O Uso das Redes Sociais para Interferir nas Democracias: Um Mapeamento Sistemático da Literatura

Yuri Luz de Almeida, Francis Spiegel Rubin, Adriana Cesário de Faria Alvim, Vânia Maria Félix Dias, Rodrigo Pereira dos Santos

Programa de Pós-Graduação em Informática – UNIRIO
Avenida Pasteur, 458 – Urca – CEP: 22290-255 – Rio de Janeiro, RJ, Brasil

{yuri.almeida, fran.spiegel}@edu.unirio.br,
{adriana, vania, rps}@uniriotec.br

Abstract. *Social networks have become the main source of information in the digital age. As a consequence, an increasing volume of information is being shared in virtual environments. In this scenario, there is evidence that social networks are being used by malicious players in a coordinated way to manipulate other users, and it is a real threat to democracies. Therefore, this paper presents a systematic mapping study on the use of social networks to interfere in democracies. The goal is to initially discover what has already been raised in the literature over this topic.*

Resumo. *As redes sociais se tornaram a principal fonte de informação na era digital. Como consequência, um crescente volume de informações vem sendo compartilhado nos ambientes virtuais. Diante desse cenário, há evidências de que as redes sociais estão sendo utilizadas por atores mal-intencionados de forma coordenada para manipular outros usuários, o que se configura como uma ameaça real às democracias. O objetivo deste artigo é apresentar os resultados de um mapeamento sistemático da literatura sobre o uso de redes sociais para interferir nas democracias a fim de descobrir o que já foi levantado sobre o tópico.*

1. Introdução

Até recentemente, as redes sociais eram vistas como possíveis responsáveis pelo fortalecimento da democracia e participação popular (Badawy et al. 2019). As pesquisas realizadas acerca dos efeitos positivos das redes sociais, como o aumento da participação dos eleitores no debate político, contribuíram para o consenso sobre o poder que essas plataformas exercem como ferramentas para promoção da democracia e do engajamento cívico. No entanto, estudiosos das redes sociais têm levantado preocupações sobre a possibilidade de manipulação da opinião pública através das redes sociais. Neste contexto, mecanismos maliciosos – e.g. contas mal-intencionadas (*trolls*) e robôs (*bots*) – vêm sendo usados para promover a desinformação nas redes sociais com o objetivo de influenciar as decisões populares (Linvill et al. 2019).

Nesse sentido, o objetivo deste artigo é apresentar os resultados preliminares de um trabalho que visou explorar na literatura estudos que analisam interferências do uso das redes sociais nas democracias. A fim de alcançar este objetivo, um mapeamento sistemático da literatura (MSL) foi conduzido. Além desta seção de introdução, este

artigo está organizado da seguinte forma: a Seção 2 descreve o método de pesquisa; a Seção 3 analisa os resultados preliminares obtidos; a Seção 4 apresenta uma discussão sobre os principais pontos; por fim, a Seção 5 conclui o artigo com algumas considerações finais.

2. Método

Para a realização deste estudo, foi executado um protocolo com base nas diretrizes para estudos do tipo revisões e mapeamentos sistemáticos de literatura de Kitchenham e Charters (2007). As questões de pesquisa foram formuladas seguindo os critérios especificados a partir da estrutura PIO (*Population*, *Intervention*, *Outcomes*), conforme a Tabela 1.

Tabela 1. Critérios do PIO

Population	Eventos democráticos
Intervention	Redes sociais
Outcomes	Manipulação / influência / interferência / impacto pelas redes sociais

Dessa forma, chegou-se a seguinte questão de pesquisa (QP): "*Como as redes sociais podem influenciar as democracias?*". Para ajudar a responder a esta QP, foram elaboradas quatro subquestões: 1) "*Em que momentos há evidências de interferência nos eventos democráticos por meio das redes sociais?"*; 2) "*Quais foram os atores identificados por atuar diretamente nesses episódios?"*; 3) "*Que plataformas de redes sociais foram mais exploradas nesses casos?"*; e 4) "*Que medidas foram tomadas por algum país, ou sugeridas na literatura, para acabar ou mitigar esse problema?"*.

A pesquisa foi realizada com buscas em bibliotecas digitais (BD) indicadas por Souza e Conte (2017), a saber: Scopus[1], ACM[2], Engineering Village[3], ScienceDirect[4] e IEEE[5]. A *string* de busca utilizada foi: ("democra*" OR "election*" OR "plebiscite*" OR "referendum*" OR "votation*") AND ("social media*" OR "social network*") AND ("manipulat*" OR "influenc*" OR "interferenc*" OR "impact*"). As Tabelas 2 e 3 apresentam os critérios de inclusão (CI) e de exclusão (CE) aplicados neste trabalho.

