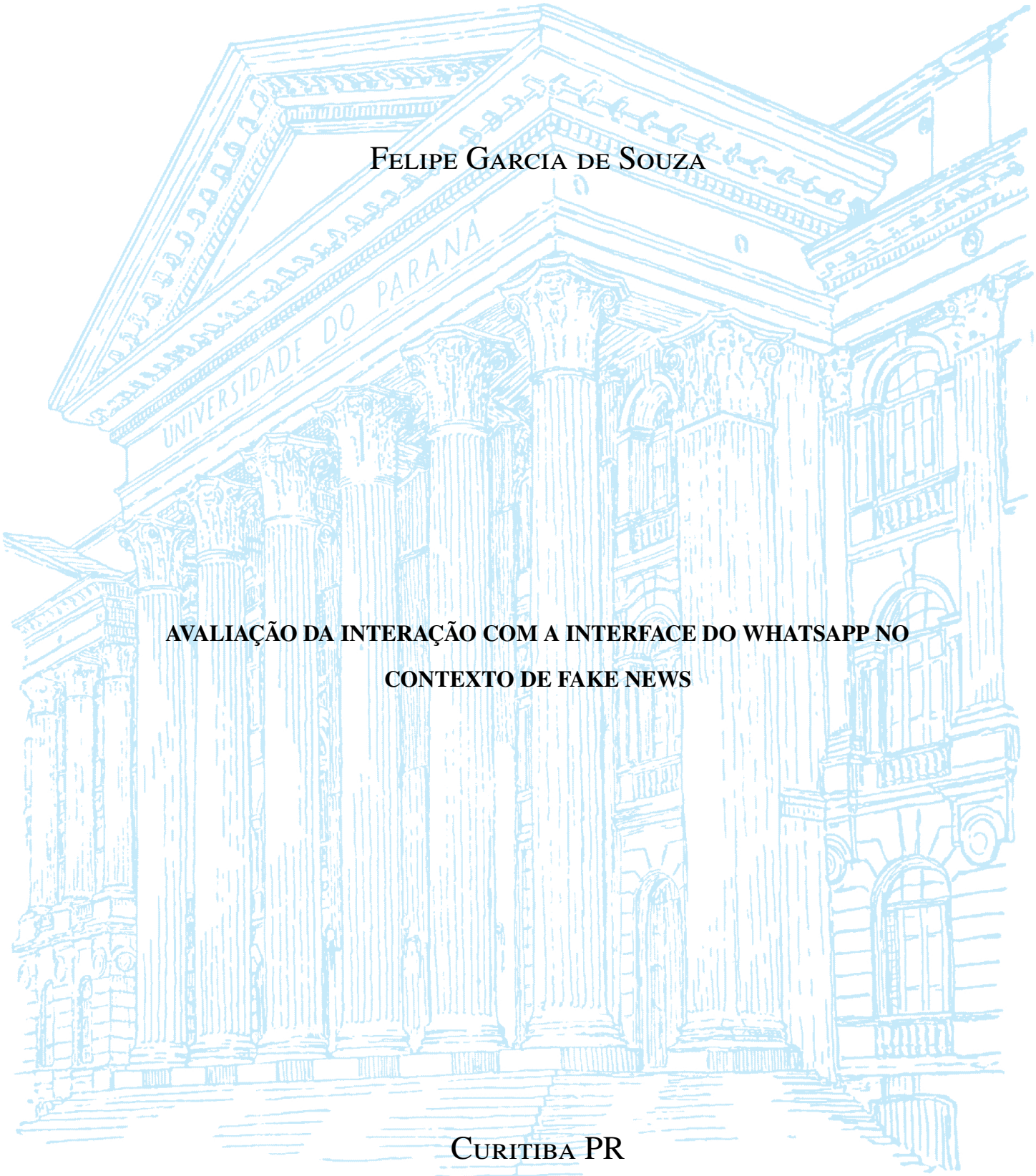
Universidade Federal do Paraná

Felipe Garcia de Souza



AVALIAÇÃO DA INTERAÇÃO COM A INTERFACE DO WHATSAPP NO CONTEXTO DE FAKE NEWS

Curitiba PR
2019

Felipe Garcia de Souza

AVALIAÇÃO DA INTERAÇÃO COM A INTERFACE DO WHATSAPP NO CONTEXTO DE FAKE NEWS

Trabalho apresentado como requisito parcial à conclusão do Curso de Bacharelado em Ciência da Computação, Setor de Ciências Exatas, da Universidade Federal do Paraná.

Área de concentração: *Ciência da Computação*.

Orientador: Prof. Dr. Roberto Pereira.

Coorientador: Alison Puska.

Curitiba PR
2019

Ficha catalográfica

Substituir o arquivo 0-iniciais/catalografica.pdf (PDF em formato A4) pela ficha catalográfica fornecida pela Biblioteca da UFPR a pedido da secretaria do PPGInf/ UFPR.

O conteúdo exato da ficha catalográfica é preparado pela Biblioteca Central da UFPR, a pedido da secretaria do PPGINF. Portanto, não "invente" um conteúdo para ela.

ATENÇÃO: por exigência da Biblioteca da UFPR, esta ficha deve ficar no verso da folha de rosto (que contém o nome do orientador e área de concentração). Cuide desse detalhe quando imprimir as cópias finais.

Ficha de aprovação

Substituir o arquivo 0-iniciais/aprovacao.pdf pela ficha de aprovação fornecida pela secretaria do programa, em formato PDF A4.

Agradecimentos

Agradeço profundamente a todos os amigos e professores que estiveram comigo nessa árdua jornada rumo a essa formação. A Pesquisadora Maíra, por realizar e disponibilizar seu incrível trabalho e os resultados da aplicação do mesmo, para que essa avaliação fosse possível. Ao Professor Roberto, pela paciência, pela orientação e por toda a sua pesquisa que serviu como base para esse trabalho e para o próprio MOvalues. Ao Pesquisador Alisson, pela orientação e permitir que eu fizesse parte do seu mapeamento referente às pesquisas que abordaram o assunto, o que serviu como uma forte base para a elaboração deste trabalho. A todos os participantes do laboratório de IHC que disponibilizaram seu tempo e atenção para que o método pudesse ser realizado. Sem eles esse trabalho não seria possível.A minha pequena por estar comigo e me ajudar a superar alguns dos momentos mais difíceis nessa jornada. Ao grande Professor Humberto que me ajudou tanto nessa etapa final e possibilitou que eu conseguisse expressar meu trabalho da melhor forma possível.

Resumo

O avanço da *Internet* e da inclusão digital vem possibilitando que cada vez mais pessoas tenham acesso a informação de maneira rápida e eficiente. Contudo, a desinformação nos meios digitais também vem se mostrando cada vez mais recorrente, principalmente após os surgimento e a popularização das redes sociais. No Brasil, o *WhatsApp* vem sendo um dos grandes facilitadores na distribuição da desinformação, onde noticias falsas e boatos ganham cada vez mais espaço em grupos e são compartilhados constantemente entre os usuários. Neste trabalho, é realizada uma análise sobre a interface do *WhatsApp* e como esta vem mostrando vulnerabilidades e poucas medidas contra o combate a *fake news*. Com o auxilio de um método desenvolvido por uma pesquisadora na área de Interação Humano Computador (IHC), para análise de valores em redes sociais, é buscado levantar os problemas que o *WhatsApp* possuí que facilitem e também permitem disseminação da desinformação. Além da análise dos problemas encontrados a partir do método, são também citados outros aspectos que fazem do *WhatsApp* um ambiente mais suscetível ao consumo e compartilhamento de noticias falsas. Dentre estes outros aspectos estão: a falta de neutralidade na rede, resultado das operadoras que oferecem uso ilimitado do aplicativo a baixos custos; a falta de informações disponibilizadas para o usuário dentro da aplicação e o uso da aplicação em disparos em massa. Ao final do trabalho, são discutidas propostas que poderiam lidar com esses problemas, a fim de mitigar a pratica do compartilhamento de desinformação e passar a informar melhor o usuário sobre a identificação desse tipo de conteúdo dentro do aplicativo.

Palavras-chave: Desinformação, Rumor, Fake News, avaliação de interface.

Abstract

The constant advance of the Internet and digital inclusion has been allowing more and more people to have access to information in a quick and efficient way. However, misinformation in digital media has also been advancing in a significant manner, especially after the creation and popularization of social networks. In Brazil, WhatsApp has been a major facilitator in the misinformation spread, where fake news and rumors are gaining forces in groups and are constantly being shared between users. This research investigates the WhatsApp interface, how it contains vulnerabilities and lack the necessary resources to deal with fake news. With the help of a method developed by a Human-Computer Interaction (HCI) researcher, with the propose of analyzing values in social networks, we seek to raise platform issues that facilitate the spread of fake news. Besides discussing the problems found using the method, other aspects that make the platform more susceptible to the consumption and sharing of fake news are also mentioned. This includes: the lack of neutrality in the WEB, as result of phone companies offering plans with unlimited use of the app at a low cost; the lack of information made available to the user in the app; the use of the app on bulk services. At the end of the research, we discuss proposals that can be used to address these issues, in order to mitigate the practice of sharing fake information and better inform the user about identifying this kind content when using the application.

Keywords: Misinformation, Rumour, Fake News, interface evaluation.

Lista de Figuras

Lista de Tabelas

Lista de acrônimos

UFPR	Universidade Federal do Paraná
UFMG	Universidade Fedral de Minas Gerais
RSO	Redes Sociais Online
API	Application Program Interface
IHC	Interação Humano-Computador
IBGE	Instituto Brasileiro de Geografia e Estatística
FGV	Fundação Getúlio Vargas
GSMA	Global System for Mobile Communications

Sumário

1 Introdução

O uso de *smartphones* vêm crescendo consideravelmente no decorrer dos últimos anos. Uma pesquisa realizada pela *Global System for Mobile Communications* (GSMA) em 2017 [14], mostra que o número de aparelhos ativos no mundo atingiu a marca de 5 bilhões. No Brasil, esse aumento também espelha a tendência global. Segundo a Fundação Getúlio Vargas (FGV) [23], em um levantamento realizado em 2018, existiam na época uma quantidade próxima a 220 milhões de celulares em funcionamento, contra aproximadamente 207 milhões de habitantes, ou seja, mais de um aparelho por habitante. A possibilidade de acesso à *Internet* pela população também acompanhou o aumento do número de *smartphones*. De acordo com dados do IBGE [18] de 2016 para 2017, o percentual de utilização da *Internet*, nos domicílios brasileiros, subiu de 69,3% para 74,9%. Três em cada quatro pessoas possuíam conexão ativa com a *Internet*. Desse total, cerca de 98% realiza acesso à rede por meio de *smartphones* e 95,5% dos usuários entram na *Internet* para trocar mensagens por aplicativos. A utilização do aparelho para consumo de notícias já ultrapassou os meios tradicionais como pode ser verificado na Figura 1.1 e na Figura 1.2.

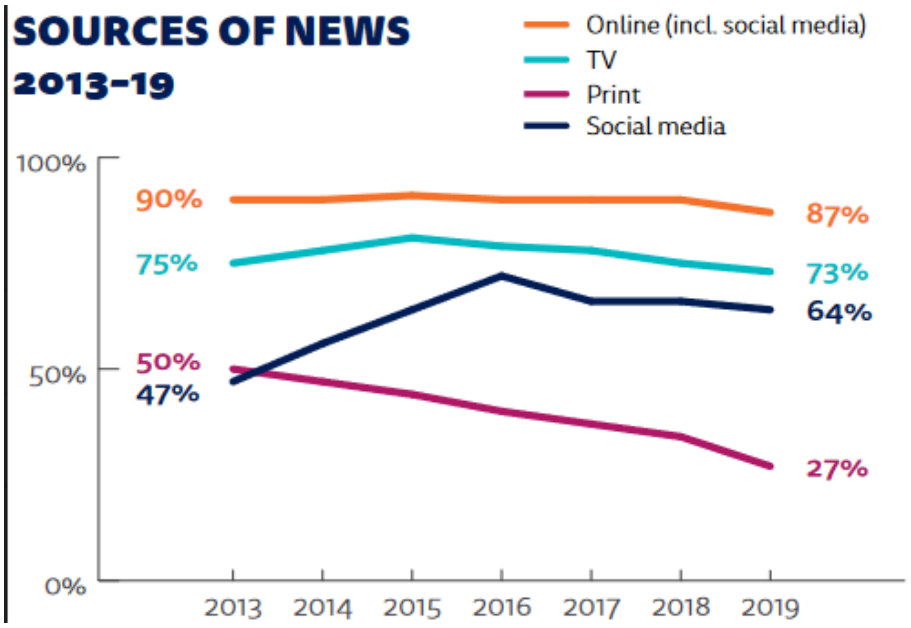


Figura 1.1: Fontes de consumo de notícia pelos brasileiros Fonte: *Reuters* 2019 [28]

Com essa enorme popularização dos *smartphones* e a maior acessibilidade de conexões de *Internet*, o aparelho se tornou algo extremamente recorrente na vida das pessoas. Além disso, 72% dos brasileiros atualmente leem notícias em redes sociais, de acordo com um levantamento do IBGE feito em 2018 [18].

