudo, adotamos o modelo, visando investigar possíveis subpartições em que a polarização seja maior, considerando os usuários mais influentes do debate.

2.2.1 Grafos de Debates

Definição 1 (Comentário e Debate no *Reddit*) Um comentário no *Reddit* é definido como uma tupla $c = (m, u)$, onde m é o texto do comentário e u é identificador do usuário. Sejam $c_1 = (m_1, u_1)$ e $c_2 = (m_2, u_2)$ comentários. Dizemos que c_1 responde a c_2 se c_1 é uma resposta ao comentário c_2 . Seja $r = (m_r, u_r)$. Um debate no *Reddit* em r é um conjunto não vazio Γ de comentários do *Reddit* tal que $r \in \Gamma$ e cada comentário em $c \in \Gamma$, tal que $c \neq r$, dizemos que r é o comentário raiz de Γ , c responde a um comentário anterior em Γ .

O sistema de debates do *Reddit* constrói uma árvore de debate, com comentários como nós e onde as arestas denotam respostas entre os comentários. Além disso, esta fase atribui rótulos para as arestas da árvore no intervalo real $[-2, 2]$. O rótulo para uma aresta (c_1, c_2) denota o sentimento expresso no texto do comentário c_1 em resposta ao texto do comentário c_2 , de modo que o valor -2 denota uma discordância total e o valor 2 um acordo total. O valor de sentimento 0 denota respostas que expressam uma posição neutra.

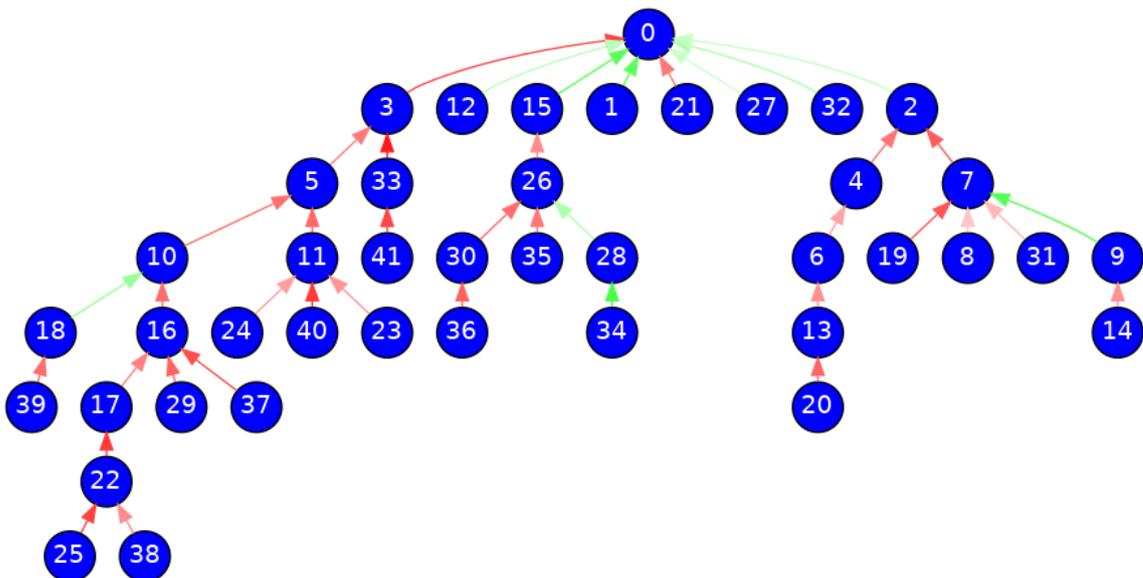
Definição 2 (Árvore de Debate) Seja Γ um debate do *Reddit* sobre um comentário raiz r . Uma Árvore de Debate (*DebT*) para Γ é uma tupla $T = (C, r, E, W)$ tal que o seguinte é verdadeiro:

- Para cada comentário em Γ , há um nó $c \in C$.

- O nó $r \in C$ é o nó raiz de T .
- Se c_1 responde a c_2 , então há uma aresta direcionada $(c_1, c_2) \in E$, e W é uma função de rotulação das arestas $W : E \rightarrow [-2, 2]$.

Na Figura 3, é mostrada a estrutura $DebT$ para uma Árvore de Debate no *Reddit*. Cada comentário é representado como um nó e cada resposta entre comentários é representado como uma aresta. O grafo tem 42 nós e 41 arestas desde cada comentário responde exatamente a um comentário, exceto o comentário raiz. Estas 41 respostas entre comentários são classificadas como arestas de concordância (pintadas em verde) e discordância (pintadas em vermelho) e neste exemplo não há resposta neutra. A instensidade da cor é diretamente proporcional ao sentimento da resposta em relação ao valor máximo. O comentário raiz da discussão é rotulado com 0, e os outros comentários são rotulados com identificadores consecutivos de acordo com sua ordem de geração. Observe que o grafo é acíclico já que um comentário apenas responde a um comentário anterior na discussão.

Figura 3 – Exemplo de Árvore de Debate no *Reddit*.



Fonte: (ALSINET et al., 2021).

Analisando as relações entre os nós do $DebT$, os comentários são classificados em dois grupos: comentários que apoiam o comentário raiz e comentários que discordam dele. Com esse objetivo, a estrutura $DebT$ é estendida com uma função de rotulagem de nós que denotam o lado de cada comentário no debate. Os comentários que suportam o comentário raiz são rotulados com 1 e o restante dos comentários com -1, ou seja, os comentários neutros e que desaprovam o comentário raiz. Esta extensão de árvore é referida como uma Árvore de Debate Bilateral.