Tabela 2. Critérios de Inclusão

Código	**Descrição**
CI1	O artigo avalia ou apresenta as influências das redes sociais nas democracias.
CI2	O artigo aborda casos em que as redes sociais foram utilizadas para interferência em eventos democráticos ou manipulação da sociedade.

O processo de seleção de estudos foi sistematizado seguindo seis etapas. Na Etapa 1, a *string* de busca foi utilizada nas BD selecionadas e foram retornados 351 estudos. Na Etapa 2, os estudos duplicados foram removidos, restando 226 resultados.

[1] https://www.scopus.com/home.uri/
[2] http://dl.acm.org/
[3] https://www-engineeringvillage-com/
[4] http://www.sciencedirect.com/
[5] http://ieeexplore.ieee.org/Xplore/home.jsp

Entre as Etapas 3 e 5, os CI e os CE foram aplicados. Na Etapa 3, a seleção se deu por meio da leitura dos títulos, resumos e palavras-chave, resultando em 86 estudos. Na Etapa 4, os estudos foram selecionados a partir da leitura da introdução e conclusão. Assim, chegou-se a um total de 25 estudos para a leitura completa na Etapa 5. Após a Etapa 5, 14 estudos foram aprovados para a Etapa 6, referente à extração de dados.

Tabela 3. Critérios de Exclusão

Código	Descrição
CE1	O artigo não avalia ou apresenta os impactos nas democracias pelo uso de redes sociais.
CE2	As redes sociais são somente citadas e não são o foco do estudo do artigo.
CE3	É um prefácio, livro, editorial, resumo, pôster, painel, palestra, mesa redonda, oficina, demonstração, tese, dissertação ou monografia de conclusão de curso.

O processo de extração dos dados se deu por meio de um formulário para o registro dos artigos. A Tabela 4 exibe, de forma sumarizada, os dados dos estudos selecionados (referidos como E*XX*, onde *XX* é o identificador do estudo).

Tabela 4. Estudos Selecionados

ID	Título	Ano	Fonte	Autores	Redes Sociais	Métodos	Eventos
E1	Filter Bubbles and Fake News	2017	XRDS23	D. DiFranzo e K. Gloria-Garcia	Facebook e Twitter	Bolhas virtuais e *Fake News*	Referendos e Eleições
E2	Impact of Internet and Social Networks on the Final Result of 2016 Presidential Election of the United States	2017	2017 NITC	L. Jayawardena	Facebook, Instagram, Facebook e Twitter	-	Eleições
E3	The Fake News Spreading Plague: Was It Preventable?	2017	WebSci 17	E. Mustafaraj e P. T. Metaxas	Facebook e Twitter	*Fake News*	Eleições
E4	Robôs, Redes Sociais e Política no Brasil: Estudo sobre Interferências Ilegítimas no Debate Público na Web, Riscos à Democracia e Processo Eleitoral de 2018	2017	FGV Dapp	M. A. Ruediger	Twitter	*Bots*	Eleições, Votações Legislativas e Manifestações
E5	Is It Time for an Offence of 'Dishonest Algorithmic Manipulation for Electoral Gain'?	2017	Alternative Law Journal 42	D. J. B. Svantesson e W. van Caenegem	Facebook, Google e Twitter	Manipulação de Algoritmo, *Fake News* e *Bots*	Referendos e Eleições
E6	From Brexit to Trump: Social Media's Role in Democracy	2018	Computer 51	W. Hall, R. Tinati e W. Jennings	Facebook e Twitter	-	Referendos e Eleições
E7	The Impact of Malicious Accounts on Political Tweet Sentiment	2018	2018 CIC	B. Heredia, J. D. Prusa e T. M. Khoshgoftaar	Facebook e Twitter	*Bots*	Eleições
E8	From Alt-Right to Alt-Rechts: Twitter Analysis of the 2017 German Federal Election	2018	2018 WWW	F. Morstatter, Y. Shao, A. Galstyan e S. Karunasekera	Twitter	*Bots*	Eleições
E9	Characterizing The 2016 Russian IRA Influence Campaign. Social Network Analysis and Mining	2019	Social Network Analysis and Mining 9	A. Badawy et al.	Facebook e Twitter	*Bots*, *Fake News* e *Trolls*	Eleições
E10	Social Media as an Opinion Formulator: A Study on Implications and Recent Developments	2019	2019 iCoMET	T. Jameel, R. Ali e K. A. Malik	Facebook, Twitter e Youtube	*Bots* e *Fake News*	Referendos e Eleições
E11	Social Distraction? Social Media Use and Political Knowledge in Two U.S. Presidential Elections	2019	Computers in Human Behavior 90	S. Leea e M. Xenosb	Facebook	*Fake News*	Eleições
E12	"THE RUSSIANS ARE HACKING MY BRAIN!" Investigating Russia's Internet Research Agency Twitter Tactics During the 2016 United States Presidential Campaign	2019	Computers in Human Behavior 99	D. L. Linvill et al.	Twitter, Facebook e Reddit	*Bots*, *Fake News* e *Trolls*	Eleições
E13	Presidential Elections in Ecuador: Bot Presence in Twitter	2019	2019 ICEDEG	D. Rofrío et al.	Twitter	*Bots*	Eleições
E14	Screening out Social Bots Interference: Are There Any Silver Bullets?	2019	IEEE Communicatio ns Magazine 57	M. Zago et al.	Facebook e Twitter	*Bots*	-

