

Uma Análise do Impacto do Anonimato em Comentários de Notícias Online

Julio Reis, Manoel Miranda, Lucas Bastos, Raquel Prates, Fabrício Benevenuto

¹Universidade Federal de Minas Gerais (UFMG) – Brasil

{julio.reis, manoelrmj, lucasmbastos, rprates, fabricio}@dcc.ufmg.br

Abstract. *In the last years, with the growth of the Internet as a mean of communication, a large number of people has changed the way they consume news, replacing the traditional newspapers and magazines by virtual and online versions and/or web blogs. One of the reasons online news have attracted users is because of the opportunity to comment and express their feelings and opinions about them. In this work, we analyze the content of comments of the online newspaper Reuters news in which we consider the users can comment anonymously, and Facebook in which comments are associated to an identity, often their real world identity. Our goal is to investigate whether users behave differently considering their anonymity or not. More specifically, we intend to verify whether there are significant linguistic differences between comments posted in an anonymous environment (Reuters) and in an environment in which their identity is usually known (Facebook). In addition, we propose to analyze the reactions of users when they are anonymous or not, to headlines with different sentiment associated with them. Our results indicate that linguistic characteristics change according anonymity, and also that users tend to be more negative when they are anonymous.*

Resumo. *Nos últimos anos, com o crescimento da Internet como meio de comunicação, um grande número de pessoas mudou a maneira de consumir notícias, substituindo os tradicionais jornais e revistas impressos por suas versões virtuais online e/ou web blogs. Uma das razões pelas quais as notícias online têm atraído usuários é a oportunidade de comentar e expressar seus sentimentos e opiniões sobre as mesmas. Neste trabalho, analisamos o conteúdo de comentários online postados em notícias do jornal Reuters, no qual consideramos que o usuário posta anonimamente, e do Facebook, onde o usuário está associado a uma identidade, frequentemente a sua identidade no mundo físico. Nossa objetivo é investigar se existe variação no comportamento do usuário considerando o seu anonimato ou não. Mais especificamente, trabalhamos no sentido de verificar se existem diferenças linguísticas significativas entre o comentário postado no ambiente onde é conferido ao usuário certo grau de anonimato (jornal Reuters) e no ambiente onde espera-se que este grau de anonimato não exista (Facebook). Além disso, nos propomos a analisar as reações dos usuários quando eles estão anônimos ou não, para manchetes com diferentes sentimentos associados a elas. Nossos resultados indicam que o conteúdo de comentários postado em notícias pelos usuários no ambiente anônimo e não anônimo, possuem características linguísticas particulares. Além disso, mostramos que o usuário tende a ser mais negativo quando encontra-se anônimo.*

1. Introdução

Nos últimos anos, com o crescimento da Internet como meio de comunicação, um grande número de pessoas mudou a maneira de consumir notícias, substituindo os tradicionais jornais e revistas impressos por suas versões virtuais online e/ou *web blogs* [Mitchelstein and Boczkowski 2009, Allan 2006, Tewksbury 2005]. Atualmente, as notícias online têm sido responsáveis pela captação ou atração de uma quantidade significativa de usuários na Internet [Tatar et al. 2014], possibilitando que o consumo de informação nestes ambientes seja cada vez mais social e interativo [Purcell et al. 2010].

Existem várias características de sistemas online em geral que permitem a participação do usuário, entretanto, a interação por meio de comentários e espaços para debates têm sido amplamente oferecida em jornais online [Domingo et al. 2008], e como consequência vem se tornando a forma mais comum de participação dos leitores neste cenário [Emmer et al. 2011]. Os comentários online em sites de notícias são fóruns particularmente importantes na esfera pública uma vez que possuem a capacidade de proporcionar a discussão de questões coletivamente relevantes para a sociedade, possibilitando a formação de opinião pública [Gerhards and Schäfer 2010]. Vários usuários são atraídos pela oportunidade de comentar e expressar seus sentimentos e opiniões sobre uma notícia, o que nos fornece, por meio desses fóruns de discussão, indicações da opinião, pensamentos e sentimentos de leitores [Kothari et al. 2013].

As notícias online, de forma geral, também têm atraído comentários que se propagam em outros meios de comunicação social, como por exemplo, nas redes sociais. Neste cenário, é comum nos depararmos com compartilhamentos e comentários que relacionam ou fazem referência a uma notícia específica [Kothari et al. 2013]. Inclusive alguns sites de jornais online incluem instalações que permitem o compartilhamento e comentários via redes sociais como Facebook ou Twitter [Diakopoulos and Naaman 2011]. Porém, a natureza dos comentários em redes sociais é consideravelmente diferente daqueles postados em sites de notícias [Kothari et al. 2013].