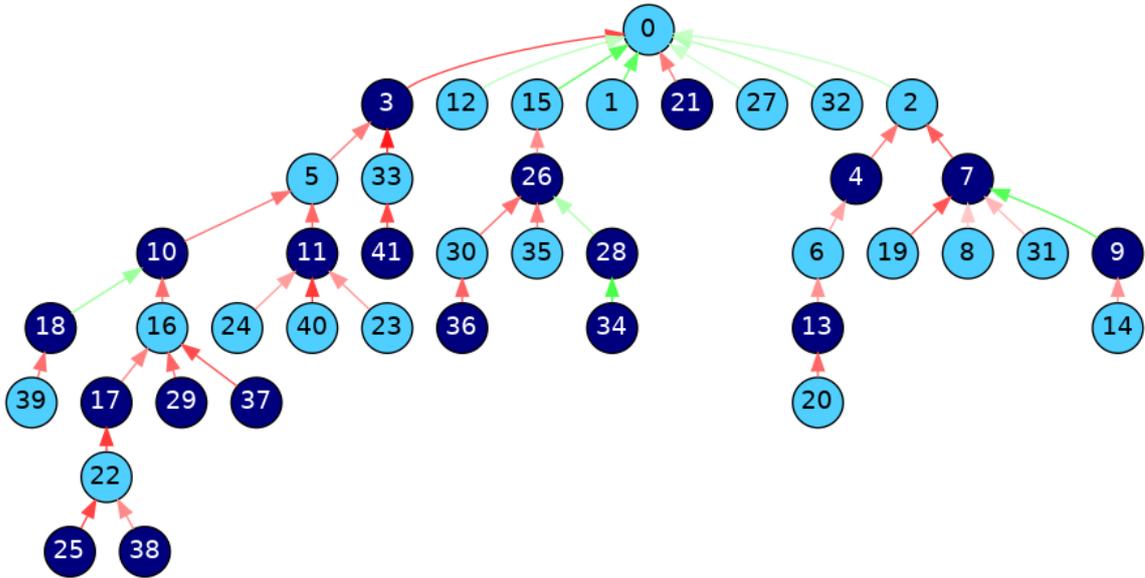
Definição 3 (Árvore de Debate Bilateral) *Seja $T = (C, r, E, W)$ uma Árvore de Debate ($DebT$) para um debate no *Reddit* sobre um comentário raiz r . Uma Árvore de Debate Bilateral ($SDebT$)*

para T é uma tupla $TS = (C, r, E, W, S)$ tal que S é uma função de rotulação dos nós $S : C \rightarrow [-1, 1]$ definida da seguinte forma:

- $S(r) = 1$.
- Para todo nó $c_1 \neq r$ em C , $S(c_1) = 1$ se, para algum nó $c_2 \in C$, $(c_1, c_2) \in E$ e qualquer um dos seguintes critérios for satisfeito: $S(c_2) = 1$ e $W(c_1, c_2) > 0$, ou $S(c_2) = -1$ e $W(c_1, c_2) \leq 0$; caso contrário, $S(c_1) = -1$.

A Figura 4 mostra a estrutura $SDebT$ que é obtida para a Árvore de Debate da Figura 3. O nós (comentários) que suportam o comentário raiz estão coloridos em ciano, enquanto o restante está colorido em azul marinho.

Figura 4 – Exemplo de Árvore de Debate Bilateral no *Reddit*.



Fonte: (ALSINET et al., 2021).

A seguir, é formalizado o grafo para representar os debates baseados nos usuários, denominado Grafo de Debate de Usuário, onde cada nó representa a média das opiniões do usuário referente ao comentário raiz e as arestas denotam interações entre usuários extraídos do sentimento predominante entre os comentários agregados dos nós.

Definição 4 (Grafo de Debate de Usuário) Seja Γ um Debate no *Reddit* sobre um comentário raiz r com identificadores de usuários $U = \{u_1, \dots, u_m\}$, e seja $TS = (C, r, E, W, S)$ uma $SDebT$ para Γ . Um Grafo de Debate de Usuário ($UDebG$) para TS é uma tupla $\mathcal{G} = (\mathcal{C}, \mathcal{E}, \mathcal{S}, \mathcal{W})$, onde:

- \mathcal{C} é o conjunto de nós de \mathcal{G} definido como o conjunto de opiniões dos usuários $\{C_1, \dots, C_m\}$, ou seja, $\mathcal{C} = \{C_1, \dots, C_m\}$, com $C_i = \{(m, u_i) \in \Gamma \mid (m, u_i) \neq r\}$, para todos os usuários $u_i \in U$.

- $\mathcal{E} \subseteq \mathcal{C} \times \mathcal{C}$ é o conjunto de arestas de \mathcal{G} definido como o conjunto de interações entre diferentes usuários no debate, i.e., há uma aresta $(C_i, C_j) \in \mathcal{E}$, com $C_i, C_j \in \mathcal{C}$ e $i \neq j$, se e somente se, para algum $(c_1, c_2) \in E$, temos $c_1 \in C_i$ e $c_2 \in C_j$.
- \mathcal{S} é um esquema de ponderação de opiniões para \mathcal{C} , que expressa o lado dos usuários no debate com base no lado de seus comentários. Definimos \mathcal{S} como o mapeamento $\mathcal{S} : \mathcal{C} \rightarrow [-1, 1]$ que atribui a cada nó $C_i \in \mathcal{C}$ o valor

$$\mathcal{S}(C_i) = \frac{\sum_{c \in C_i} \mathcal{S}(c_i)}{|C_i|}$$

no intervalo real $[-1, 1]$, que expressa o lado do usuário u_i em relação ao comentário raiz, desde discordância estrita (-1) até concordância estrita (1), passando por opiniões neutras (0).

- \mathcal{W} é um esquema de ponderação de interações para \mathcal{E} , que expressa tanto a proporção de interações positivas entre as opiniões dos usuários quanto o sentimento geral entre os usuários, combinando os valores individuais de sentimento atribuídos às respostas entre seus comentários. Seja \oplus um operador de agregação de dimensão n no intervalo real $[-2, 2]$, ou seja, um mapeamento limitado, monotônico, simétrico e idempotente $\oplus : [-2, 2]^n \rightarrow [-2, 2]$. Definimos \mathcal{W} como o mapeamento $\mathcal{W} : \mathcal{E} \rightarrow ([0, 1] \times [-2, 2])$ que atribui a cada aresta $(C_i, C_j) \in \mathcal{E}$ o par de valores $(p, w) \in ([0, 1] \times [-2, 2])$ definidos da seguinte forma:

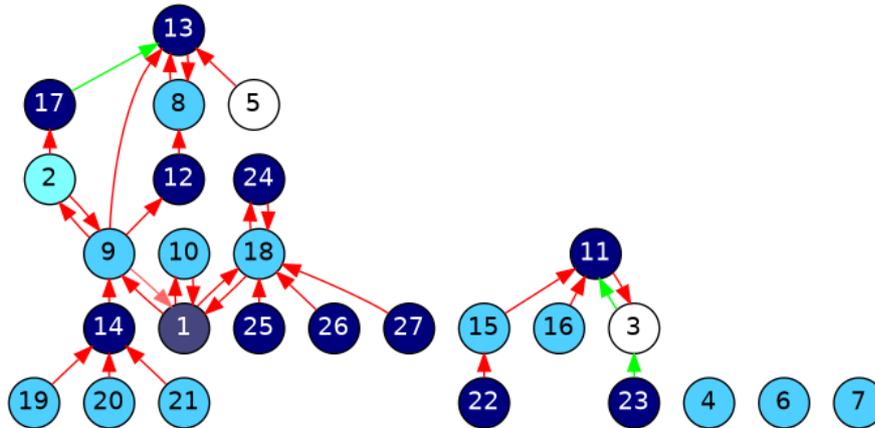
$$p = \frac{|(c_1, c_2) \in E \cap (C_i \times C_j) \text{ com } W(c_1, c_2) > 0|}{|(c_1, c_2) \in E \cap (C_i \times C_j)|} e$$

$$w = \oplus_{\{(c_1, c_2) \in E \cap (C_i \times C_j)\}} W(c_1, c_2).$$

onde p expressa a proporção de respostas positivas do usuário u_i para o usuário u_j no debate, e w expressa o sentimento geral do usuário u_i em relação aos comentários do usuário u_j , desde altamente negativo (-2) até altamente positivo (2).