O acelerado avanço da inclusão digital, combinado com as novas maneiras de se consumir noticias, (**e.g., páginas da web, aplicativos**) podem trazer consequências negativas

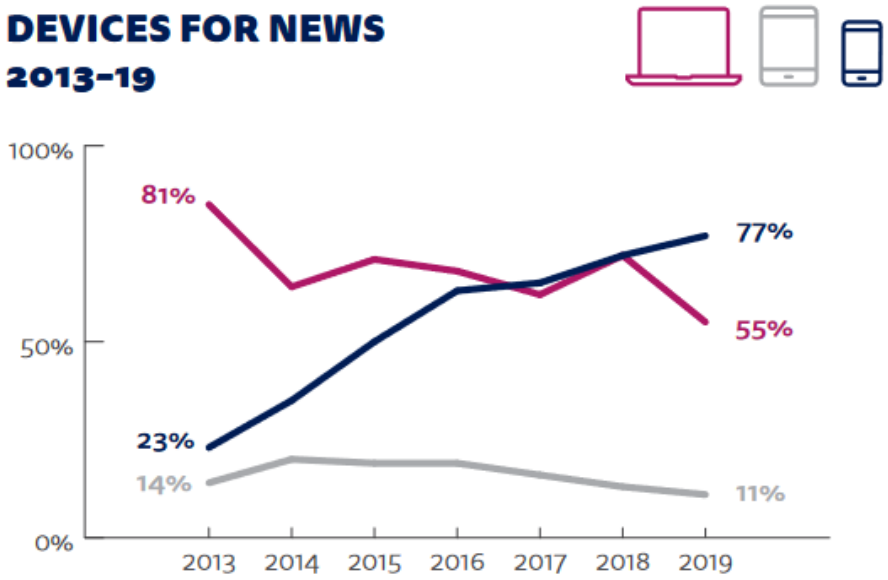


Figura 1.2: Aparelhos utilizados para o consumo de notícias pelos brasileiros Fonte: *Reuters* 2019. [28]

para a sociedade. As informações contidas nas redes sociais não possuem garantia de se originar em fontes confiáveis. Apesar de rumores, boatos e difamação sempre estarem presentes em qualquer meio de difusão de informação, a facilidade do acesso e da transmissão de notícias digitais elevou o problema a um novo patamar, no qual qualquer individuo tem a capacidade de transmissão de informação em nível global com o uso da *Internet*. Qualquer conteúdo transmitido dentro desse ambiente tem uma capacidade de difusão imensa, devido ao alcance global dessas plataformas.

Atualmente a desinformação e a disseminação de notícias falsas tem atingindo proporções alarmantes, atentando de forma rápida contra a vida das pessoas, como em casos de linchamento que aconteceram na Índia devido a boatos espalhados por redes sociais [1]. Além disso, o compartilhamento de noticias falsas vem dividindo pessoas, desde grupos de amigos até familiares, conflitos resultantes do ato de compartilhar *fake news* tem sido frequentes fazendo a distribuição de tal conteúdo estar associado a problemas nas relações interpessoais. Dessa forma, existe uma responsabilidade sobre a aplicação de como seu uso pode afetar as relações de seus usuários pela forma que o aplicativo é usado, esses aspectos referentes a influencia do aplicativo devem ser abordados adequadamente no projeto, desenvolvimento e manutenibilidade da aplicação.

A desinformação é descrita como informação que tem como objetivo enganar, iludir e manipular os sistemas de crenças e opiniões do indivíduo [34]. Redes sociais e aplicativos de mensagem (**e.g., *Telegram*, *WhatsApp*, *Facebook Messenger***) são hoje um dos grandes facilitadores na propagação de *fake news*, que são apresentados em *links* falsos, notícias exageradas, perfis fabricados e entre outros diversos meios de articulação [4]. Tais meios sãos provenientes desse novo ambiente digital, e estão presentes na vida de qualquer usuário conectado a *WEB*, especialmente após o advento de plataformas sociais como *Twitter*, *Facebook* e *WhatsApp*.

No Brasil, o *WhatsApp* mostra-se ser o ambiente mais alarmante [8]. Segundo a própria empresa responsável pelo *software*, o país possui cerca de 120 milhões de usuários que vêm utilizando cada vez mais a plataforma para compartilhamento e discussão de notícias. Um relatório digital da *Reuters* [28] aponta que em 2019 cerca de 53% dos usuários entrevistados usavam a plataforma com o objetivo de se informar, sendo que o mesmo tipo de levantamento realizado em 2018 [27] mostrava uma quantia de 48%. A pesquisa mostra o aumento desse

comportamento, o que é um fato preocupante, tendo em vista como a plataforma facilita a distribuição de conteúdo em massa e como ela carece de formas para checagem de informação.

O WhatsApp vem recebendo novos recursos com o intuito de conter problemas relacionados ao uso da plataforma, como a adição de limites para o número de destinatários em que uma mensagem possa ser enviada [1], tal medida foi implementada em decorrência das recentes mortes provocadas por linchamento na Índia. A empresa também tem planos de incluir pesquisa reversa por imagem, este recurso ainda encontra-se em fase de testes [35].

Devido ao grande aumento do acesso a *internet* e tecnologia, torna-se necessária a aplicação de abordagens para lidar com a transmissão de dados falsos em redes sociais. É preciso desenvolver maneiras de auxílio e reavaliar como as redes sociais são estruturadas, para que os usuários não sejam levados a disseminar notícias falsas recebidas por outros usuários sem saber, possam identificar com mais facilidade conteúdos duvidosos e para que a distribuição proposital de notícias falsas seja dificultada.

O objetivo desse trabalho é realizar uma análise da interface do *WhatsApp* e discutir como ela atualmente pode contribuir no compartilhamento de *fake news*. A relação da plataforma com o consumo de *fake news* pode ser caracterizado por falhas e falta de recursos que fazem com que a plataforma facilite a disseminação de desinformação. Além disso, falta dentro da aplicação ferramentas de auxílio e informativos para ajudar os usuários a validar informações duvidosas. É importante, em aplicações como essa, que seja levado em conta não apenas os aspectos convencionais de usabilidade, adotados para qualquer tipo de *software*, mas também aspectos que considerem como o sistema influenciará na vida pessoal do usuário e na de todos dentro e fora do seu meio de convivência. A relação entre famílias, amigos, casais e qualquer grupo de pessoas que convivem no mesmo ambiente são afetadas pela disseminação de *fake news*.

Para auxiliar no levantamento do uso, aspectos éticos e valores dentro da plataforma, foram analisados os resultados do Método de Avaliação de Valores em Redes Sociais [3], elaborado pela pesquisadora da área de IHC, Maíra Codo Canal. O método, intitulado **MOvalues**, buscou fazer o levantamento dos valores da aplicação sob o contexto de *fake news*. A avaliação foi realizada com seis pessoas, contando com a participação de especialistas do laboratório de IHC da Universidade Federal do Paraná. Na aplicação do método, os avaliadores realizaram o levantamento dos valores das partes interessadas no contexto do *WhatsApp*, e a partir daí foram extraídos problemas e propostas de soluções para serem aplicadas no aplicativo.

Este trabalho busca iniciar a discussão definindo o conceito de *fake news* e mostrando como a desinformação em redes sociais vem sendo um problema ao redor do mundo. Em seguida, serão discutidos alguns trabalhos analisados, mostrando as maneiras com que esse problema tem sido tratado no meio acadêmico. Depois, serão destacadas as questões referentes apenas ao *WhatsApp* que será o foco deste trabalho, sendo a aplicação a principal fonte de propagação de desinformação no Brasil [13]. Apesar do problema no Brasil ser o foco do trabalho, as observações presentes podem ser aplicadas em qualquer aplicação de mensagem instantânea (**e.g., *WhatsApp*, *Telegram*, *Facebook Messenger*, *Line***), e as melhorias podem ser consideradas para qualquer outro país. Serão apontados alguns problemas referentes a utilização da aplicação a partir de resultados levantados pelo uso do método MOvalues, além de outros problemas identificados sem o uso do método. Por fim, serão citadas propostas de como a aplicação pode ser projetada para que o mau uso e a distribuição de conteúdos falsos sejam menos danosos, ou seja, sejam mais controlados ou ate mesmo evitados.

2 A Disseminação de Notícias Falsas

O que hoje conhecemos como *fake news*, consiste em notícias completamente criadas ou manipuladas de maneira tendenciosa. Estas notícias são comumente elaboradas de modo a parecer como vindas de uma fonte jornalística legítima e possuem, na maior parte das vezes, um caráter sensacionalista para que consigam atrair o máximo de atenção [16]. Essas notícias falsas são difundidas em grande escala por meio de redes sociais e aplicativos de mensagens instantâneas. Possuem objetivos como o de obtenção de lucro, por meio de publicidade direcionada a usuários que acessam os sites, [19, 29] ou com propósitos políticos e ideológicos [9], atacando um determinado grupo ou ideia.

As eleições de 2016 nos Estados Unidos foram marcadas por ondas de publicidade duvidosa [38]. Nesse período, surgiram diversas páginas e perfis falsos em diversas redes sociais, compartilhando quantidades massivas de propaganda com viés político a fim de difamar a imagem de seus opositores. Em uma pesquisa feita sobre o acesso dos eleitores nesse período [37] cogita-se que cerca de 27% dos americanos leram pelo menos uma notícia falsa durante a campanha eleitoral. A pesquisa considerava o acesso a um site com notícia comprovadamente falsa por um serviço de *fact-checking* e levando em conta uma quantidade de tempo gasto dentro do site com a postagem, o que pudesse ser o mínimo suficiente para que o conteúdo pudesse ser consumido. Os usuários que buscavam acessar esse tipo de conteúdo eram em maior parte de uma linha de pensamento político mais radical.

O estudo apontou uma relação direta entre o uso de *Facebook* e as visitas a sites de *fake news*. O acesso para esse tipo de sites foram mais numerosos entre eleitores que tinham atividade mais intensa no *Facebook* [37].

O interesse de grupos extremistas também pode ser verificado em um caso ocorrido na Alemanha, no qual o assassinato de uma menina foi utilizado como uma forma para atacar os imigrantes no país, acusando de forma injusta um jovem muçulmano do crime [2]. Um homem, fundador do grupo de extrema-direita anti-Islã *Pegida*, compartilhou uma foto, *link* da página pessoal e o nome completo de um jovem muçulmano, alegando que esse havia cometido o crime, sendo que nenhuma informação tinha sido divulgado pela polícia até o dado momento. A postagem em questão, foi compartilhada diversas vezes e precisou ser desmentida pela polícia, que já havia encontrado e aprendido o verdadeiro suspeito.

A disseminação de notícias falsas também pode ser um risco para a vida da população em geral. Recentemente na Índia, o compartilhamento de *fake news* levou pessoas a serem agredidas, chegando a consequências fatais [1]. Notícias falsas de sequestros de crianças compartilhadas no *WhatsApp* criaram um estado de pânico entre a população fazendo com que grupos de pessoas atacassem homens que estivessem andando sozinhos, afirmando que os mesmos eram sequestradores de crianças, como os grupos de homens que vinham aparecendo em vídeos compartilhados pelo aplicativo. Além disso, também circulam com frequência no país boatos falsos referentes a saúde. Mensagens como a afirmação de que absorventes poderia provocar a morte das mulheres que os usassem ou que sutiãs da cor preta pudessem causar câncer de

mama [7]. A Índia é o país com maior número de usuários do *WhatsApp* possuindo mais de 200 milhões de pessoas utilizando a aplicação.