Fóruns online existem desde o início da Internet, no passado, eles eram muitas vezes dedicados a certas questões ou temas sensíveis e sua população de usuários era relativamente pequena e muitas vezes limitada para usuários tecnicamente sofisticados com preocupações específicas. No Facebook, por sua vez, espera-se que cada usuário envie uma mensagem utilizando uma única identidade online (conta), que corresponde à sua identidade no mundo físico. Em contraste, no ambiente de um jornal online, um usuário pode enviar mensagens públicas usando qualquer nome de conta do utilizador, o que pode não deixar clara a sua real identidade podendo proporcionar assim um certo grau de anonimato para a postagem de mensagens por usuários. Trabalhos recentes evidenciam que o anonimato estimula o discurso negativo [Wang et al. 2014], entretanto, pouco se sabe sobre o impacto do anonimato no comportamento do usuário no contexto de comentários em notícias online. E é parte desta lacuna que endereçaremos neste trabalho.

Nosso objetivo é investigar se existe variação no comportamento do usuário considerando o anonimato ou não. Parte do nosso esforço depende do uso de análise de sentimentos como uma forma para capturar o sentimento expresso nos comentários de notícias online. Mais especificamente, trabalhamos no sentido de verificar se existem diferenças linguísticas significativas entre o comentário postado no ambiente onde é conferido ao usuário certo grau de anonimato (jornal Reuters) e no ambiente onde espera-se que este grau de anonimato não exista (Facebook). Além disso, nos propomos a analisar

as reações dos usuários quando eles estão anônimos ou não, para manchetes com diferentes sentimentos associados a elas. Por exemplo, será que a polaridade do sentimento dos comentários dos usuários inseridos nestes ambientes (e.g. anônimo – Reuters, não anônimo – Facebook) refletem a polaridade das manchetes (notícias)¹?

Para respondermos a estas questões, analisamos o conteúdo de cerca de 360 mil comentários postados em notícias do jornal Reuters, onde consideramos o anonimato do usuário, e no Facebook, onde o usuário normalmente possui um identificador associado à sua identidade no mundo físico. Nossos resultados evidenciam que o conteúdo de comentários postados em notícias pelos usuários no ambiente anônimo e não anônimo, possuem características linguísticas particulares. Além disso, mostramos que o usuário tende a ser mais negativo quando encontra-se anônimo. Nossas observações têm aplicações para concepção ou manutenção de sistemas que apoiem a interação do usuário neste contexto, uma vez que fornece indícios do comportamento dos usuários nestes cenários. Em última análise, nossos resultados contribuem para a compreensão dos ambientes de comunicação mediada contemporâneos, bem como a sociedade em que vivemos.

O restante desse artigo está organizado da seguinte forma. A Seção 2 descreve os trabalhos relacionados. Em seguida, na Seção 3, descrevemos nossa metodologia experimental elaborada para este trabalho. Depois, na Seção 4, apresentamos e discutimos os resultados. Por fim, a Seção 5 apresenta conclusões e direções para trabalhos futuros.

2. Trabalhos Relacionados

Antes do surgimento da Internet, o efeito do anonimato sobre o comportamento do usuário já atraía diversos pesquisadores. Em 1969 por exemplo, por meio de um experimento com humanos, P. Zimbardo mostrou que as pessoas se tornam mais agressivas ou violentas em ambientes menos limitados por normas sociais [Zimbardo 1969]. Com o advento da Internet e a crescente popularidade de sistemas de computação, foi permitido ao usuário recriar algumas de suas convenções comportamentais sociais do mundo real no ambiente online. Para essa interação com o conteúdo na Web, o usuário normalmente faz uso de um identificador único que o representa dentro do contexto específico [Jøsang and Pope 2005]. Entretanto, grande parte destes cadastros requerem somente informações arbitrárias de perfil (e.g. email, nome), e este aspecto pode ser utilizado para preservar diversos aspectos do usuário (e.g. privacidade), dentre eles o anonimato, cujo impacto continua sendo analisado em diferentes contextos. Em [Moore et al. 2012], por exemplo, o anonimato em posts de um fórum (ou seja, mensagens em fóruns sem nenhum identificador) foi relacionado como um importante critério para se distinguir/identificar ataques na Internet. Já em [Silva et al. 2016] uma rede social anônima é utilizada como cenário para realização de um estudo sistemático que mede em larga escala os principais alvos de discurso de ódio nos meios de comunicação social online.

Para melhor compreensão da diferença entre o conteúdo postado em mídias anônimas e em outras mídias sociais, o trabalho realizado em [Correa et al. 2015] caracterizou mensagens postadas pelos usuários no *Whisper*² e no *Twitter*³. Com o uso do *Amazon Mechanical Turk*⁴, os autores introduziram o conceito de nível de anonimato para de-

¹É válido ressaltar que a polaridade da notícia é obtida por meio da inferência da polaridade da manchete da mesma, considerando sua importância já atestada e discutida em trabalhos anteriores [Reis et al. 2015].