A Figura 5 mostra a estrutura $UDebG$ obtido para a Árvore de Debate Bilateral da Figura 4. A opinião de cada usuário é representada como um nó, e cada relacionamento entre eles é representado como uma aresta. O grafo possui 27 nós e 31 arestas; o sentimento geral para as arestas são avaliadas usando algum operador de agregação (DETYNIECKI, 2001) (como por exemplo média, máximo ou mínimo) o que levou a três relações de concordância (em verde), com o resto sendo desacordo (em vermelho). A quantidade de nós do $UDebG$ tende a ser menor do que $SDebT$ pois um usuário pode responder a um outro usuário várias vezes durante um debate e em $UDebG$ essa discussão será representada apenas por uma aresta avaliada pela agregação de todas as arestas entre esses usuários.

Figura 5 – Exemplo de Grafo de Debate de Usuário no *Reddit*.



Fonte: (ALSINET et al., 2021).

2.2.2 Polarização de Debates

Dado um Grafo de Debate de Usuário \mathcal{G} , em (ALSINET et al., 2021) foi proposto um modelo para medir o nível de polarização no debate entre seus usuários. Nele, um debate polarizado contém uma bipartição de \mathcal{C} em dois conjuntos (L, R) de modo que o conjunto L contenha principalmente usuários em desacordo e o conjunto R contém principalmente usuários de acordo e ambos os conjuntos devem ter tamanho semelhante. A segunda característica é o sentimento entre os usuários de L e R .

Para capturar essas duas características com um único valor, foram definidas diferentes medidas e suas combinações em um nível final denominado nível de polarização bipartida. Primeiro, dada uma bipartição (L, R) de \mathcal{C} , é definido seu nível de consistência e quão balanceado os tamanhos de L e R como:

$$SC(L, R, \mathcal{G}) = LC(L, \mathcal{G}) \cdot RC(R, \mathcal{G}),$$

onde $LC(L, \mathcal{G})$ avalia a força de discordância dos comentários em L (em relação ao comentário raiz) e $RC(R, \mathcal{G})$ a força de concordância dos comentários em R . Eles são definidos da seguinte forma:

$$LC(L, \mathcal{G}) = \frac{\sum_{C_i \in L, \mathcal{S}(C_i) \leq 0} -\mathcal{S}(C_i)}{|\mathcal{C}|} \quad \text{e} \quad RC(R, \mathcal{G}) = \frac{\sum_{C_i \in R, \mathcal{S}(C_i) > 0} \mathcal{S}(C_i)}{|\mathcal{C}|}.$$

O valor de $SC(L, R, \mathcal{G})$ está no intervalo $[0, 0.25]$. O valor mínimo (0) é alcançado quando não há usuários em L com $\mathcal{S}(C_i) < 0$ ou nenhum em R com $\mathcal{S}(C_i) > 0$. O valor máximo (0.25) é alcançado quando metade dos usuários estão em desacordo com $\mathcal{S}(C_i) = -1$ e são encontrados em L , e a outra metade estão de acordo com $\mathcal{S}(C_i) = 1$ e são encontrados em R .

Em segundo lugar, o sentimento das interações entre usuários de diferentes lados é definido do seguinte modo:

$$SWeight(L, R, \mathcal{G}) = \frac{\sum_{e_i \in \mathcal{E} \cap (L \times R) \cup (R \times L)} -c(p(e_i)) \cdot w(e_i)}{|\mathcal{E}|} + 2.$$

com $c(p(e_i)) = 2(p(e_i) - 0.5)^2 + 0.5$, e onde $p(e_i)$ e $w(e_i)$ denotam os valores de p e w , respectivamente, em $\mathcal{W}(e_i) = (p, w)$. Vale ressaltar que o valor de $SWeight(L, R, \mathcal{G})$ está no intervalo $[0, 4]$. O valor mínimo (0) é alcançado quando todas as interações $e_i \in E$ são interações puramente positivas com valor de sentimento máximo, e todas elas são encontradas entre L e R . De maneira análoga, o valor máximo (4) é alcançado quando todas as interações $e_i \in E$ são interações puramente negativas com valor de sentimento mínimos entre L e R .

Finalmente, as medidas $SC(L, R, \mathcal{G})$ e $SWeight(L, R, \mathcal{G})$ são combinadas para definir o nível de polarização bipartida de uma determinada bipartição (L, R) como segue:

$$BipPol(L, R, \mathcal{G}) = SC(L, R, \mathcal{G}) \cdot SWeight(L, R, \mathcal{G}).$$

Assim, dado o intervalo de valores das medidas anteriores, o valor de $BipPol(L, R, \mathcal{G})$ está no intervalo $[0, 1]$:

- O valor mínimo (0) é alcançado quando $SC(L, R, \mathcal{G}) = 0$ ou $SWeight(L, R, \mathcal{G}) = 0$. Se isso acontecer com $SC(L, R, \mathcal{G}) = 0$ e $SWeight(L, R, \mathcal{G}) > 0$ então em L , temos apenas usuários neutros e positivos, pelo menos um usuário em R e algumas interações entre L e R que não são extremamente positivos (de tal forma que $SWeight(L, R, \mathcal{G})$ pode ser positivo).
- O valor médio (0.5) pode, obviamente, ser alcançado para muitas combinações destes dois fatores, mas há dois casos canônicos que refletem bem o significado pretendido de este caso intermediário. São os casos em que um dos dois fatores tem seu máximo valor e o outro tem seu valor médio. Isso pode acontecer em dois casos. Primeiro, isso acontece quando $SC(L, R, \mathcal{G}) = 0.25$ e $SWeight(L, R, \mathcal{G}) = 2$. Ou seja, quando os usuários em desacordo e usuários em acordo estão perfeitamente equilibrados entre L e R . Em segundo lugar, isso também acontece quando $SC(L, R, \mathcal{G}) = 0.25 \cdot 0.25$ e $SWeight(L, R, \mathcal{G}) = 4$. Ou seja, ambas as partições possuem uma representação equilibrada de usuários em acordo e desacordo, e temos interações negativas puras entre usuários de acordo com uma partição e usuários em desacordo com a outra.
- O valor máximo (1) é alcançado quando ambos os fatores possuem seu valor máximo. O equilíbrio perfeito entre usuários em desacordo em L e usuários em concordância em R ($SC(L, R, \mathcal{G}) = 0.25$) e todas as interações possuem $\mathcal{W}(e_i) = (0, -2)$ e são encontradas apenas entre L e R . Este caso extremo representa a polarização máxima, uma situação em que existem dois lados homogêneos perfeitos no debate (lados de acordo e desacordo) com sentimentos que são maximamente opostos (+1 e -1) e a única conversa entre deles é criticar o outro lado com força máxima, para que não haja respostas entre os dois lados.