3. Resultados

3.1. Análise dos Resultados

3.1.1. Como as redes sociais podem influenciar as democracias?

Diversos mecanismos de manipulação de opinião e seus impactos nas redes sociais são estudados na literatura analisada. A técnica mais citada é a atuação de robôs (*bots*), que está presente nos estudos E4, E5, E7-E10 e E12-E14. A segunda técnica mais observada é a divulgação de informação manipulada (*fake news*), que pode ser vista nos estudos E1, E3, E5 e E9-E12. Apesar de menos recorrentes, foram identificados estudos sobre *trolls* no mapeamento realizado.

Faz-se necessário apresentar as definições dos termos mais relevantes identificados na literatura explorada neste MSL. *Bots* são contas automáticas que atuam com diversas táticas como inflação de *hashtags*, criação de tendências artificiais, campanhas de difamação e propaganda política - que podem ser usadas para violar os direitos de livre expressão de outras pessoas de uma maneira diferente (Svantesson e Caenegem 2017).

Por sua vez, *fake news* são notícias falsas publicadas e compartilhadas como se fossem verdades. Apesar de ser um termo simples de entender, é mais complicado definir objetivamente o que são *fake news*, uma vez que existem desde mentiras diretas até distorções - que são mais difíceis de classificar como falsas. Por fim, *trolls* são contas virtuais operadas manualmente com o objetivo principal de manipular a opinião pública. Para rotular algumas contas ou fontes de informação como *trolls*, é necessário haver uma clara intenção de enganar ou criar conflito (Badawy et al. 2019).

3.1.2. Em que momentos há evidências de interferência nos eventos democráticos por meio das redes sociais?

Os fatos recentes mais citados nos estudos foram as eleições presidenciais dos EUA e o Brexit, ambos ocorridos no ano de 2016. Nos estudos E1-E3, E5-E7 e E9-E12, pelo menos uma dessas votações foi analisada. Outros dois artigos, E8 e E13, analisaram, respectivamente, o uso de *bots* nas eleições federais da Alemanha e do Equador do ano de 2017. Por fim, o estudo E4 faz uma análise da presença de *bots* em eleições brasileiras, além também de analisar o uso de *bots* para impulsionar manifestações políticas ou demonstrar apoio contra ou a favor de uma votação no legislativo.

3.1.3. Quais foram os atores identificados por atuar diretamente nesses episódios?

Em todos os eventos democráticos analisados pelos estudos mapeados, nenhum responsável direto pela manipulação é determinado. As exceções são os estudos E9 e E12, que analisaram a eleição presidencial dos EUA de 2016 e apontam a participação da IRA – uma companhia de pesquisa russa – na interferência da eleição no país, tendo como base a investigação aberta pelo Congresso Estadunidense. Nos estudos E5 e E14, que discutem o Brexit, a empresa de análise de dados Cambridge Analytica é citada como personagem envolvida em uma possível manipulação política.

3.1.4. Que plataformas de redes sociais foram mais exploradas nesses casos?

As duas plataformas de redes sociais mais citadas nos estudos foram o Facebook e o Twitter, mas elas são exploradas por motivos diferentes. É possível observar na literatura que o Facebook é explorado por sua popularidade. Conforme citado no estudo E2, até novembro de 2016, o Facebook detinha 42,1% do total do mercado de redes sociais. O estudo E14 relata que o Twitter é a plataforma de rede social com a maior proporção de *bots* por usuário. Nesse mesmo estudo, os autores apontam que a razão para o Twitter ser a plataforma de rede social mais explorada por esse método malicioso é o fato de que, geralmente, os seus usuários costumam agir de maneira recíproca quando recebem um novo seguidor. Conforme discutido no estudo E3, um *bot* ainda pode atingir algum público no Twitter fazendo réplicas em publicações de perfis com muitos seguidores.