²<https://whisper.sh/>

³<https://twitter.com/>

⁴<https://www.mturk.com/>

terminado texto e a análise revelou que este aspecto, em alguns casos, não é binário. Formalmente, os pesquisadores realizaram um experimento com seres humanos e propuseram uma métrica que representa o teor do anonimato de determinado texto a partir de uma percepção do próprio usuário. Além disso, diferenças linguísticas entre os conteúdos anônimos e não anônimos foram identificadas, e essas características foram utilizadas para treinamento de um classificador que distingue ambos os tipos.

Já o contexto de notícias online especificamente, tem sido bastante explorado em trabalhos recentes (e.g [Reis et al. 2014b, Reis et al. 2014a, Elgesem et al. 2016, Buhl et al. 2016]). Uma pesquisa com 583 jornalistas dos EUA, que fundamentou-se na formação social das teorias tecnológicas, é apresentada em [Nielsen 2014]. O objetivo do referido trabalho era investigar o impacto dessas tecnologias no comportamento do usuário, incluindo cobertura e interação dos usuários no contexto de notícias. Os resultados mostram que os jornalistas não se opõem ao fato de que as plataformas web forneçam mecanismos que permitam o compartilhamento de comentários pelos leitores, entretanto consideram que o anonimato do usuário impulsiona a interação dos leitores entre si.

O impacto do anonimato em comentários foi também explorado em trabalhos recentes [Santana 2014, Cho and Kwon 2015]. Enquanto [Cho and Kwon 2015] correlaciona o anonimato com a qualidade dos comentários, [Santana 2014] compara comentários postados por usuários anônimos e não anônimos investigando o nível de civilidade em ambos. Os resultados fornecem uma evidência empírica de que o anonimato tem efeito nos comentários dos usuários. Já em [Diakopoulos and Naaman 2011] são caracterizados comentários postados no SacBee.com. O estudo apresenta uma discussão sobre as motivações dos usuários para leitura e escrita de comentários em notícias online, além de diferentes perspectivas para a gestão da qualidade dentro do contexto jornalístico.

Neste trabalho damos um foco diferente na direção de analisar o impacto do anonimato no comportamento do usuário. Considerando o contexto de notícias online, investigamos o impacto do anonimato com foco na polaridade expressa nestes comentários. Além disso, inspirados por algumas análises efetuadas para o contexto de redes sociais, investigamos características linguísticas que diferenciam o comportamento dos usuários nestes dois cenários (e.g. anônimo e não anônimo). Nossos resultados fornecem importantes evidências do comportamento humano sobre essas perspectivas com inúmeras aplicações para projetos futuros, tais como sistemas que atuem neste contexto.

3. Metodologia

Para que fosse possível analisar o impacto do anonimato nos comentários postados pelos usuários, inicialmente, coletamos dados de um ambiente anônimo e outro não anônimo. Em seguida, definimos uma estratégia para caracterização do conteúdo nestes dois cenários e inferimos o sentimento de cada um dos registros. A seguir, descrevemos detalhes relacionados à metodologia experimental adotada.

3.1. Base de Dados

Foram utilizados comentários oriundos do jornal Reuters Online⁵, onde o usuário pode postar mensagens utilizando qualquer nome de conta que na maioria dos casos não deixa clara a sua real identidade (ambiente anônimo), e do Facebook⁶, onde o usuário possui

⁵<http://www.reuters.com>

⁶<http://www.facebook.com>

Tabela 1. Sumário de informações dos datasets

Fonte	Período de Coleta	# Notícias	# Comentários
Reuters	Jan/2013-Set/2014	27.051	166.329
Facebook	Mar/2014-Mar/2014	617	195.157

uma identidade, que normalmente corresponde à sua identidade no mundo físico (ambiente não anônimo).

A agência Reuters, em seus vários idiomas de publicação, ficou conhecida por reportar furos jornalísticos estrangeiros. Atualmente com sede em Nova Iorque, a empresa tem expandido seus produtos de negócio e rede de reportagem global para incluir serviços de mídia, financeiros e econômicos e tem abrangência mundial. Já o Facebook, é a rede social mais utilizada no mundo atualmente⁷. As relações entre seus bilhões de usuários têm atraído a atenção de vários pesquisadores no mundo inteiro, como se pode atestar pelo grande número de pesquisas recentes realizadas no Facebook (e.g. [Ellison et al. 2007, Hayes et al. 2016, Marder et al. 2016])⁸.