Assim como nos EUA, redes sociais foram fortemente utilizadas para influenciar os usuários na época das eleições brasileiras em 2018 [6]. Porém, diferente do país Norte Americano que possuía a maior quantidade de mensagens repassadas por redes sociais como *Twitter* e *Facebook*, é por meio do *WhatsApp* que as informações falsas são difundidas em nosso país.

O período eleitoral mostrou que a desinformação e principalmente a sua difusão são muito bem presentes no país. A corrida presidencial foi tomada por acusações e discussões relacionadas a notícias difamatórias, que se espalharam pela rede de forma descontrolada. Durante esse período, circulavam entre os grupos um volume elevado de notícias construídas: *links* falsos, capas de revistas modificadas, áudios sobre algum pronunciamento de uma suposta autoridade e muitos outros [13]. Ocorreu o surgimento de diversos grupos de *WhatsApp* relacionados a política nesse período, esses grupos eram muito ativos e uma fonte produtiva de *fake news* [13].

Mais recentemente, no primeiro semestre de 2019, cortes de R$ 5.8 bilhões estabelecidos pelo governo afetaram todas as universidades federais brasileiras. Diversas instituições ficaram ameaçadas de ter seu funcionamento comprometido, além de milhares de bolsas de pós graduação terem sido canceladas. A partir do momento em que os cortes foram anunciados, as universidades passaram a ser vítimas de ataques provenientes da grande quantidade de notícias falsas compartilhadas no *WhatsApp*. Diversos materiais tentando desmoralizar as universidades e justificar os cortes foram compartilhados em grupos dentro da plataforma de maneira massiva. Segundo o site de checagem “Aos Fatos”, por meio de uma ferramenta de monitoramento desenvolvida na UFMG, foi verificado que o compartilhamento de imagens de estudantes nus em 350 grupos do *WhatsApp* cresceu pelo menos 950 % nas 24 horas que se seguiram após o anuncio dos cortes [33].

As imagens compartilhadas nos grupos do *WhatsApp* mostravam situações fora de contexto, afirmavam que esses tipos de práticas eram recorrentes dentro das instituições. Nunca era mencionado seus reais propósitos, que eram por vezes protestos ou algum tipo de apresentação artística, ambas as situações sem qualquer vínculo com as universidades.

Os protestos que se seguiram em defesa das verbas das instituições tiveram o mesmo tipo de ataque. Novamente, imagens de situações distintas eram usadas para atacar os manifestantes e eram apresentadas fora de contexto, [26, 25] tentando assim, manipular a opinião pública e colocando-a contra as universidades.

3 Trabalhos Analisados

A intensidade na qual as notícias falsas têm se propagado digitalmente e gerado impacto levou a um considerável aumento no número de trabalhos relacionadas ao tema. A partir de um Mapeamento Sistemático da Literatura coordenado pelo pesquisador Alisson Puska, foram analisados trabalhos que abordem questões ligadas a informações falsas e desinformação dentro da área da Computação. A partir dessa pesquisa, foram encontradas soluções que vão desde classificações de credibilidade [5] até detecção de boatos [10]. O foco na analise da desinformação vem ganhando cada vez mais espaço no meio científico nos últimos três anos. Palavras chaves como '*misinformation*' e '*rumours*' que antes apresentavam resultados sem muita relevância dentro do tópico abordado nesse trabalho agora estão se tornando recorrentes e buscam lidar com o problema da grande propagação de informações falsas.

Um tema muito estudado em boa parte das pesquisas diz respeito a detecção das fontes, credibilidade da informação e conteúdo [10, 40, 41]. As formas que lidam com o conteúdo em sua maioria abordam técnicas dentro do aprendizado de máquina tais como processamento de linguagem natural as vezes aliado a fatores humanos para identificar desinformações [5, 15, 17]. As com base em encontrar fontes ou definir a credibilidade da origem de uma notícia consideram modelos de propagação de informação e fatores sociais para mitigação, como a identificação de indivíduos influentes e *"echo chambers"* [32] ou uma análise da rede de contatos de um usuário [10].

Chatterjee e Agarwal [5] propõem um método para a análise da credibilidade de postagens no *Twitter* a partir da extração dos *tweets* e a aplicação de classificadores lineares para filtrar o conteúdo relevante. Por meio de *APIs* utilizadas no *Twitter* os *posts* são extraídos e tratados para que sejam aplicados os classificadores adequados em seguida a credibilidade da informação é calculada usando uma hipótese de decisão majoritária na qual os *tweets* resultantes são classificados como positivos e negativos por uma amostra de pessoas.

O método no entanto, é limitado para o uso em uma plataforma como o *Twitter* na qual toda a informação é compartilhada de maneira pública. No caso de aplicativos de conversa privada como *WhatsApp*, tal solução não é possível dada a maneira que a comunicação é feita, por mensagens privadas sem nenhum interceptador.

Ghafari et al. [10] apresentam um conjunto de algoritmos para predição de confiança usando fatores de contexto social, inspirados em teorias de psicologia social e realizando análises baseadas na decomposição tensorial. O artigo propõe uma nova abordagem a partir do conceito de *web-of-trust*, em que cada usuário possui uma rede de confiança em que cada nó representa uma pessoa e as arestas a confiança entre elas. O método considera vários fatores de contexto social que serão levados em conta para realizar a previsão como o nível de especialidade dentro de um determinado domínio. São considerados dados como tópicos de interesse e comentários, número de seguidores, interação com outros usuários, curtidas e finalmente o nível de intimidade com outros de maneira que eles revelem informações pessoais. Após a coleta deses dados a partir do perfil do usuário, é proposto um algoritmo para prever a confiança na rede social, ou seja o nível de confiança possuído por aquele indivíduo dentro da rede. Experimentos com o algoritmo

mostraram uma previsão mais precisa do que abordagens gerais de *web-of-trust*. Devido o uso de conjunto de algoritmos de predição de confiança a partir de uma análise social juntamente com técnicas de decomposição tensorial para a classificação dos dados, os autores classificam esse método como uma abordagem híbrida.

Essas técnicas demonstram resultados interessantes, porém como a grande maioria das soluções sugeridas para lidar com o problema, são apresentados meios com o foco apenas em dados e métricas quantitativas. Mesmo na proposta de *web-of-trust* [10] os fatores sociais e de relacionamento são convertidas em valores numérico para que se possa ser aplicado o devido algoritmo. Visto isso, seria interessante o uso de abordagens distintas e considerar fatores que sejam pessoais para o usuário, ou seja, como o uso da plataforma pode afetar sua vida ou seu comportamento de forma negativa dependendo do tipo de conteúdo que esse usuário venha a receber.

Pereira e Baranauska [31] descrevem que a tecnologia pode causar impacto no ambiente em que é inserida, podendo resultar em consequências negativas nas pessoas e em todo o ambiente caso não sejam considerados os aspectos culturais de onde se está sendo introduzida a tecnologia. Por isso, é importante em um contexto de design de *software* um comprometimento ético relativo a quem ou onde essa nova tecnologia será aplicada quando a mesma estiver sendo elaborada. Os autores defendem que existem poucas iniciativas relacionadas com os valores das pessoas dentro do estudo de engenharia de *software*, na qual a necessidade de atender a esses aspectos só são notadas quando uma regra social é violada ou um padrão comportamental está quebrado. Assim, os autores elaboram uma forma de se classificar os valores a serem considerados na implementação de um sistema no qual os aspectos culturais dos usuários devem ser levados em consideração. Além disso, esse trabalho foi fundamental para o desenvolvimento do método de aplicação que será utilizado mais adiante.

Hyman [17] argumenta sobre a preocupação a respeito da capacidade do usuário checar a veracidade da informação. Ele discute uma abordagem de união entre automação e checagem humana, na qual é levantada a ideia de que a informação deveria ser passada para o usuário de forma simples, para que o mesmo consiga compreender. Para isso, é necessário um esforço conjunto com a indústria para que métodos melhores e de maior acesso para os usuários sejam desenvolvidos, visto que boa parte dos usuários não se darão ao trabalho de conferir uma notícia uma vez que muitos não tem habilidade, tempo ou até mesmo motivação para isso.

Assim, tal como Hyman argumenta, precisamos levantar as necessidades e propor maneiras que facilitem e estimulem o acesso a verificação da informação garantindo um sistema no qual boatos estejam sempre sendo combatidos.

4 Levantamento de Problemas

Nesta seção, são apresentados os problemas dentro da plataforma que vão desde a falta de notificações e informativos até a possibilidade de usar a plataforma para envios massivos de mensagens. Tais problemas vão além das questões relacionadas ao *design* de usabilidade convencional, por exemplo a aplicação pode desempenhar suas funções corretamente, mas ainda assim influenciar ou prejudicar todos que estão em contato direta ou indiretamente com ela.

4.1 Metodologia de Avaliação

Com o intuito realizar uma análise do uso da aplicação e como a mesma pode afetar a vida de todos que possuem contato direto ou indireto com a aplicação, desde os próprios usuários até qualquer individuo ou grupo que este usuário tenha contato, foram analisados os resultados da execução do Método Orientado a Valores para Avaliação de Redes Sociais **(MOvalues)** [3].

O Modelo apresenta uma proposta de análise mais ampla se comparado aos modelos de análise de usabilidade comumente aplicados para avaliação de aplicações. Outros modelos tendem a realizar análises de caráter predominantemente técnico, sem considerar aspectos como o ambiente em que a tecnologia esta sendo inserida ou qualquer aspecto cultural dos usuários, buscando considerar os aspectos mais informais sob o uso da aplicação. O MOvalues tendo como objetivo a analise do aplicativo sob uma perspectiva orientada a valores, propõe avaliar a influência que o aplicativo possa ter na vida do usuário e de qualquer outra pessoa ou instituição envolvida. Dessa forma, o método pode nos auxiliar na avaliação da influência do aplicativo em todo o meio social de seus usuários. Tal posicionamento é interessante para a análise do software dentro de contexto de *fake news* pois as características que tornam a plataforma mais vulnerável a esse tipo de problemas vão além de um nível técnico e são identificadas se levarmos em conta todo o ambiente em que a aplicação é utilizada.

O MOvalues foi aplicado pela sua criadora, a pesquisadora Maíra Codo Canal. O método é um artefato de inspeção em IHC e busca a realização da analize de aplicações sob uma perspectiva orientada a valores [30, 31]. O método foi originalmente projetado com base em mídias sociais *online*, como *Facebook* e *Twitter*. Esta aplicação do método foi um experimento do uso do MOvalues para avaliação de mídias sociais móveis focando na perspectiva de desinformação e *fake news*. Neste trabalho, serão discutidos os resultados levantados no experimento referentes ao *WhatsApp*.

A elaboração do MOvalues foi baseado nos fundamentos da Semiótica Organizacional [22] lidando com aspectos das dimensões informal, formal e técnica [30]. O informal diz respeito a costumes, crenças e hábitos individuais dos membros de um grupo. Os conceito formais consideram os conceitos já estabelecidos em um meio social, as atitudes que já se tornaram uma convenção ou uma regra a ser seguida por seus membros. E finalmente os aspectos técnicos são definidos como aqueles que são estritamente formalizados, regras e definições escritas e procedimentos que possam ser descritos e executados de maneira concreta.

O MOvalues tem como objetivo, buscar indícios sobre as questões relacionadas às dimensões formais e informais na investigação das causas que podem ser responsáveis ou indutáveis aos diversos comportamentos dos usuários de redes sociais [3], incluindo os relacionados a emoção e afeto, como o ódio e a violência por exemplo.

O método é composto por três etapas, como pode ser visto na figura 4.1. A descrição de cada etapa segue como: Preparação, Análise e Coleta de Dados e por fim Consolidação e Relato dos Resultados e Sugestões de Melhorias. Estas etapas consistem em estágios a serem realizados pelos avaliadores sendo a primeira e a terceira etapa realizadas em grupo e a segunda individualmente. Em seguida serão listados os procedimentos para cada uma das três etapas.