3.1.5. Que medidas foram tomadas por algum país, ou sugeridas na literatura, para acabar ou mitigar esse problema?

Os autores do estudo E5 sugeriram duas opções de medidas que poderiam ser tomadas. A primeira opção é focada no conteúdo das publicações para proteger as atividades políticas *online*. A segunda opção é classificar como crime se o indivíduo fizer algo com a intenção de obter ganhos eleitorais por manipulação algorítmica desonesta. Entretanto, para ambas as opções, a lei e todos os seus termos devem ser bem definidos e claros para não haver margens para más interpretações. Por fim, os autores alertam que as soluções para esse tema podem soar como um mecanismo de amarra para o debate político, atingindo o "coração" de uma sociedade democrática: a liberdade de expressão.

4. Discussão

Os resultados deste mapeamento mostraram que há muitos estudos sobre como as redes sociais podem ser exploradas a fim de exercer interferência em democracias e manipulação social, tanto na área da Computação como na área das Ciências Sociais, por ser uma problemática sociotécnica. Os estudos levantados neste trabalho destacam como as tecnologias empregadas nas redes sociais contribuíram na disseminação de *fake news*. O estudo E3 explora como essas técnicas não só servem para conseguir um voto a favor de um determinado candidato em período eleitoral, como também impulsionam correntes de opinião. O estudo E5 expõe que *big data* foi explorada no Brexit para identificar os perfis de pessoas e esse rastreamento foi usado para interferir no referendo com a produção de *fake news* voltadas para esse público. Apesar dos métodos maliciosos terem sido abordados isoladamente nos estudos, isso não significa que eles são empregados de forma exclusiva. Essa análise da associação dos métodos maliciosos é especialmente relevante quando falamos em *fake news*, pois elas são disseminadas por *bots* e *trolls* a fim de manipular com mais sucesso a população.

O presente estudo levanta a questão sobre que ações devem ser tomadas pelas instituições públicas e privadas para frear a propagação de conteúdo malicioso. Uma discussão importante (e ainda pouco explorada) se refere a como as leis governamentais devem cuidar da proteção da população *online*, principalmente em países de livre expressão democrática, e em quais circunstâncias se deve tratar o uso de *bots* e o funcionamento dos algoritmos das redes sociais como ameaça à democracia.

Durante o mapeamento, foi possível notar que a percepção sobre as redes sociais para uso político mudou desde a eleição presidencial dos EUA de 2016, como apontado por Badawy et al. (2019). Todos as vulnerabilidades expostas pelo processo americano e referendo britânico trouxeram uma grande preocupação e busca por respostas. Há ainda muito o que se explorar sobre esse tema como, por exemplo, identificar as organizações que buscam causar turbulência nos ritos democráticos e observar que ações estão sendo tomadas nos países para proteger suas democracias do ambiente virtual. Esses são alguns pontos que ainda precisam ser aprofundados e estudados mais detalhadamente a partir de pesquisas sobre análise e mineração de redes sociais com dados reais.

5. Considerações Finais

Este artigo explorou na literatura estudos que indicam ou analisam interferências do uso das redes sociais nas democracias por meio do planejamento e execução de um MSL, cujos resultados preliminares foram discutidos à luz de algumas questões. Como resultado, foram encontrados estudos sobre práticas maliciosas e uso de técnicas que exploram divisões da sociedade para manipular opiniões e dificultar o debate de temas importantes nas redes sociais. Este mapeamento expõe ainda uma mudança negativa da expectativa acerca da relação entre as redes sociais e seu impacto nas democracias, o que pretende ser objeto de investigação em trabalhos futuros desta pesquisa, além da análise e discussão completa dos resultados deste mapeamento.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Referências Bibliográficas

Badawy, A. et al. (2019). "Characterizing the 2016 Russian IRA influence campaign", Soc. Netw. Anal. Min. 9. DOI: https://doi.org/10.1007/s13278-019-0578-6.

Kitchenham, B. e Charters, S. (2007). “Guidelines for Performing Systematic Literature Reviews in Software Engineering”. Technical Report EBSE 2007-001, Keele University and Durham University Joint Report. DOI: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.117.471.

Linvill, D. L. et al. (2019). “‘THE RUSSIANS ARE HACKING MY BRAIN!’ investigating Russia’s internet research agency twitter tactics during the 2016 United States presidential campaign”, Computers in Human Behavior 99, p. 292-300. DOI: https://doi.org/10.1016/j.chb.2019.05.027.

Souza, E. T. B. e Conte, T. (2017). “Estimativa de Projetos de Aplicativos Móveis: Um Mapeamento Sistemático da Literatura”, In: Anais do 16º Simpósio Brasileiro de Qualidade de Software (SBQS), Rio de Janeiro, Brasil.

Svantesson, D. J. B. e Caenegem, W. V. (2017). “Is it time for an offence of ‘dishonest algorithmic manipulation for electoral gain’?” Alternative Law Journal 42, p. 184–189. DOI: https://doi.org/10.1177/1037969X17730192.