2.2.3 Encontrando a Partição mais Polarizada

Aqui é apresentado um algoritmo guloso de otimização de busca local para encontrar uma bipartição (L, R) de $\mathcal{G} = (\mathcal{C}, \mathcal{E}, \mathcal{S}, \mathcal{W})$ com um valor alto para $BipPol(L, R, \mathcal{G})$ que segue a mesma ideia básica de algoritmos guloso comuns para o problema *maxcut* (TREVISAN, 2009), com o objetivo de encontrar uma partição com um valor de polarização próximo ao mais alto.

O algoritmo segue estas etapas:

1. Inicialize os vértices em L ou R . A ideia mais óbvia é ordenar os vértices por polaridade e, em seguida, atribua os negativos a L e o restante a R . Este passo pode ser feito de várias formas, mas a mais simples consiste de colocar cada usuário de maneira uniforme e aleatória em L ou R .
2. Execute etapas de melhoria local (em relação ao valor do *BipPol*): a cada iteração, encontrar um bom candidato de L (para ser movido para R) e um bom candidato de R (para ser movido para L).

O pseudocódigo é mostrado no Algoritmo 1. Ele é usado para encontrar uma solução local ótima para uma bipartição de $\mathcal{G} = (\mathcal{C}, \mathcal{E}, \mathcal{S}, \mathcal{W})$ com alta polarização bipartida.

Algoritmo 1 $BipPol(\mathcal{G})$

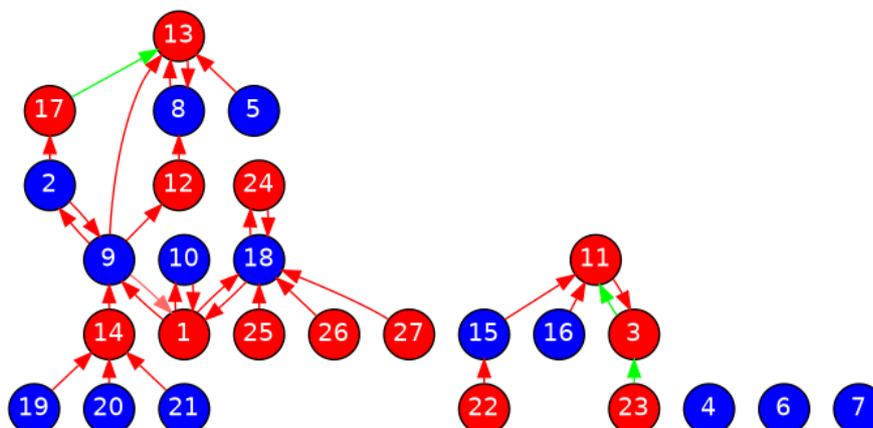
```

1: Entrada:  $\mathcal{G} = (\mathcal{C}, \mathcal{E}, \mathcal{S}, \mathcal{W})$ 
2: Saída: Bipartição  $(L, R)$  de  $\mathcal{G}$  com alta polarização bipartida
3:  $(L, R) \leftarrow getInitBPart(\mathcal{G})$ 
4:  $improving \leftarrow true$ 
5:  $steps \leftarrow 0$ 
6: while  $improving$  and  $steps \leq maxsteps$  do
7:    $improving \leftarrow false$ 
8:   if  $\exists v \in L, BipPol(L \setminus v, R \cup v, \mathcal{G}) > BipPol(L, R, \mathcal{G})$  then
9:      $(L, R) \leftarrow (L \setminus v, R \cup v)$ 
10:     $improving \leftarrow true$ 
11:   end if
12:   if  $\exists v_0 \in R, BipPol(L \cup v_0, R \setminus v_0, \mathcal{G}) > BipPol(L, R, \mathcal{G})$  then
13:      $(L, R) \leftarrow (L \cup v_0, R \setminus v_0)$ 
14:      $improving \leftarrow true$ 
15:   end if
16:    $steps \leftarrow steps + 1$ 
17: end while
18: return  $(L, R)$ 

```

A Figura 6 mostra um exemplo de resultado da aplicação do Algoritmo 1 utilizando o UDebG da Figura 5 como entrada. Os nós na cor vermelha são os nós resultantes do conjunto L e os nós na cor azul são os nós resultantes do conjunto R .

Figura 6 – Exemplo de Aplicação de Polarização Máxima.



Fonte: (ALSINET et al., 2021).

3 TRABALHOS RELACIONADOS

Esta seção concentra trabalhos relacionados com polarização em redes sociais. Diferente da proposta apresentada por (ALSINET et al., 2021), que é voltada na construção de um modelo léxico e na noção de Grafos de Debate de Usuário, os trabalhos a seguir apresentam abordagens baseadas em Aprendizado de Máquina e Modelos Híbridos.

No trabalho (AMIN et al., 2017), os autores apresentam um algoritmo de fatorização de matriz e descida de gradiente baseada em conjunto para descobrir a polarização em redes sociais. O estudo avalia o algoritmo em contextos de disputas, conflitos e controvérsias, demonstrando uma capacidade de separar mensagens no *Twitter* de diferentes polaridades com uma precisão superior a 90%. Os resultados indicam um desempenho notável em comparação com técnicas supervisionadas, como análise de sentimento, e supera em cerca de 20% a 30% as abordagens de detecção de comunidades em cenários com uma grande presença de redes neutras. A discussão sobre a identificação correta de casos desafiadores, limites de confiança e a possível integração de informações adicionais para estimar conjuntamente a polaridade do *tweet* com sua veracidade também são abordados.

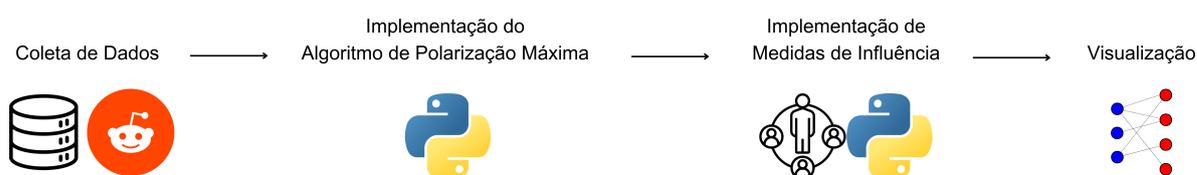
O artigo de (MORALES et al., 2015) propõe uma metodologia para estudar e medir o surgimento da polarização política em redes sociais, com foco na análise de dados gerados por usuários em plataformas como o *Twitter*. O autor introduz um modelo para estimar opiniões em que uma minoria de indivíduos influentes propaga suas opiniões por meio de uma rede social, influenciando as opiniões dos demais. O objetivo é inferir uma distribuição de opiniões na rede social e, a partir disso, medir a polarização política. O resultado do trabalho é uma função de densidade de probabilidade $p(X)$, que determina a fração de indivíduos que têm uma opinião X . O autor introduz um índice de polarização para medir o quão polarizada é a distribuição resultante de opiniões. Quando a população está polarizada, espera-se que a distribuição de opiniões seja bimodal, com dois picos em torno de opiniões dominantes e opostas.