Figura 4.1: Etapas do método de inspeção de valores. Imagens de *Alisson Puska.*

1. **Preparação:** São definidas as partes interessadas no sistema, os valores relevantes a tais partes e porquê cada valor é importante para cada indivíduo, os valores são selecionados pelos participantes a partir de uma lista fornecida pela avaliadora . Tais indivíduos serão qualquer pessoa ou grupo envolvido com o *WhatsApp*, de qualquer forma que seja. Parceiros, fornecedores e concorrentes, qualquer outra aplicação ou organização que contribua ou dispute o espaço do aplicativo são considerados. As possíveis partes interessadas são divididas em quatro categorias: contribuição, fonte ,mercado e comunidade. Na categoria "contribuição", estão os indivíduos envolvidos na criação da mídia social *online*. O grupo "fonte", engloba clientes e fornecedores de recursos. A classe mercado associa parceiros e concorrentes. Por fim, a categoria "comunidade", se refere a comunidade em geral, espectadores e legisladores. Essa etapa é realizada de maneira conjunta entre os avaliadores, para que haja acordo na associação dos valores para cada parte interessada. A Tabela 4.1 mostra a descrição para cada categoria segundo a autora.

2. **Análise e Coleta de Dados:** Cada avaliador, inspeciona as interfaces do sistema, com o intuito de levantar problemas na rede social relacionando tais problemas com os valores de cada parte interessada, selecionadas na etapa anterior. Cada problema identificado deve ser caracterizado dentro dos aspectos formal, informal e técnico [30].

3. **Consolidação e Relatos dos Resultados:** Todos os avaliadores se reúnem para discutir os resultados e apresentar um relatório consolidado único. Juntos, os participantes discutem os seus resultados, julgando os problemas encontrados e decidindo quais deles são considerados importantes o suficiente para serem selecionados e integrados no relatório final. Cada problema escolhido deve ser classificado com um grau de

Contribuição	Fonte	Mercado	Comunidade
Atores ou responsáveis diretamente envolvidos com a concepção e uso da RSO, partes interessadas que usam a RSO diretamente e que são decisivas em sua existência	Clientes e fornecedores de recursos e de informação para a RSO, pessoas que se beneficiarão da informação produzida, outras partes que fornecerão recursos	Parceiros e concorrentes da RSO, potenciais colaboradores ou competidores, outras soluções existentes ou alternativas ao uso da RSO	Expectadores e legislados, comunidade em geral, pessoas que não usarão a RSO mas podem ser afetadas de algum modo por ela

Tabela 4.1: Partes interessadas a RSO que devem ser consideradas para o Movalues

severidade, definido de 0 a 4 conforme a Tabela 4.2. Por fim, o grupo deve discutir e elaborar possíveis soluções para os problemas.

Grau	Descrição
0	"problema sem importância"
1	"problema cosmético - correção ocasional"
2	"problema pequeno - baixa prioridade na correção"
3	"problema grave - alta prioridade na correção"
4	"problema catastrófico - correção obrigatória"

Tabela 4.2: Grau de severidade para os problemas encontrados

4.1.1 Aplicação da Avaliação

A avaliação teve como objetivo a identificação de problemas no *WhatsApp* durante a comunicação particular ou dentro de grupos envolvendo o compartilhamento de mensagens falsas. A aplicação do método foi realizada por sua idealizadora e pesquisadora na área de IHC. A avaliação contou com a participação de especialista de usabilidade do laboratório de Interação Humano Computador, totalizando seis participantes. Entre os participantes: dois doutorandos, três mestrandos e um graduando. Todos os participantes eram da área de computação e com alguma experiência com técnicas de inspeção em IHC (**e.g., Avaliação de Heurística, Teste de Usabilidade, Prototipação**) . O trabalho foi realizado no laboratório do departamento de informática da UFPR onde os participantes não tinham qualquer contato ou influência de aspectos externo e durou cerca de 1 hora e 20 minutos.

Os participantes tiveram inicialmente um pouco de tempo para explorar o sistema. Depois, receberam uma amostra de uma notícia falsa e utilizaram o *WhatsApp* para compartilhar essa notícia, entre os participantes e em um grupo também composto pelos participantes, para então analisar como o aplicativo pode auxiliar ou dar brechas para que a atitude seja feita com mais facilidade. As atividades foram divididas em duas partes: (a) compartilhamento e verificação de desinformação e (b) análise do *WhatsApp* na perspectiva de valores. As atividades da parte (a) foram executadas pelos participantes individualmente e continham 4 tarefas:

1. Compartilhar a informação falsa a um amigo ou grupo;

2. Verificar a veracidade da informação;

3. Buscar informação sobre a mensagem compartilhada dentro do aplicativo (Metadados).

4. Postar a notícia no status do *WhatsApp*.

Durante a aplicação do método, cada avaliador foi instruído a compartilhar uma notícia falsa em um grupo especialmente criado para esse fim, contendo os demais avaliadores, ou então compartilhar a notícia falsa diretamente com outro avaliador por mensagem privada. A notícia falsa foi fornecida pela pesquisadora que estava aplicando o método.

A avaliação resultou no levantamento de diversos problemas ligados a valores dentro da aplicação porém no contexto deste trabalho se tem apenas interesse naqueles que se relacionam com fake news. Desta forma são apresentados em seguida os resultados de (b) de maneira breve que levaram a escolha desses problemas específicos relacionados a desinformação presentes no *WhatsApp*.

4.1.2 Levantamento das Partes Interessadas

Os avaliadores dividiram os grupos que impactam direta ou indiretamente o *WhatsApp*. Para aqueles que têm uma atuação direta dentro da rede, foram selecionados os usuários e a equipe de desenvolvimento. Empresas de marketing e a imprensa foram citados como aqueles que produzem conteúdo para a rede ou que consigam algum benefício com a informação produzida na mesma. Para o mercado, são levados em conta todas as aplicações ou meios que disputem o usuário, oferecendo o mesmo tipo de serviço e que tendem a competir com a RSO analisada. Por fim, a comunidade, que consiste desde pessoas dentro do convívio do usuário até o próprio governo.

A seguir, são descritos os valores escolhidos para algumas das partes interessadas selecionadas juntamente com uma breve descrição sobre o que aquele valor representa para a parte interessada. Serão listadas apenas as partes interessadas e os valores que parecem estar relacionados com a questão de *fake news* e desinformação. As descrições do significado de cada valor foram extraídas do artigo referente ao *modelo* [3], essas descrições foram levantados pela autora do modelo e se originam de diversas fontes diferentes. Foram listados apenas os valores escolhidos pelos avaliadores, que foram associados de alguma forma a qualquer uma das partes interessadas selecionadas pelos avaliadores.

- **Usuários, Pessoa Física**

  a ***Awareness***: Percepção a respeito do sistema, sobre a fácil percepção do usuário sobre o sistema, o que está acontecendo e o quão fácil é ficar a par de mudanças. Notificação, atualizações e registros de atividades são exemplos da aplicação esse valor. Para os participantes esse valor foi selecionado pois o usuário deve sentir que tem a aplicação sob controle. Para isso, o ambiente deve ser claro para os mesmos e bem notificado sobre o que está e o que pode ser feito em tela. No caso do *WhatsApp*, a falta de *Awareness* pode ser evidenciada na falta de informações sobre os recursos da plataforma. Novos recursos como o de possibilitar o bloqueio de usuários e recentemente de bloquear convites para grupos indesejados passam despercebidos pelos usuários da aplicação.

  b **Privacidade**: O valor referente a definir as informações do usuário que podem ser visualizadas por terceiros. É indicado para os avaliadores como o controle de tudo aquilo que parece ser importante para o usuário, que tenha cunho pessoal e privativo. Os usuários não podem definir quais pessoas especificamente poderão

visualizar seus status e seus nomes, quando estão em uma conversa em grupo por exemplo. Tais informações estão abertas a todos os contatos ou até mesmo pessoas que simplesmente possuem o número de telefone da pessoa em questão.

c **Grupos**: Um conjunto de indivíduos com características, situações, interesses ou propósitos em comum. Na aplicação são nos grupos em que estão presentes a maior quantidade de compartilhamento de *fake news*. Os usuários com interesses semelhantes se juntam nesses grupos, em que são discutidas e compartilhadas informações que sejam mais de acordo com as ideias dos integrantes. Essa situação pode vir a resultar na aceitação apenas nos fatos apresentados que são do agrado desse grupo, fazendo com que acreditem no que for mais interessante para essas pessoas.

- **Imprensa**

a **Emoção e Afeto**: Sentimentos dos indivíduos, características que permitem aos usuários expressar seu estado afetivo. O conteúdo produzido pela imprensa pode influenciar diretamente o uso da plataforma pelos usuários. Sendo que um material produzido pela imprensa que contrarie o posicionamento de determinado usuário pode causar uma resposta de raiva pessoal contra esse meio. Isso pode vir a se tornar uma reação coletiva, envolvendo todas as pessoas que tenham uma ração negativa semelhante.

- **Família**

a **Confiança:** Se diz respeito a assegurar a confiança no caráter ou habilidade em alguém e algo. Grupos familiares tem constantemente introduzindo diversos usuários nos meios digitais e principalmente há redes sociais. A falta de costume em uso de meios digitais faz com que esses usuários menos experientes compartilhem e consumam noticias falsas com mais frequência.

- **Governo**

a **Histórico:** Refere-se ao histórico de ações do usuário na rede. A capacidade do sistema de registrar ações do usuário. É de importância para o governo o armazenamento de informações para que as mesmas possam eventualmente ser utilizadas para dados, levantamentos, planejamentos e estatísticas.

b **Normas, Regras e Políticas:** O governo é responsável pela elaboração e aplicação de leis que devem ser seguidas pela população. Tais normas devem assegurar que a aplicação mantenha uma ambiente mais seguro e de uso consciente. O governo poderia estipular limites e responsabilidades para as empresas que atuam dentro do seu território, como em questões de garantir a neutralidade da rede e a garantia da proteção de dados de usuários. Tudo isso é claro, sem ultrapassar o limites da liberdade de cada individuo.

c **Colaboração:** Indicando a possibilidade de se trabalhar juntos, tendo algo que possa ser construído de forma compartilhada. Governos podem auxiliar as redes sociais disponibilizando informações e canais de informação para a população, alertando sobre possíveis notícias falsas e as boas práticas para poder evitá-las no nosso uso diário.

- **Empresas de telefonia móvel**

a **Disponibilidade:** A empresa deve manter o serviço sempre disponível, caso contrário irá perder o interesse e confiança do usuário. É importante não apenas a disponibilidade da plataforma para a troca de mensagens, mas também a disponibilidade de meios para que informações possam ser conferidas pelo usuário.

b **Normas, Regras e Políticas:** Responsáveis por controlar e direcionar o tráfego para todos os usuários, sendo necessária a adoção de políticas que visem melhorar a experiência dos usuários. Tais serviços devem levar em consideração a possibilidade do usuário poder se manter informado não só apenas dentro do ambiente da aplicação.

4.1.3 Inspeção e Consolidação dos Dados:

Na ultima etapa os participantes se reúnem e discutem a conclusão de sua inspeção. A seguir será descrito as considerações a partir da análise e da visão dos participantes. Após o levantamento das partes interessadas, os participantes realizaram uma análise individual da aplicação e levantaram os problemas que foram encontrados. Em seguida, todos os participantes se reuniram e discutiram quais desses problemas deveriam ser considerados para a próxima etapa. Os problemas selecionados, referentes a cada parte interessada, foram então classificadas com um grau de severidade seguindo o modelo da Tabela 4.2.

Compartilhamento: O valor consiste em poder realizar ações e disponibilizem mídias e recursos entre si. O sistema dispõe de diversas ações que permitem o compartilhamento mas algumas questões ligados a esse tipo de ação dentro da plataforma foram classificados como grave (3) pelos avaliadores. Qualquer conteúdo pode ser compartilhado na aplicação, sem qualquer restrição. Os conteúdos compartilhados podem ser desde informações forjadas, com objetivos de persuadir um usuário, até imagens fortes como acidentes automobilísticos ou pedofilia (para esse caso específico o grau de severidade foi taxado como ‘4’). Dessa forma, as informações não possuem nenhuma forma de verificação e a interpretação é de total responsabilidade do usuário, o que abre brechas para que pessoas usem tais recursos para persuadir outros usuários com mais facilidade.

Consentimento Informado: Esse valor descreve a conscientização do usuário por parte do sistema. A capacidade do sistema informar o usuário a respeito de suas ações e o que poderá ser produzido a partir dessas ações. A aplicação permite que mensagens sejam enviadas para qualquer outro usuário, independente de ser alguém fora da lista de contatos. Não existem mecanismos para detecção ou prevenção de spam. O bloqueio deve ser realizado manualmente pelo usuário depois de ter recebido mensagem de um número específico. O *WhatsApp* não possui qualquer opção para aceitar ou recusar novas conversas ou algum tipo de modo privativo ou restrito para receber mensagens, existe apenas a sugestão de bloqueio de contatos fora da lista de mensagem após determinado contato enviar a mensagem. Assim, a qualquer momento um indivíduo pode entrar em contato com outro usuário e compartilhar ou postar o que ele quiser até que o receptor escolha a opção de bloquear o contato. Tal problema foi classificado como grave (3).

Privacidade: A possibilidade de entrar em um grupo com diversas pessoas ou a forma que qualquer grupo público acaba expondo os números telefônicos de seus membros. Ao entrar no grupo, o número de telefone do usuário fica visível para todos os participantes. A falta de privacidade possibilita que qualquer membro envie mensagens e conteúdos de mídia para qualquer membro do grupo. Esse problema também foi classificado como grave (3) pelos avaliadores.

Grupos: Podemos dizer que é nos grupos que o *WhatsApp* apresenta alguns dos problemas mais graves no que diz respeito as *fake news*. Quando a avaliação foi realizada, foi

levantado que não existia um sistema de convites para entrar em um determinado grupo. Qualquer usuário pode incluir outro, sem o consentimento do mesmo, sendo assim, não é possível recusar a entrada em um grupo sendo necessária a saída dele manualmente. No momento, é possível selecionar nas configurações do aplicativo a possibilidade de não ser imediatamente adicionado em um novo grupo, sendo necessário a validação do convite por parte do usuário [36]. A entrada em grupos sem o consentimento dos usuários foi classificada como alta prioridade (4) pelos avaliadores.

Existe também, a falta de controle dos administradores de um grupo criado. Um administrador tem controle de apenas promover outros membros a administradores e retirar membros do grupo. Não há uma forma de controlar a entrada de novos membros, a não ser o método de excluí-los a medida que membros indesejáveis entrem ou sejam convidados para o grupo. Essa questão foi avaliada como baixa prioridade(2).

O terceiro ponto citado sobre problemas no grupo é a falta de se poder notificar alguma atitude ou postagem inapropriada. Se um membro posta conteúdo falso, tendencioso, que proporcione ódio, misoginia, racismo ou seja obsceno, esse membro não corre o risco de ter alguma punição ou advertência pela aplicação. Tendo em vista que não existe nenhuma forma de realizar denúncias dentro do *WhatsApp*. Nesse caso a única forma de lidar com o problema seria a expulsão do membro por um dos administradores. Esse problema foi classificado como grave (3). Apesar desse ponto ser levantado, existe de fato uma opção para a denúncia de usuários dentro da aplicação. Porém, como essa questão foi levantada dentro da aplicação do método podemos concluir que ela não está facilmente acessível para os usuários, o que já pode ser considerado como um problema dentro da aplicação.

O último ponto levantado sobre os grupos é que a forma como o chat é apresentado. A medida que o grupo passa a ter um elevado número de membros e de postagens, se torna impossível de acompanhar ou encontrar coerência em meio tantas mensagens, vídeos, fotos e áudios. Dessa forma, informações podem ser facilmente perdidas e/ou mal interpretadas. Sem contar que essa forma sem qualquer filtro ou organização impossibilita muitas vezes que haja uma discussão organizada.

Confiança: As informações dentro da aplicação não apresentam confiança. Não existe qualquer meio na aplicação que verifique ou afirme a autenticidade de uma informação que vem a ser compartilhada massivamente. Diferentes redes sociais vem implementando formas de indicar quando determinada postagem ou notícia foi checada, podendo ser considerada verdadeira. Contudo, dentro do *WhatsApp* não existe qualquer indicação de confiabilidade, principalmente porque as mensagens são transmitidas de ponto a ponto sem ter qualquer tipo de intermediário.

4.1.4 Conclusão da Avaliação

Pelos problemas levantados com a avaliação, pode se verificar que não existe controle com relação a conteúdos dentro da plataforma. Qualquer informação pode ser compartilhada deliberadamente e sem qualquer aviso no sistema. Não existem formas de validar informações, o aplicativo não disponibiliza qualquer tipo de mensagem de ajuda ou chama a atenção para que os usuários se atentem a este tipo de conteúdo. Qualquer imagem, vídeo ou conteúdo pode ser distribuído sem nenhuma restrição ou bloqueio, dada a natureza da aplicação que funciona com um sistema de criptografia em que apenas a fonte e o destino tem acesso a conversa.

Outro ponto levantado na avaliação cita que: apesar de não haver uma exposição dos dados públicos do perfil de usuário, como ocorre em redes sociais como o *Facebook*, qualquer pessoa que possuir o número de telefone de um usuário pode entrar em contato e enviar conteúdo

falso. Não existindo assim, qualquer tipo de controle sobre os usuários não registradas na lista de contatos que possam vir a transmitir conteúdo.

Mensagens são passadas sem que sua origem possa ser identificada. Não existe uma forma de se conferir a fonte de uma dada mensagem compartilhada na aplicação. Conteúdos compartilhados também não podem ter sua credibilidade confirmada. Portanto, não é possível por meio apenas do *WhatsApp* fazer qualquer tipo de verificação ou pesquisa a respeito da informação recebida. Não há como ter certeza que dada informação foi originada de uma fonte confiável, podendo vir de qualquer usuário desinformado ou mal intencionado.

Muito foi citado na conclusão dos avaliadores sobre como funcionam os grupos e como os administradores poderiam ter mais controle, podendo vir a definir quem pudesse ser aceito em um grupo. Isso poderia reduzir o número de membros indesejáveis compartilhando informações falsas. A possibilidade de um membro bloquear determinada pessoa em um grupo para que suas postagens não sejam disponibilizadas também foi levado em conta, tendo vista que certos usuários podem compartilhar conteúdo que seja considerado ofensivo ou de forma excessiva.

Estes foram os problemas levantados de acordo com a avaliação da aplicação por meio dos avaliadores. Nos próximos tópicos serão discutidos outras questões relacionadas a aplicação que podem vir a contribuir com a distribuição de *fake news*. Tais questões são adicionadas devido a grande repercussão midiática que tem recebido. [11, 21, 24, 6]. Os problemas não se limitam apenas a recursos dentro da aplicação, mas também a serviços que oferecem disponibilidade de uso do aplicativo, além de formas discutíveis de se usar a aplicação comercialmente (**e.g., compartilhamento de *spams*, disparos em massa**) indo até mesmo contra termos legais de compromisso definidos pelo *WhatsApp*[39].

As questões a seguir não foram levantadas a partir do método, porém, acredito que sejam de importante relevância para a avaliação e melhoria do aplicativo. Parte desses problemas, como a falta de neutralidade por parte das operadoras e os problemas com disparos em massa tem sido recorrentemente noticiados nos veículos de noticias [8, 11] enquanto que a questão de informativos na aplicação possa ser verificadas pela falta de informações presentes na aplicação, tanto referente a novos recursos quanto a como fazer bom uso do aplicativo.

4.2 Neutralidade de Rede em Planos de Celulares

Com a grande popularização da aplicação, algumas das principais operadoras de telefonia móvel do país, entre elas Vivo, Tim, Claro e Oi vem oferecido planos para o uso ilimitado do *WhatsApp* [12] sem que esse consuma a franquia contratada. Um exemplo desse tipo de plano pode ser visto na Figura 4.2. Dessa forma, mesmo se o usuário não possuir mais acesso a internet, devido ao esgotamento de seu pacote de dados, o aplicativo continuará funcionando.

Em um primeiro momento esse tipo de plano parece ser bom para o consumidor final, visto que o uso ilimitado permite a utilização da aplicação sem o risco de bloqueio de acesso por falta de franquia. Contudo, essa prática pode gerar problemas pois esses planos podem vir a gerar um ambiente isolado com conexão apenas ao *WhatsApp*, na qual informações não podem ser confirmadas em nenhum outro meio *online*, *site* de notícias ou aplicações. Além de ferir os princípios de neutralidade da rede em que: todas as informações que trafegam pela internet devem ser tratadas da mesma forma, navegar à mesma velocidade, garantir o livre acesso a qualquer tipo de conteúdo na rede, não ferir a autonomia do usuário e não discriminar determinadas aplicações por meio de limitações de consumo de banda.

Esta forma de negócio pode ser associada ao aumento da distribuição de *fake news* [12], na qual o acesso apenas ao *WhatsApp* impossibilita a forma de qualquer checagem *online* pelo usuário. Assim sendo, até mesmo *links* que são compartilhados dentro da plataforma não podem

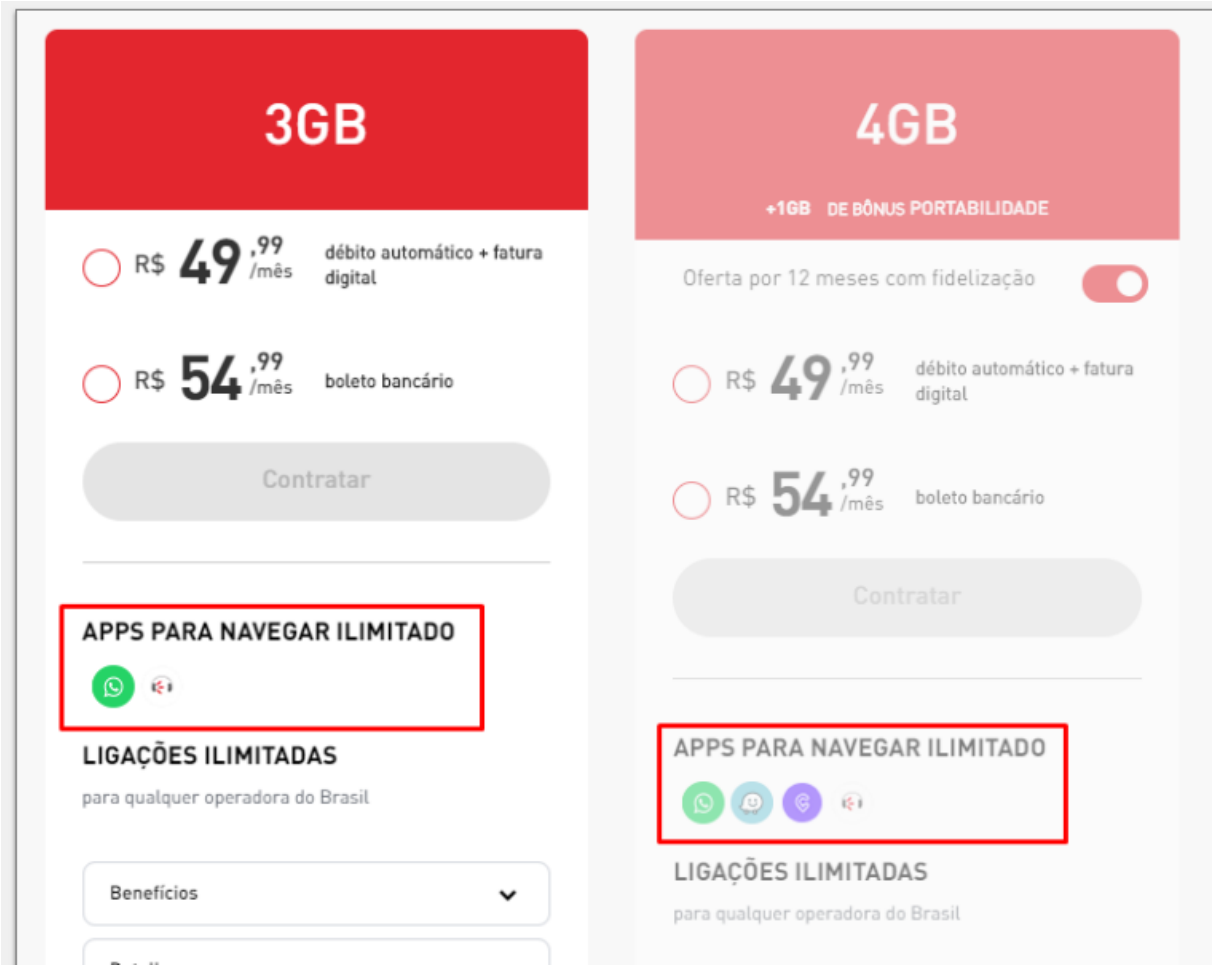


Figura 4.2: Plano oferecido por operadora com uso ilimitado do *WhatsApp*. Retirados do Site da operadora claro (Acessado em 18/11/2019).

ser visualizados, informações não podem ser verificadas e tudo está sujeito a interpretações rasas sem contextualização prejudicando como as pessoas entendem as notícias.