Já no trabalho proposto por (AL-AYYOUB et al., 2018) é construída uma abordagem híbrida para identificar o nível de controvérsia em tópicos em tendência no *Twitter*, utilizando dados de mídias sociais escritos em árabe sobre eventos relevantes no Oriente Médio. A proposta envolve a combinação de abordagens que consideram tanto as interações na forma de um grafo quanto o conteúdo das mensagens trocadas. A ideia é superar as limitações das abordagens que se baseiam apenas em uma dessas perspectivas. O trabalho inclui a coleta de *tweets* relacionados a tópicos em tendência, a construção de um grafo de *retweets* para representar as relações de endosso entre os usuários envolvidos na conversa e a aplicação de medidas de controvérsia para estimar o grau de controvérsia de cada tópico. Além disso, o trabalho aborda o desafio específico do processamento de texto em árabe, incorporando análise de sentimento e detecção de postura para enriquecer a detecção de controvérsia.

4 METODOLOGIA

Nesta seção, descrevemos a metodologia adotada neste trabalho, que se concentra na aplicação de técnicas para mensurar a polarização e a influência dos usuários em debates *online*. O projeto foi conduzido seguindo as etapas delineadas na Figura 7. Essas etapas incluem (i) a coleta de dados do *Reddit*; (ii) a implementação do modelo de polarização máxima mencionado na Seção 2.2; (iii) a aplicação de medidas de influência e um algoritmo de busca linear que visa identificar as subpartições do grafo que maximizam a polarização, de acordo com a influência; e (iv) a visualização dos dados obtidos. Cada uma dessas etapas é detalhada a seguir.

Figura 7 – Fluxograma da Metodologia Proposta.



Fonte: Elaborado pelos autores.

4.1 Coleta de Dados

Inicialmente realizamos uma pesquisa no *Reddit* para encontrar um *subreddit* que apresentasse diversidade de perspectivas e informações compartilhadas pelos usuários. O *subreddit Political Discussion*⁵ se destaca como um espaço especialmente dedicado a debates políticos amplos e embasados. Os participantes desse *subreddit* compartilham notícias, análises e recursos relacionados à política, engajando-se em discussões abrangentes sobre eleições, políticas governamentais e questões sociais relevantes.

Com uma equipe de moderação ativa, esse *subreddit* busca promover discussões fundamentadas em evidências e argumentos respeitosos. Ele proporciona aos participantes uma

⁵ <https://www.reddit.com/r/PoliticalDiscussion/>

oportunidade valiosa de aprendizado, troca de ideias e aprofundamento do conhecimento sobre os assuntos políticos atuais.

Para a coleta de dados, utilizamos a ferramenta *PRAW*⁶ (*Python Reddit API Wrapper*). Essa ferramenta consiste em uma API que permite interagir com a plataforma *Reddit* de maneira programática. Ao fornecer um *link* de uma postagem, o *PRAW* retorna o conteúdo completo desta postagem, incluindo informações como ID, título, texto, sentimento e comentários. Os comentários, por sua vez, são estruturados como *arrays* contendo informações como ID, usuário, texto, sentimento, interação e respostas. As respostas seguem a mesma estrutura dos comentários.

Essa abordagem utilizando o *PRAW* facilita a coleta eficiente e sistemática de dados, permitindo-nos construir uma base sólida para análise e pesquisa no âmbito do *subreddit Political Discussion*.

4.2 Implementação do Modelo

Nesta etapa do projeto, avançamos com a implementação do modelo delineado na Seção 2.2. O código correspondente está disponível no repositório do projeto no *GitHub*⁷. Durante o desenvolvimento, contamos com o suporte das principais bibliotecas, tais como *pandas*, *numpy*, *nlk*, *praw* e *math*. Essas bibliotecas desempenham um papel fundamental na análise de sentimento, processamento e manipulação de dados, e facilitam o acesso à API do *Reddit*.

4.2.1 Árvore de Debate

Para representar uma Árvore de Debate (DebT), realizamos a conversão dos dados obtidos pelo *PRAW* em uma estrutura tabular. Nessa abordagem, cada comentário é representado como uma linha, contendo informações como o nome do **Usuário** (Autor do Comentário), **C1** (ID do Comentário), **C2** (ID do Comentário de Resposta), **W** (sentimento) e **Interação** (nome do usuário que respondeu o comentário). Essa representação tabular proporciona uma organização eficaz dos dados, essencial para a análise subsequente (ver Figura 8).

Figura 8 – Representação da Árvore de Debate.

Usuário	C1	C2	W	Interação
AutoModerator	kaxlu33	184u316	0.5547	184u316
Yvaelle	kb09pjw	kazu4ku	0.5637	ddoyen
A-Wise-Cobbler	kb288nt	kb1zs6q	-0.2644	III-Description3096
teh_maxh	kb7nrq6	kb4pt9z	-0.7506	two-wheeled-dynamo
two-wheeled-dynamo	kb4pt9z	kb4dm7m	0.0000	III-Description3096

Fonte: Elaborado pelos autores.

⁶ <https://praw.readthedocs.io/en/stable/>

⁷ <https://github.com/bryanlimac/TCC-algoritmo>

4.2.2 *Árvore de Debate Bilateral*

Com a tabela DebT, torna-se possível identificar concordâncias ou discordâncias em relação ao comentário raiz. Para viabilizar essa análise, introduzimos duas novas colunas: **Sc1** (indicando concordância e discordância em relação ao comentário raiz) e **Sc2** (indicando concordância e discordância em relação ao comentário pai do comentário atual).

Por meio dessas colunas adicionais, atribuímos o valor 1 à coluna **Sc2** para todos os comentários cujo **C2** seja igual ao do comentário raiz. Com base nos valores em **Sc2**, estabelecemos uma condição para verificar se os sentimentos são correspondentes. Caso afirmativo, atribuímos o valor 1 à coluna **Sc1** desse comentário e também atribuímos o valor 1 a todas as linhas da coluna **Sc2** em que o **C1** do comentário atual é igual ao **C2** de todos os comentários da tabela. Se a condição não for atendida, **Sc1** recebe -1. Além disso, para todas as linhas da coluna **Sc2** em que o **C1** do comentário atual é igual ao **C2** de todos os comentários da tabela, também atribuímos o valor resultante da condição. Essa abordagem proporciona uma estrutura lógica para avaliar as relações entre os comentários na tabela SDebT, destacando concordâncias e discordâncias em diferentes níveis da discussão (ver Figura 9).