Juntamente com o grande aumento do acesso a internet no Brasil, as classes de pessoas de menor renda, chamadas de classes D e E, vem demonstrando o maior crescimento do número de pessoas com acesso a internet [20], além do aumento sempre constante da quantidade de pessoas que realizaram acesso apenas a partir do *smartphone*. Os planos que oferecem este tipo de pacote com utilização ilimitada do aplicativo, possuem preços chamativos o que certamente estimula usuários de renda mais baixa a contratar o pacote. Quanto mais barato o pacote menor é a franquia de internet oferecida, porém o uso ilimitado do *WhatsApp* sempre esta presente. Ligando esse fato com o grande aumento de pessoas que utilizam a internet apenas pelo *smartphone* [20], Figura 4.3, e o constante aumento do acesso das classes mais baixas a internet, Figura 4.4, é natural que a adesão por pacotes mais baratos e com menos franquia seja maior. Como resultado, temos um ambiente muito mais propício para que as noticias falsas sejam disseminadas, no qual muitas pessoas ficarão isoladas com a conexão apenas do *WhatsApp*.

4.3 Falta de Informação na Aplicação

Existe carência com relação as informações disponibilizada dentro do aplicativo. Dado nosso cenário atual de desinformação massiva na plataforma, era de se esperar que houvesse

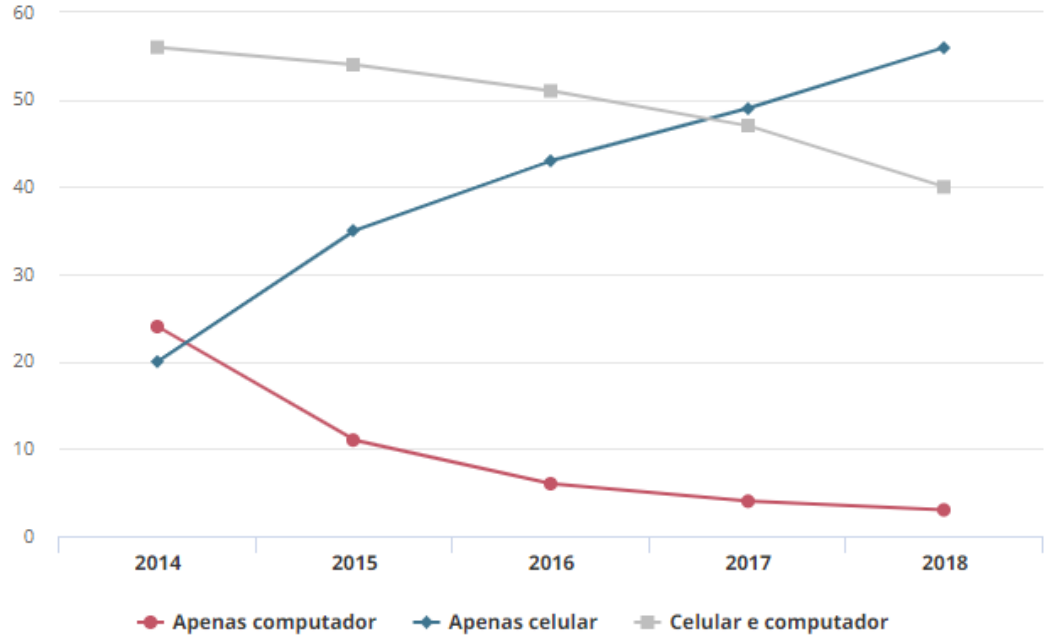


Figura 4.3: Aparelhos usados para realizar acesso a internet no Brasil *Fonte G1* [20]

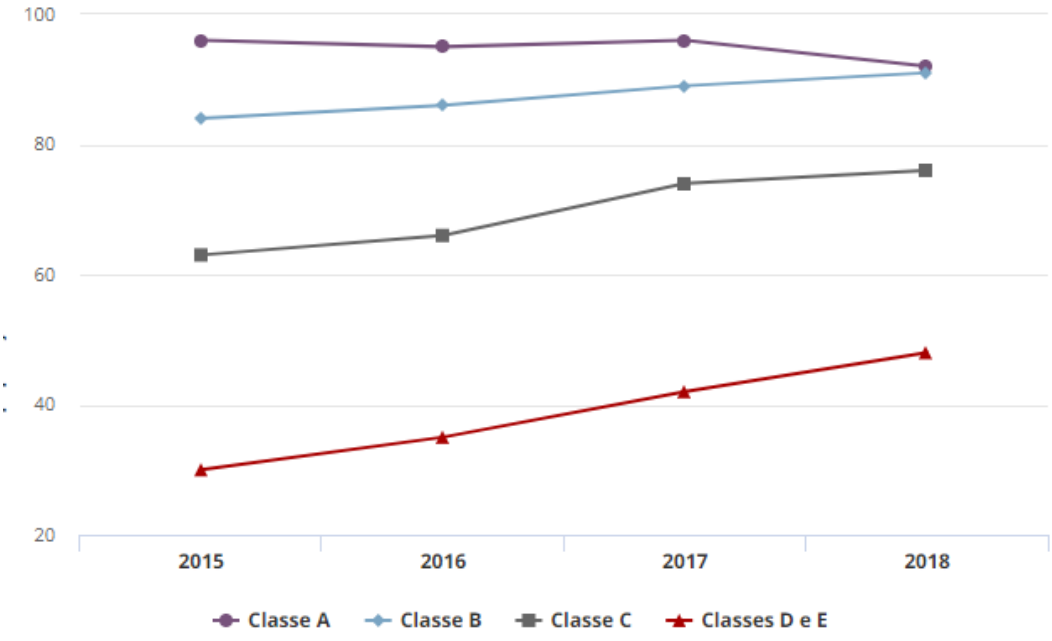


Figura 4.4: Conexão por classe social *Fonte G1* [20]

notificações e mensagens de alertas para conscientizar e advertir os usuários de funcionalidades que possam ajudar na experiência de uso e evitar a disseminação de *fake news*.

A deficiência dessa comunicação pode ser verificada na avaliação de valores em que os avaliadores indicaram que a aplicação não possui formas de denunciar usuários e conteúdo indevido. Contudo a função de fato existe mas não é facilmente encontrada pelos usuários.

A aplicação também carece de informativos de novos recursos. Com a chegada de períodos como eleições presidências, novos recursos são adicionados para tentar reduzir qualquer dano causado por informações falsas. O problema é que não existem avisos informando esses

novos recursos. Algumas dessas novas funções podem ser passivas, não adicionando uma função inteiramente nova para o usuário, como a redução de destinatários no encaminhamento de mensagens, mas outros recursos como o de desativar entradas em grupos, deveriam obrigatoriamente ser notificados de forma clara aos usuários.

4.4 Questões Legais e o Disparo em Massa

Uma prática problemática aplicada com frequência no *WhatsApp* é o disparo em massa. Nesse tipo de prática, uma empresa é contratada para enviar uma quantidade imensa de mensagens para diversos aparelhos. Os números que receberão as mensagens, em alguns casos, são providos por quem contrata o serviço, mas é comum empresas terem um banco de dados próprios. Existem ainda casos em que a aquisição dessa lista tenha sido feita de maneira ilegal, sendo vendida por pessoas dentro de agências telefônicas ou de cobranças por exemplo.

Essa prática não é nova, sendo anteriormente utilizada para distribuição de *e-mails*, muitas vezes como forma de *marketing*. Contudo, esse tipo de serviço pode ser utilizado de forma mais ilícita, como foi o caso durante as eleições nacionais de 2018 que propagaram várias informações falsas [24]. Esta forma de repassar mensagens não é prejudicial apenas por questões políticas, mas porque qualquer indivíduo que tenha a capacidade de pagar por esse serviço pode distribuir a mensagem que quiser, o que pode ser danoso devido a forma com que as informações e noticias estão sendo consumidas hoje, na grande maioria por *smartphones* e cada vez mais a partir de redes sociais e aplicativos de mensagens instântaneas.

De acordo com os termos de serviço do *WhatsApp*, essa prática não é permitida. Segundo a plataforma a aplicação:

> *"(..)não deve ser usada de forma ilícita, obscena, difamatória, ameaçadora, intimidadora, assediante, odiosa, ofensiva em termos raciais ou étnicos, ou instigue ou encoraje condutas que sejam ilícitas ou inadequadas, inclusive a incitação a crimes violentos; (c) envolvendo declarações falsas, incorretas ou enganosas; (d) para se passar por outrem; (e) para enviar comunicações ilícitas ou não permitidas, mensagens em massa, mensagens automáticas, ligações automáticas e afins; ou (f) de forma a envolver o uso não pessoal dos nossos Serviços, a menos que esteja autorizado por nós."* [39].

Entretanto, isso não está impedindo que diversas empresas atuem nesse tipo de ramo e disponibilizem seus serviços para a distribuição em massa de informações falsas.

Até então, não existiam medidas eficientes contra quem abuse e ofereça esse tipo de serviço. Caso uma empresa fosse flagrada fazendo tal atividade o *WhatsApp* simplesmente realizaria o bloqueio dos números em questão, o que não desestimula a prática. Sendo assim, medidas mais efetivas poderiam ser tomadas.

5 Propostas de Soluções

Tendo em vista toda essa gama de problemas levantados anteriormente, são apresentados nessa seção algumas ideias a fim de combater a proliferação de notícias falsas e melhorar como o usuário é informado sobre a natureza de tais notícias, sobre a existência desse tipo de conteúdo dentro da plataforma e quais são as devidas precauções a serem tomadas para não ser levado a acreditar nesse tipo de mensagem.

5.1 *Fact Checking* Dentro da Plataforma

Um dos aspectos a serem listados como alternativa é o uso de *Fact checking*, com acesso dentro da própria aplicação. Essa proposta busca lidar principalmente com o problema causado pela falta de neutralidade da rede, no qual os usuários ficam limitados a apenas as informações que são trazidas diretamente dentro do aplicativo. Com a implementação desse recurso seria possível a consulta de fontes externas sem a necessidade do uso de um pacote de dados de *Internet*.

Esse recurso da a possibilidade de checagem da informação de forma ampla, com acesso fácil, no aplicativo. A consulta se efetivaria a partir de uma busca por palavras chave ou títulos de noticias, que retornariam com um texto de um *site* de checagem de fatos previamente verificado discutindo o assunto buscado pelo usuário. A proposta pode ajudar em muitos casos mas ela não é perfeita. A principio, é necessário que as palavras chaves sejam corretamente ligados a um texto presente no *site* de checagem. Uma palavra muito genérica pode fazer com que uma informação diferente seja trazida para o usuário causando confusão. Além do mais, não há garantia que todo o tipo de notícia buscada tenha uma publicação no *site* buscado. Mas é de se esperar que notícias que já estejam populares estejam prontamente descritas e checadas.

Considerando as vantagens do uso dessa abordagem, o acesso e a transparência de informação dentro da aplicação seriam ainda mais garantidos se essa consulta pudesse ser feito de forma direta no próprio aplicativo, independentemente de acesso a internet por meio de *browsers*. O aplicativo poderia ter uma seção, que a partir de uma *string*, seria possível buscar informações dentro de sites de checagem de fatos variados. Um *bot* coletaria o conteúdo desses diversos sites e o enviaria de forma de mensagem para o usuário. Todos os usuários têm direito a consulta da informação, dessa forma mesmo com as limitações de pacotes de dados oferecidos com a liberação apenas do *WhatsApp*.

Outro aspecto relacionado à checagem dos fatos, seria a indicação de que algum conteúdo já tenha sido consultado e validado em alguma plataforma de checagem. Tal fato oportunizaria, possivelmente induziria os usuários a verificar também a mesma informação. Mensagens devidamente verificadas, exibiriam algum tipo de marcação que elevaria sua credibilidade perante os usuários e criaria o efeito oposto naquelas que não possuem tal marcação. Algo semelhante já acontece no *Facebook*, onde noticias verificadas são sinalizadas. Contudo, esse tipo de verificação ficaria sob a responsabilidade da própria empresa, o que pode não garantir total

confiança caso, por exemplo, a empresa queira agir conforme seus próprios interesses. Em uma situação como essa, a empresa teria facilidade de definir o que é dado como verdadeiro ou não segundo seus próprios interesses. A tecnologia muito tem a contribuir com a informação e o conhecimento, sem que se fira a liberdade, por meio de ferramentas tecnológicas. Contudo, é preciso cautela dentro desses avanços repentinos para que tais aspectos sejam preservados e que as próprias empresas responsáveis pela verificação ou dadas como "neutras"sejam verificadas.

5.2 Responsabilização da Empresa

Para casos como o uso da aplicação de maneira irregular por empresas terceiras, a maior responsabilização do *WhatsApp* poderia ser uma prática para que haja uma maior preocupação em cobrar o uso correto do aplicativo e agir de maneira efetiva quando tais empresas agirem de maneira ilegal. Uma das questões que poderia ser levantada por meios legais, ou por acordos sejam eles celebrados por diferentes atores sociais ou mesmo junto às autoridades, de que de alguma forma a empresa tem responsabilidade pelo uso da ferramenta. A empresa que utilizar a aplicação de maneira ilegal deveria arcar com as consequências e receber uma forma de bloqueio efetivo sobre o uso dessa ou de qualquer aplicação semelhante, para isso seria preciso de uma atuação não apenas do *WhatsApp* mas também de órgãos governamentais. A empresa deve pensar em formas de uma utilização socialmente responsável. Para isso, investimentos, disponibilização de recursos, canais de comunicação, sistemas de registro de ocorrências, instruções de utilização e riscos e orientações aos usuários devem sempre ser buscados.

5.3 Notificações e Informativos

Dada a falta de informações dentro do aplicativo e a falta de transparência a respeito de novos recursos alguns pontos deveriam ser melhorados no projeto do aplicativo. O *WhatsApp* deveria trabalhar com notificações e avisos educativos sobre o uso consciente da aplicação, como alertas sobre possíveis notícias falsas circulando entre os usuários. Uma aba nova poderia ser implementada com informações sobre leitura crítica de notícias, com dicas sobre informações que parecem ser adulteradas. Outro meio de implementação seria a constituição de uma janela modal com informativos diários oferecendo *links* de checagem e colocando em destaque as últimas *fake news*.

Em muitos casos os novos recursos do aplicativo não são conhecidos pelos usuários. No caso de uma campanha extensa na sociedade de combate às *fake news*, deveria se valorizar a notificação ao usuário de novos recursos disponíveis relativos à verificação da veracidade de informações. Essas notificações deveriam ser claras e repetitivas, dando consciência dos novos recursos de maneira mais enfática. Uma solução simples poderia ser utilizar o próprio sistema de mensagens do aplicativo e notificar o usuário de informativos e novos recursos por um perfil que possuísse algum tipo de identificação, que garantisse sua credibilidade. A Figura 5.1 mostra um esquema simples de como a empresa poderia notificar o usuário de novos recursos.

Além das notificações fornecidas pelo aplicativo, seria interessante se os contatos também dispusessem de um meio de alertar uns aos outros de maneira eficiente. Um recurso semelhante aos *stories* poderia ser utilizado, onde um usuário indicaria uma mensagem ou mídia como falsa, notificando todos seus contatos uma possível identificação de noticias falsas.

Com relação a mensagens encaminhadas, existe atualmente um indicativo de quando uma mensagem provém de um compartilhamento, não sendo escrita diretamente pelo usuário que a enviou. Poderia haver mais informações nesse tipo de mensagem, como por exemplo, sua

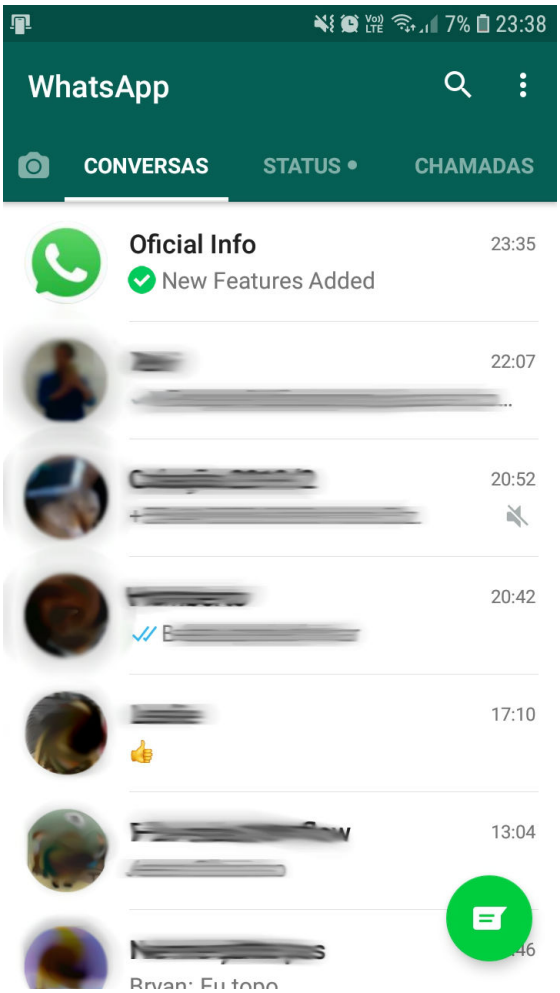


Figura 5.1: Proposta de notificações por mensagens oficiais

origem e se a mesma foi alterada desde seu envio original, juntamente com a informação de quem as alterou e principalmente por quantas pessoas aquela mensagem foi repassada até chegar o seu destino. Informações como essas poderiam vir a se tornar de grande auxilio, fazendo com que pelo menos alguma parcela dos usuários se atente devido a natureza de uma mensagem compartilhada de forma massiva ou por uma informação alterada.

5.4 Controle de Disparos em Massa

Devido o grande números de usuários do *WhatsApp* e sua grande influência social, se faz necessário o acompanhamento dos disparos massivos - e ilegais - de mensagens no aplicativo. O controle de disparos massivos não é algo particular, não é algo que irá interferir na intimidade de ninguém. Pelo contrário, o fato de ser massivo já demonstra a sua dimensão social. Dessa forma precisa ser acompanhado, fiscalizado e devidamente lidado. O que se deve discutir é em quais níveis as empresas donas de aplicativos de mensagens devem acompanhar esses disparos massivos, e a partir de que nível deve ser acompanhado pelas autoridades.

Nos parece consueto que o acompanhamento e fiscalização desses eventos de disparos em massa deve forçosamente ocorrer de alguma forma.

Em tempos recentes a empresa do *WhatsApp* já vem tomando algumas medidas nesse sentido [8]. A organização pretende adotar medidas legais contra as empresas que recorrerem

ilegalmente a essa prática [21]. Anteriormente a única medida tomada era o bloqueio do número por meio do qual foram disparadas as mensagens.

Em nível de implementação do sistema, poderia-se fiscalizar os trafego de mensagens repetidas. A ocorrência de disparo de uma mesma mensagem, por diversos aparelhos, em um curto espaço de tempo já pode ser identificado como uma ação de disparo em massa. Idealmente, o aplicativo poderia impedir o envio dessas mensagens antes que chegassem aos seus destinatários. Contudo, tal ação implicaria em problemas referentes a questão da privacidade, no qual seria necessário a intervenção das mensagens.

5.5 Limitação de Mensagens

A limitação das mensagens compartilhadas é outra medida adotada pela plataforma na tentativa de se conter o avanço das *fake news*. O encaminhamento de mensagens foi limitado para até 5 destinatários por vez. Tal medida foi aprovada em decorrência das recentes mortes provocadas por linchamento na Índia [1]. Diante da onda de correntes passadas entre os contatos foi decidido a implementação dessa limitação. Contudo, tal impedimento se apresenta ainda muito raso e não tão eficaz quanto deveria.

O limite de mensagens compartilhadas define que apenas cinco encaminhamentos sejam feitos por vez, mas não estabelece nenhuma restrição de tempo de espera entre eles, além de que, grupos são considerados como sendo um único contato. Dessa forma, é possível compartilhar para 5 grupos diferentes, em que cada grupo pode conter até 256 membros. Portanto, apesar dessa função estar implementada e de fato auxiliar no problema, se vê a necessidade de impor restrições mais fortes e eficazes como impedir que a mesma mensagem seja compartilhada dentro de um intervalo de tempo definido. Também deve existir algum tipo de controle de *spam*, onde um usuário que venha compartilhar mensagens repetidas vezes dentro de um curto espaço de tempo, perca a capacidade de usar o recurso durante algum tempo.

5.6 Verificação e Busca Reversa de Imagem

As imagens são recorrentes nos casos de desinformação. Um mecanismo que possibilite a verificação de uma imagem seria uma ótima adição ao aplicativo. De forma semelhante, já existe um recurso presente em ferramentas de busca em navegadores não *mobile*, no qual um algoritmo varre a rede em busca de imagens que sejam semelhantes a escolhida pelo usuário. Esse novo recurso permitira o usuário a pelo menos encarar mais de uma interpretação e versão para uma mesma imagem, permitindo uma melhor análise dessa imagem em questão. Tal recurso já vem sendo desenvolvido pela empresa e se encontrava em fase de testes no momento da escrita desse trabalho. [35].

5.7 Maior Controle de Privacidade

Outro aspecto a ser analisado pelo aplicativo diz respeito à questão da privacidade, que deve ser trabalhada de forma mais satisfatória. O usuário poderia definir como outros usuários possam enviar mensagens para sua conta. Restringindo dessa forma, o recebimento de mensagens daqueles que estão fora da lista de contatos, exigindo assim algum tipo de identificação e apresentação acompanhada de uma mensagem quando o envio de mensagem é efetuado pela

primeira vez. Esses recursos possibilitariam um melhor controle de *spams* como os resultados de mensagens multi-direcionadas.

5.8 Funções de Acesso a Grupos

Por fim, vale também citar como o aplicativo está redefinindo as opções de acesso a grupos. Antes, qualquer indivíduo tinha a possibilidade de adicionar outro em um grupo, bastasse possuir o número de telefone. Agora, usuários podem dar ou negar a permissão para ser adicionado em um grupo, podendo definir que apenas contatos possam adiciona-los diretamente em grupos ou apenas os contatos de uma lista pré-definida. No entanto, é necessário habilitar esta opção dentro dos ajustes do aplicativo, como pode ser visto na Figura 5.2. Um recurso como esse poderia ser padronizado para estar ativado desde o primeiro uso, pelo menos restrito por padrão para que apenas os contatos tivessem tal permissão, A possibilidade de poder ser adicionado a qualquer grupo e por qualquer outro usuário do aplicativo podem ser prejudicial para a privacidade de quem usa o aplicativo.

Figura 5.2: Passos necessários para desativar entradas em grupos sem autorização do usuário.

6 Conclusão

Devida a facilidade de acesso, baixo custo e aumento da inclusão digital, o *WhatsApp* se tornou uma das principais formas de comunicação e consumo de notícias no Brasil.

Inserida na crescente onda de desinformação e propagação de boatos pela rede, essa plataforma também veio a se estabelecer como o predominante meio de se transmitir notícias falsas. O compartilhamento de tais noticias vem resultando tanto em problemas dentro da sociedade, levanto a afetar a segurança e bem estar da população quanto que o hábito do compartilhamento tem dividindo as pessoas e prejudicando relações interpessoais. A partir disso, se vê como necessária e imprescindível a discussão sobre como essa plataforma pode facilitar esse tipo de disseminação e quais as medidas cabíveis para diminuir tais problemas. Nesse contexto precisamos afirmar que a empresa deve arcar com a responsabilidade social de investir em mecanismo e recursos que possibilitem o melhor controle na plataforma, garantindo o seu uso de maneira mais ética.

Neste trabalho, foram apresentados os impactos e a facilidade da disseminação de notícias falsas no Brasil por meio do *WhatsApp*. Buscando analisar os problemas da plataforma considerando também o contexto social, verificamos o resultado do método para avaliação do design/projeto sob uma perspectiva orientada a valores em redes sociais *online* desenvolvido pela pesquisadora Maíra Codo Canal, MOvaules, aplicado no consumo de *fake news* pelo *WhatsApp*. A partir dos resultados de sua aplicação, foram levantados diversos problemas no uso da plataforma. Além das falhas encontradas com o método, foram também apresentadas outras questões consideradas de extrema importância, tais como: A fraca capacidade da aplicação de informar o usuário de problemas com desinformação e de compartilhar novos recursos implementados na plataforma; falta de contenções referentes aos envios de mensagem em massa; a falta de neutralidade em pacotes de telefonia que proporcionam um ambiente sujeito a informações contidas apenas dentro do aplicativo, sem qualquer forma de consulta externa.