Figura 9 – Representação da Árvore de Debate Bilateral.

Usuário	C1	C2	W	Interação	Sc1	Sc2
AutoModerator	k3oadec	1712iuh	0.5547	1712iuh	1	1
GiantPineapple	k8m0qrg	k8lpurm	0.9185	SmoothCriminal2018	-1	-1
SmoothCriminal2018	k8lpurm	k8li6bu	0.9246	GiantPineapple	-1	-1
fishman1776	k8qq86g	k8li6bu	0.0000	GiantPineapple	1	-1
GiantPineapple	k8li6bu	1712iuh	-0.8193	1712iuh	-1	1

Fonte: Elaborado pelos autores.

4.2.3 *Grafo de Debate de Usuário*

Para construir a tabela referente ao Grafo de Debate do Usuário, empregamos as tabelas previamente desenvolvidas para calcular os valores de **Sc(i)**, **p**, e **w**, conforme mencionado na Definição 4. Para calcular **Sc(i)**, iniciamos agrupando a tabela SDebT pela coluna **Usuário** e determinamos a média da coluna **Sc1**. No que diz respeito ao cálculo de **p**, foi realizada a divisão da quantidade de interação positiva pela quantidade total de interações. Para calcular **w**, procedemos agrupando os usuários e suas respectivas interações da tabela SDebT e calculamos a média dos sentimentos (ver Figura 10).

Figura 10 – Representação do Grafo de Debate de Usuário.

Usuário	Interação	Sc(i)	p	w
2000thtimeacharm	DoctorChampTH	1.000000	1.0	0.5719
Ambitious_Clue_3201	1712iuh	1.000000	1.0	0.7609
AutoModerator	1712iuh	1.000000	1.0	0.5547
Blackmass91	No-Touch-2570	0.333333	1.0	0.6093
Brit-Crit	1712iuh	1.000000	1.0	0.8355

Fonte: Elaborado pelos autores.

4.2.4 Encontrando a Partição mais Polarizada

A partir da tabela de Debate dos Usuários, aplicamos o algoritmo de polarização descrito no Algoritmo 1. O resultado desse processo é a geração de duas tabelas, L e R, que compõem o grafo bipartido. Essas partições são otimizadas para maximizar a polarização, proporcionando uma representação eficaz das dinâmicas presentes nos debates entre os usuários (ver Figura 11).

Figura 11 – Representação das Partições L e R.

Partição L

Usuário	Interação	Sc(i)	p	w
ExtruDR	184u316	-1.000000	0.0	-0.812600
TheLastCoagulant	boogi3woogie	0.500000	1.0	0.750600
TheLastCoagulant	Avatar_exADV	0.500000	0.0	-0.493900
TheLastCoagulant	RingAny1978	0.500000	0.0	-0.611500
djm19	Obvious_Chapter2082	1.000000	0.0	-0.319700

Partição R

Usuário	Interação	Sc(i)	p	w
Njorls_Saga	newsreadhjw	0.333333	1.0	0.67790
Njorls_Saga	Tremor_Sense	0.333333	1.0	0.07720
Njorls_Saga	LookAnOwl	0.333333	0.0	0.00000
Saanvik	OprahtheHutt	-1.000000	1.0	0.93580
Tremor_Sense	Aacron	0.200000	0.5	0.08785

Fonte: Elaborado pelos autores.

4.3 Medidas de Influência Social

Para desenvolver essas métricas, utilizando como base o trabalho de (HAMZEHEI et al., 2017), o trabalho propõe uma abordagem de medição de influência baseada em tópicos. Essa

abordagem visa integrar os relacionamentos entre usuário-tópico, informações de conteúdo de tópicos e conexões sociais entre usuários em um único modelo fundamentado. Ao considerar a influência e os interesses individuais em um tópico, essa abordagem confere a capacidade de prever a influência de alguém em um novo tópico. Para isso, o autor desenvolve um algoritmo utilizando métricas de influência. As métricas utilizadas incluem a escala de seguidores, atividade do tópico, atratividade baseada em tópicos e centralidade na rede. Essas medidas são posteriormente combinadas para formar uma única métrica de influência. No âmbito do nosso trabalho, adaptamos essas métricas para serem aplicadas ao contexto específico do *Reddit*, resultando em *karma*, atividade, atratividade e centralidade como indicadores-chave.

4.3.1 *Karma*

Essa medida representa o número de *upvotes* e *downvotes* nas postagens e comentários de um usuário na rede. Se um usuário tiver mais *downvotes* do que *upvotes*, seu *karma* será negativo; caso contrário, será positivo. A forma como o *karma* é calculado não é divulgada pelo *Reddit* para evitar que os usuários encontrem maneiras de burlar essa métrica. Esse valor é constante em todos os tópicos para o usuário e é independente dos tópicos, demonstrando a força do engajamento do usuário. Um exemplo da medida de *karma* que foi extraído pelo *PRAW* é mostrado na Figura 12.

Figura 12 – Representação do *Karma* dos Usuários.

Usuário	Karma
3rdtimeischarm	53137.0
Aacron	33571.0
AdhesivenessCivil581	44103.0
AgentMonkey	35310.0
Altruistic-Text3481	289773.0

Fonte: Elaborado pelos autores.

Observe que o *karma* extraído possui um valor consideravelmente elevado. Para integrá-lo em nosso trabalho, optamos por realizar uma normalização, conforme ilustrado na Figura 13. Para realizar a normalização, procedemos calculando a razão entre o *karma* de cada usuário e o somatório do valor dos *karmas* de todos os usuários, resultando um valor normalizado no intervalo de 0 a 1.

Figura 13 – Representação de *Karma* Normalizado.

Usuário	Karma
3rdtimeischarmmy	0.002384
Aacron	0.001506
AdhesivenessCivil581	0.001979
AgentMonkey	0.001584
Altruistic-Text3481	0.013002

Fonte: Elaborado pelos autores.

4.3.2 *Atividade*

Essa medida captura a quantidade de comentários do usuário no debate. Para calculá-la, realizamos a somatório dos comentários do usuário e dividimos pelo total de comentários do debate. Dessa forma, obtemos um valor normalizado no intervalo de 0 a 1 (ver Figura 14).

Figura 14 – Representação da Atividade Normalizada.

Usuário	Atividade
SteelmanINC	0.040179
I405CA	0.004464
ScaryBuilder9886	0.022321
Morat20	0.008929
SeekingAugustine	0.031250

Fonte: Elaborado pelos autores.