Após levantar os problemas da plataforma, foi buscado indicar possíveis soluções a partir da aplicação do método, que direcionou a análise do aplicativo para que fossem levados em conta questões técnicas e sociais que também afetem o cotidiano dos usuários e de qualquer grupo ou individuo envolvido com mesmo. Assim, as soluções para lidar com os problemas nas redes sociais podem ser implementadas ou colocadas em prática para conter a grande quantidade de compartilhamento de informações duvidosas e informar melhor o usuário de como evitar esse tipo de notícia. Também foram levantados problemas envolvendo a plataforma que vem sendo noticiado com frequência na mídia, como os disparos em massa, e como tais problemas também poderiam ser trabalhados para conter a disseminação da desinformação. Foram discutidos soluções que já vem sendo elaboradas pela equipe de desenvolvimento do *WhatsApp* e questionada a eficiência dos recursos já implementados. Parece que se precisa de maior empenho e compromisso com a questão, que deve ser levada muito a sério, tendo em vista os efeitos que podem e que têm sido causados na sociedade.

A avaliação de um ponto de vista social foi importante para o levantamento dos resultados. Foi possível fazer a avaliação de como mensagens falsas podem ser facilmente compartilhadas em

conversas privadas e grupos. Aspectos relacionadas a privacidade e ao desconforto que algumas questões podem causar aos usuários também foram consideradas, tais como o recebimento de mensagens de qualquer outra pessoa e a inserção não consensual em grupos. Esses tipos de problemas foram além dos aspectos técnicos nos quais, apesar dos recursos do aplicativo funcionarem como deveriam sob uma perspectiva funcional, os mesmos não prezam totalmente pelos valores e aspectos culturais dos usuários que vem a ter contato com a aplicação.

Percebemos que é necessário maior empenho e compromisso com a questão do uso do aplicativo por terceiros. Sociedade, autoridades e empresários devem reconhecer e trabalhar com maior efetividade para o equacionamento de tais problemas.

Nesse contexto, argumentamos novamente que a empresa e o aplicativo não podem ser eximidos de suas responsabilidades. Ainda é importante sustentar que alguns controles, respeitando a liberdade e a privacidade da ferramenta, devem ser implementados a fim de se coibir as práticas de desinformação. Controles sobre os disparos massivos, recursos para denúncias e para checagem das informações são fundamentais nesse sentido.

Assim, o *WhatsApp* ou qualquer outra aplicação social de mensagens direta, devem ser elaboradas considerando os impactos que seus recursos podem causar no meio social. Não somente no Brasil, mas em qualquer outro lugar que aplicações desse tipo venham a ser usadas, visto que todos os aspectos citados no texto podem ser considerados em qualquer tipo de cultura que essas aplicações sejam inseridas. Todo o desenvolvimento de *software* precisa se atentar a essa nova realidade, na qual os aplicativos tem uma influência considerável na sociedade. O desenvolvimento desses *softwares* deve se adaptar para que problemas como os levantados aqui sejam levados em conta desde a concepção do sistema e as soluções devem ser mantidos e melhorados durante toda a vida útil do sistema. Visto que em uma era na qual desinformação pode ser transmitida de maneira instantânea é preciso planejamento para que não haja chances de que boatos e informações enganosas passem despercebidos. Sendo assim, métodos, técnicas e artefatos que sejam informados por uma visão sociotécnica, como o MOvalues, são atualmente essenciais para que aplicações sejam projetada/avaliadas levando em consideração todos os aspectos socioculturais de seus usuários, garantindo que os aplicativos apresentem o menor dano possível a sociedade.

Referências

[1] Ant Adeane. Whatsapp limita mensagens na Índia após notícias falsas levarem a linchamentos. *BBC*, 2018. Acessado em 24/06/2019.

[2] Ant Adeane. How the far right hijacked a teenager's murder. *BBC*, 2019. Acessado em 24/06/2019.

[3] Maíra Codo Canal. *MOvalues: Um método Orientado a Valores para a Avaliação de Redes Sociais Online*. PhD thesis, Universidade Federal do Paraná (UFPR), 2019.

[4] Deanna D Caputo, Shari Lawrence Pfleeger, Jesse D Freeman, and M Eric Johnson. Going spear phishing: Exploring embedded training and awareness. *IEEE Security & Privacy*, 12(1):28–38, 2013.

[5] Ram Chatterjee and Samiksha Agarwal. Twitter truths: Authenticating analysis of information credibility. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, pages 2352–2357. IEEE, 2016.

[6] Deborah Ramos da Silva135 and Maria Teresa Miceli Kerbauy136. Eleições 2018 e a forte influência das redes sociais. *Liberdade de Expressão Questões da atualidade*, page 125, 2019.

[7] Shobita Dhar. Fake health advice forwards a new headache for doctors. *Times of India*, 2018. Acessado em 24/06/2019.

[8] Eduardo Militão e Aiuri Rebello. Whatsapp bane ao menos 1,5 mi de contas no brasil por robôs e fake news... - veja mais em https://noticias.uol.com.br/politica/ultimas-noticias/2019/09/30/whatsapp-fake-news-robos-envio-em-massa-eleicoes-2018-contas-banidas.htm?cmpid=copiaecola. `https://noticias.uol.com.br/politica/ultimas-noticias/2019/09/30/whatsapp-fake-news-robos-envio-em-massa-eleicoes-2018-contas-banidas.htm`, 2019. Acessado em 26/11/2019.

[9] Nicholas Fandos and Kevin Roose. Facebook identifies an active political influence campaign using fake accounts. *New York Times*, 2018.

[10] Seyed Mohssen Ghafari, Shahpar Yakhchi, Amin Beheshti, and Mehmet Orgun. Social context-aware trust prediction: Methods for identifying fake news. In *International Conference on Web Information Systems Engineering*, pages 161–177. Springer, 2018.

[11] Juliana Gragnani. Como planos de celular com facebook e whatsapp ilimitados podem potencializar propagação de notícias falsas. *BBC Brasil*, 2018.

[12] Juliana Gragnani. Como planos de celular com facebook e whatsapp ilimitados podem potencializar propagação de notícias falsas p. *BBC*, 2018. Acessado em 24/06/2019.

[13] Juliana Gragnani. Um brasil dividido e movido a notícias falsas: uma semana dentro de 272 grupos políticos no whatsapp. *BBC*, 2018. Acessado em 26/11/2019.

[14] GSMA Intelligence. Number of mobile subscribers worldwide hits 5 billion. https://www.gsma.com/newsroom/press-release/number-mobile-subscribers-worldwide-hits-5-billion/, 2017. Acessado em 24/06/2019.

[15] Arushi Gupta and Rishabh Kaushal. Improving spam detection in online social networks. In *2015 International conference on cognitive computing and information processing (CCIP)*, pages 1–6. IEEE, 2015.

[16] Elle Hunt. What is fake news? how to spot it and what you can do to stop it. *The Guardian*, 17, 2016. Acessado em 24/06/2019.

[17] Joshua Hyman. Addressing fake news: Open standards & easy identification. In *2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON)*, pages 63–69. IEEE, 2017.

[18] IBGE. Internet chega a três em cada quatro domicílios do país. https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/23445-pnad-continua-tic-2017-internet-chega-a-tres-em-cada-quatro-domicilios-do-pais, 2018. Acessado em 24/06/2019.

[19] Nir Kshetri and Jeffrey Voas. The economics of "fake news". *IT Professional*, 19(6):8–12, 2017.

[20] Thiago Lavado. Uso da internet no brasil cresce, e 70% da população está conectada. https://g1.globo.com/economia/tecnologia/noticia/2019/08/28/uso-da-internet-no-brasil-cresce-e-70percent-da-populacao-esta-conectada.ghtml, 2019. Acessado em 15/10/2019.

[21] Ronaldo Lemos. Whatsapp diz que vai tomar medidas legais contra envio de mensagens em massa. *Folha de São Paulo*, 2019. Acessado em 24/06/2019.

[22] Kecheng Liu. *Semiotics in information systems engineering*. Cambridge University Press, 2000.

[23] Fernando S. Meirelles. 29ª pesquisa anual, 2018 administração e uso da ti nas empresas. https://eaesp.fgv.br/sites/eaesp.fgv.br/files/pesti2018gvciappt.pdf, 2018. Acessado em 24/06/2019.

[24] Patrícia Campos Mello. Whatsapp admite envio maciço ilegal de mensagens nas eleições de 2018. *Folha de São Paulo*, 2019. Acessado em 25/11/2019.

[25] Luiz Fernando Menezes. Foto de faixa com 'fora bolsonaro' e 'liberdade para lula' é de março, não de ato pela educação. *Aos Fatos*, 2019. Acessado em 24/06/2019.

[26] Luiz Fernando Menezes. Meme que critica manifestações pela educação usa fotos de protestos antigos. *Aos Fatos*, 2019. Acessado em 24/06/2019.

[27] Newman Nic, Richard Fletcher, Antonis Kalogeropoulos, David AL Levy, and Rasmus Kleis Nielsen. *Reuters Institute Digital News Report 2018*. Reuters Institute for the Study of Journalism, 2018.

[28] Antonis Kalogeropoulos Nic Newman with Richard Fletcher and Rasmus Kleis Nielsens. *Reuters Institute Digital News Report 2019*. Reuters Institute for the Study of Journalism, 2019.

[29] Abby Ohlheiser. This is how facebook's fake-news writers make money. *Washington Post*, 18, 2016.

[30] Roberto Pereira, M Cecília C Baranauskas, and Sergio Roberto P da Silva. Social software and educational technology: informal, formal and technical values. *Journal of Educational Technology & Society*, 16(1):4–14, 2013.

[31] Roberto Pereira and Maria Cecília Calani Baranauskas. Value pie: a culturally informed conceptual scheme for understanding values in design. In *International Conference on Human-Computer Interaction*, pages 122–133. Springer, 2014.

[32] Alexander Petrov and Olga Proncheva. Modeling propaganda battle: Decision-making, homophily, and echo chambers. In *Conference on Artificial Intelligence and Natural Language*, pages 197–209. Springer, 2018.

[33] Amanda Ribeiro. Após cortes no mec, envio de imagens de estudantes nus cresce 950em grupos de whatsapp em 24 horas. *Aos Fatos*, 2019. Acessado em 24/06/2019.

[34] Miroslav Tudjman and Nives Mikelic. Information science: Science about information, misinformation and disinformation. *Proceedings of Informing Science+ Information Technology Education*, 3:1513–1527, 2003.

[35] Felipe Ventura. Whatsapp testa busca reversa de imagem no google para combater fake news. `https://tecnoblog.net/281977/whatsapp-busca-reversa-imagem-google/`, 2019. Acessado em 24/06/2019.

[36] Felipe Ventura. Whatsapp: todo mundo agora pode exigir convite para ser adicionado a grupos. `https://tecnoblog.net/313361/whatsapp-todos-podem-exigir-convite-adicionar-a-grupos/`, 2019. Acessado em 22/11/2019.

[37] Lilian Venturini. Qual o impacto das fake news sobre o eleitor dos EUA, segundo este estudo. *Nexo Jornal*, 2018. Acessado em 24/06/2019.

[38] Mike Wendling. The (almost) complete history of'fake news'. *BBC Trending, January*, 22, 2018.

[39] WhatsApp. Dados jurídicos do whatsapp. `https://www.whatsapp.com/legal/?lang=pt_br`, 2016. Acessado em 24/06/2019.

[40] H Xia and J Liu. Credibility analysis of comments of virtual community based on text similarity computing. *Journal of Modern Information*, 31(9):33–37, 2011.

[41] Kaiyong Xu, Jiayan Wang, Qihan Xu, Jinquan Lu, and Qingjun Yuan. Micro-blog user trustworthiness evaluation. In *2017 11th IEEE International Conference on Anti-counterfeiting, Security, and Identification (ASID)*, pages 79–84. IEEE, 2017.