4.3.3 *Atratividade*

Para calcular a atratividade, analisamos os comentários e avaliamos a quantidade de respostas associadas a ele. Uma vez que um usuário pode contribuir com vários comentários, calculamos a média do engajamento gerado por cada usuário. Posteriormente, aplicamos uma normalização, obtida pela divisão da interação atual pelo total das interações gerais. O resultado final é um valor situado no intervalo de 0 a 1, proporcionando uma representação normalizada do engajamento considerando a participação de cada usuário (ver Figura 15).

Figura 15 – Representação da Atratividade Normalizada.

Usuário	Atratividade
SteelmanINC	0.041185
I405CA	0.037067
ScaryBuilder9886	0.018276
Morat20	0.028143
SeekingAugustine	0.007301

Fonte: Elaborado pelos autores.

4.3.4 Centralidade

A centralidade de um usuário é um indicador de influência em uma rede social. Para avaliar essa centralidade, aplicamos a Centralidade de Intermediação⁸, que é calculada considerando o número de menores caminhos de todos os vértices para quaisquer outros vértices que passam por aquele nó. Em termos mais simples, um nó com alta Centralidade de Intermediação está geometricamente próximo a muitos outros nós na rede, indicando sua eficiência em termos de interação. A fórmula para calcular a Centralidade de Intermediação $g(v)$ para um nó v é expressa como segue:

$$g(v) = \sum_{s \neq v, t \neq v} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

onde σ_{st} é o número total de menores caminhos do nó s para o nó t , e $\sigma_{st}(v)$ é o número desses caminhos que passam pelo nó v .

No processo de calcular a centralidade de intermediação para cada usuário no grafo, empregamos o algoritmo de Floyd-Warshall (CORMEN et al., 2012). Esse algoritmo desempenha um papel crucial ao determinar as distâncias mais curtas entre todos os pares de nós em um grafo. Sua aplicação permite mapear eficientemente a proximidade relativa entre todos os vértices, estabelecendo a base para a avaliação precisa da centralidade de intermediação. Assim como as outras medidas precisamos normalizar o valor no intervalo de 0 a 1. A normalização é aplicada da seguinte forma:

$$normal(g(v)) = \frac{g(v) - \min(g)}{\max(g) - \min(g)}$$

A Figura 16 representa o valor normalizado da centralidade de intermediação.

⁸ <https://pt.wikipedia.org/wiki/Intermediação>

Figura 16 – Representação da Centralidade de Intermediação Normalizada.

Usuário	Centralidade de Intermediação
SteelmanINC	0.380802
I405CA	0.212922
ScaryBuilder9886	0.198298
Morat20	0.006581
SeekingAugustine	0.165394

Fonte: Elaborado pelos autores.

4.3.5 Agregação das Métricas

As métricas mencionadas anteriormente, como atividade, atratividade e centralidade, serão agregadas para criar uma única métrica de influência para cada usuário no debate. Adicionalmente, o *karma* será utilizado como critério de desempate caso a medida de influência de um usuário seja igual à de outro. A consolidação das métricas é realizada por meio da média entre as três métricas mencionadas. Os resultados desse processo de agregação estão visualmente apresentados na Figura 17.

Figura 17 – Representação da Agregação das Métricas.

Usuário	Karma	Influência
2manyfelines	0.021208	0.000782
AutoModerator	0.000235	0.000782
AwesomeScreenName	0.034871	0.000782
Awesomeuser90	0.015774	0.095557
Awesomeuser90	0.015774	0.095557

Fonte: Elaborado pelos autores.

4.4 Aplicação da Métrica de Influência Social

Para a aplicação dessa métrica, adotamos um algoritmo de busca linear, operando em um grafo bipartido derivado do Algoritmo 1. O Algoritmo 2 atua na busca da sub-bipartição que resulta na máxima polarização bipartida. Este algoritmo é designado como *MaxSubPartition* e é apresentado a seguir.

Algoritmo 2 *MaxSubPartition*(\mathcal{G})

```

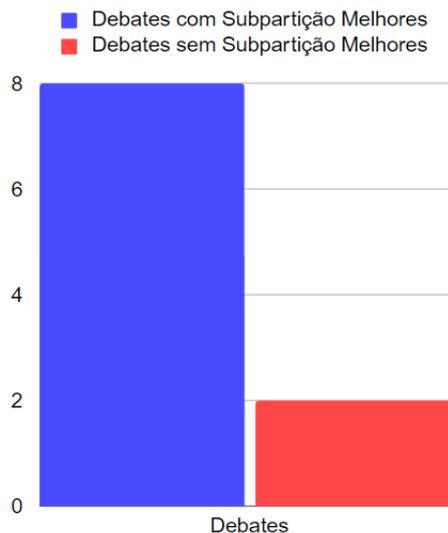
1: Entrada:  $\mathcal{G} = (\mathcal{C}, \mathcal{E}, \mathcal{S}, \mathcal{W})$ 
2: Saída: SubBipartição  $\mathcal{G}'$  de  $\mathcal{G}$  tal que  $(L, R)$  de  $\mathcal{G}'$  tem a polarização bipartida máxima
3:  $steps \leftarrow 1$ 
4:  $maxsteps \leftarrow |\mathcal{C}|$ 
5:  $best \leftarrow \mathcal{G}$ 
6: while  $steps \leq maxsteps$  do
7:    $\mathcal{G}' \leftarrow getMostInfluent(\mathcal{G}, (maxsteps - steps))$ 
8:   if  $BipPol(\mathcal{G}') > BipPol(best)$  then
9:      $best \leftarrow \mathcal{G}'$ 
10:  end if
11:   $steps \leftarrow steps + 1$ 
12: end while
13: return  $getInitBPart(best)$ 

```

5 RESULTADOS

Nesta seção, destacamos os resultados obtidos por meio da implementação da metodologia proposta para a identificação de subpartições com polarização substancialmente maior em comparação com as duas partições originais nos debates políticos do *Reddit*. Com o objetivo de otimizar a estrutura de debates, foram realizadas análises de influência entre os participantes em um total de 10 debates. Na Figura 18, é possível notar que ocorreram 8 debates com uma sub-bipartição aprimorada, enquanto 2 debates mantiveram as partições inalteradas.

Figura 18 – Representação do Total de Debates Analisados.



Fonte: Elaborado pelos autores.

A Tabela 1 apresenta os resultados com mais detalhes, oferecendo uma visão aprofundada do desempenho das métricas aplicadas. Observa-se uma diversidade significativa nos resultados,

refletindo a heterogeneidade dos debates considerados. Análisisando os debates, observamos que existe três tipos de resultados.

O primeiro tipo de resultado compreende os casos em que há uma pequena redução percentual, porém não ocorreu melhora na polarização, ou houve alguma melhora, mas esta se manteve próxima ao valor inicial. Isso ocorre porque o debate já está com polarização máxima; mesmo ao aplicarmos o algoritmo de redução, o valor da polarização permanecerá próximo.

O segundo caso envolve debates nos quais houve uma considerável redução percentual, mas a polarização se manteve próxima. Isso indica que o debate já estava polarizado, ou seja, os usuários que entraram no debate após aquele ponto não exerceram uma forte influência sobre ele.

O terceiro tipo de resultado abarca os casos nos quais houve uma grande redução percentual e a polarização também teve uma melhora significativa. Isso indica que o debate estava polarizado até certo ponto e um usuário ou um grupo de usuários que chegou depois contribuiu de forma negativa para a polarização.

ID do Debate	BipPol	BipPol Após a Redução	Porcentagem de Redução
17x1my6	0.2738	0.2899	02.04%
189b64a	0.2900	0.3592	06.99%
189ba2a	0.3707	0.5085	77.06%
189bymj	0.1722	0.2078	02.11%
1893sww	0.3888	0.4110	33.34%
18650ll	0.1936	NaN	0%
18653z1	0.2622	0.4840	47.15%
185zikk	0.4047	0.4443	85.48%
184ugfx	0.4124	0.4413	46.93%
183m8z3	0.2574	NaN	0%

Tabela 1 – Resultado do grau de polarização máxima de debates *online* após a aplicação do algoritmo de *MaxSubPartition*.

6 CONSIDERAÇÕES FINAIS

Este estudo introduziu métricas de influência baseadas no *Reddit* para a detecção de subpartições com polarização maior em comparação com as partições originais em debates *online*. O modelo propõe a remoção de usuários menos influentes como uma estratégia para encontrar subpartições com polarização mais aprimorada em comparação às iniciais.

Ao longo da pesquisa, foram coletados dados de debates *online* na plataforma *Reddit* usando um script em *Python* com o *PRAW*. Em seguida, foi desenvolvido um algoritmo de busca linear em *Python* no *Google Colab*, capaz de identificar as partições que maximizam a polarização em um grafo bipartido e calcular a influência dos usuários no debate. Os resultados deste trabalho apontam para um desempenho positivo, destacando uma melhoria na polarização ao retirar os usuários menos influentes.

Durante a condução desta pesquisa, foram identificadas algumas dificuldades, sendo a complexidade computacional uma delas, o que dificulta uma análise mais detalhada e pode

limitar a compreensão desses debates.

Como perspectiva para trabalhos futuros, sugere-se a realização de uma análise individual dos usuários e sua influência na polarização, visando uma compreensão mais aprofundada e refinada dos mecanismos que regem os debates *online*. Além disso, são visadas estudar estratégias de diminuição da complexidade dos algoritmos utilizados. Um outro tópico importante de estudo consiste em analisar a estrutura dos grafos nos grafos de debates e como eles podem influenciar no resultado da polarização e na complexidade dos algoritmos.

REFERÊNCIAS

AL-AYYOUB, M. et al. Studying the controversy in online crowds' interactions. **Applied Soft Computing**, Elsevier, v. 66, p. 557–563, 2018.

ALSINET, T. et al. Measuring polarization in online debates. **Applied Sciences**, MDPI, v. 11, n. 24, p. 11879, 2021.

AMIN, M. T. A. et al. Unveiling polarization in social networks: A matrix factorization approach. In: IEEE. **IEEE INFOCOM 2017-IEEE Conference on Computer Communications**. [S.l.], 2017. p. 1–9.

CHA, M. et al. Measuring user influence in twitter: The million follower fallacy. In: **Proceedings of the international AAAI conference on web and social media**. [S.l.: s.n.], 2010. v. 4, n. 1, p. 10–17.

CORMEN, T. H. et al. **Introduction to Algorithms**. 3rd. ed. Cambridge, Massachusetts: The MIT Press, 2012.

DETYNIECKI, M. Fundamentals on aggregation operators. **This manuscript is based on Detyniecki's doctoral thesis and can be downloaded from**, 2001.

DIXIT, A. K.; WEIBULL, J. W. Political polarization. **Proceedings of the National Academy of sciences**, National Acad Sciences, v. 104, n. 18, p. 7351–7356, 2007.

EVANS, J. H. Polarization in abortion attitudes in us religious traditions, 1972–1998. In: SPRINGER. **Sociological Forum**. [S.l.], 2002. v. 17, p. 397–422.

FLORES, A. R.; BARCLAY, S. Backlash, consensus, legitimacy, or polarization: The effect of same-sex marriage policy on mass attitudes. **Political Research Quarterly**, SAGE Publications Sage CA: Los Angeles, CA, v. 69, n. 1, p. 43–56, 2016.

HAMZEHEI, A. et al. Topic-based social influence measurement for social networks. **Australasian Journal of Information Systems**, Australian Computer Society, v. 21, 2017.

ISENBERG, D. J. Group polarization: A critical review and meta-analysis. **Journal of personality and social psychology**, American Psychological Association, v. 50, n. 6, p. 1141, 1986.

MCCRIGHT, A. M.; DUNLAP, R. E. The politicization of climate change and polarization in the american public's views of global warming, 2001–2010. **The Sociological Quarterly**, Taylor & Francis, v. 52, n. 2, p. 155–194, 2011.

MILLER, S. V. What americans think about gun control: evidence from the general social survey, 1972–2016. **Social Science Quarterly**, Wiley Online Library, v. 100, n. 1, p. 272–288, 2019.

MORALES, A. J. et al. Measuring political polarization: Twitter shows the two sides of venezuela. **Chaos: An Interdisciplinary Journal of Nonlinear Science**, AIP Publishing, v. 25, n. 3, 2015.

MOUW, T.; SOBEL, M. E. Culture wars and opinion polarization: the case of abortion. **American Journal of Sociology**, The University of Chicago Press, v. 106, n. 4, p. 913–943, 2001.

SHAFIQ, M. Z. et al. Identifying leaders and followers in online social networks. **IEEE Journal on Selected Areas in Communications**, IEEE, v. 31, n. 9, p. 618–628, 2013.

SUNSTEIN, C. R. The law of group polarization. **University of Chicago Law School, John M. Olin Law & Economics Working Paper**, n. 91, 1999.

TREVISAN, L. Max cut and the smallest eigenvalue. In: **Proceedings of the forty-first annual ACM symposium on Theory of computing**. [S.l.: s.n.], 2009. p. 263–272.

WAUGH, A. S. et al. Party polarization in congress: A social networks approach. 2009.

WEINBERGER, D. **Too Big to Know: Rethinking Knowledge Now That the Facts Aren T the Facts, Experts Are Everywhere, and the Smartest Person in the Room Is the Room**. [S.l.]: Basic Books, 2011.