3.2. Processo de Coleta

A coleta dos dados do jornal Reuters foi realizada com a execução de um *web crawler* cujo propósito era abranger comentários de notícias em destaque no referido jornal, permitindo uma análise mais robusta considerando o objetivo de pesquisa proposto. Acessamos o *archive* deste jornal online e selecionamos o “*Top News*” de publicações para realização de uma coleta de dados retroativa. Para cada uma das notícias recuperadas, foi realizado um *parser* do conteúdo retornado e recuperadas informações relativas às notícias e aos comentários.

Já a coleta de dados no Facebook, foi realizada por meio da API oficial da rede⁹. Foram selecionadas páginas oficiais de jornais conhecidos internacionalmente¹⁰. A partir daí, acessamos os posts destas páginas que compartilharam notícias. Para cada notícia recuperada, coletamos todos os comentários inseridos pelos usuários, além de informações relativas às próprias notícias. A Tabela 1 sumariza informações relacionadas ao conjunto de dados coletado.

3.3. Caracterização

Com o objetivo de identificarmos características próprias de comentários em ambiente anônimo e não anônimo, damos um passo na caracterização da nossa base de dados nestes cenários. Neste trabalho, primeiramente propomos uma caracterização de um conjunto de comentários coletados com foco em características linguísticas deste conteúdo. Nossa propósito é entendermos o uso de atributos linguísticos (e.g. estrutura e gramática das mensagens) em comentários postados nos dois ambientes analisados. Para isso, utilizamos como estratégia para essa caracterização a ferramenta *Linguistic Inquiry and Word*

⁷De acordo com: <http://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/> - última visita em março de 2016.

⁸Uma rápida pesquisa por artigos publicados a partir de 2012 contendo a palavra “Facebook” retorna mais de 3800 resultados.

⁹<https://developers.facebook.com/>

¹⁰Reuters (<https://www.facebook.com/Reuters>), BBC News (<https://www.facebook.com/bbcnews>), DailyMail (<https://www.facebook.com/DailyMail>) e The NY Times (<https://www.facebook.com/nytimes>)

Tabela 2. Exemplos de comentários rotulados pelo SentiStrength

Força de Polaridade	Comentário
≤ -3	No hope, this people love to kill their own ...very sad
-2	Well, it might be a headache for US sooner or later. These are your puppets sweet US
0	Only what u do for God will last forever
2	The world feels like a better place already.
≥ 3	We will never forget you !!! TOTAL RESPECT !!!

Count (LIWC) [Tausczik and Pennebaker 2010]. O uso dessa ferramenta¹¹ é importante e adequado para nosso estudo pois ele nos permite estimar componentes emocionais, cognitivos, estruturais e gramaticais de um texto fornecido como entrada, baseado na utilização de dicionários contendo palavras e categorias associadas a cada uma.

3.4. Inferência de Sentimento

Depois de definirmos a estratégia para caracterização linguística do conteúdo, propomos um procedimento para inferência do sentimento de cada um dos comentários postados no jornal Reuters e no Facebook. Esforços recentes mostram que não existe um método de análise de sentimentos que seja mais adequado para todos os tipos de dados [Gonçalves et al. 2013]. Assim, ao invés de utilizarmos um método de análise de sentimentos popular, nós brevemente investigamos qual o método seria mais adequado para o nosso conjunto de dados. Para isso selecionamos aleatoriamente do nosso conjunto de dados, sumarizado na Tabela 1, 100 comentários aleatórios com quantidades proporcionais para cada um dos cenários, que foram rotulados manualmente por 3 voluntários do nosso laboratório de pesquisa como positivo, negativo ou neutro. Cada um dos voluntários rotulou as mensagens de forma independente, sem a influência dos demais, permitindo o alcance de um percentual de concordância 92%.

Nós utilizamos este pequeno conjunto de dados rotulado como uma *baseline* para a comparação de desempenho de 8 métodos de análise de sentimento implementados e disponibilizados no sistema *iFeel* [Araújo et al. 2014, Araújo et al. 2016]. O *Sentistrength* [Thelwall et al. 2010] foi o método de análise de sentimentos com o melhor desempenho para o cenário em questão, alcançando uma acurácia de 74% para comentários. Outros métodos apresentaram resultados competitivos com os do *Sentistrength*, entretanto, uma propriedade interessante deste método é que ele mede a força da polaridade do sentimento em textos em uma escala de -5 (muito negativo) a 5 (muito positivo). Acreditamos que isso nos forneça uma perspectiva interessante sobre a intensidade da opinião expressa nos comentários. A Tabela 2 fornece exemplos de comentários, em ambos os cenários, classificados de acordo com a pontuação do *Sentistrength* (optamos por manter a nomenclatura no idioma original, em inglês). Não houve mensagens rotuladas com pontuação 1 e -1 em nosso conjunto de dados.

Uma das limitações deste trabalho está relacionada ao fato de que o conjunto de comentários não representa de forma balanceada o conjunto de notícias. Por exemplo, os comentários coletados no Facebook compreendem mais de um jornal. Entretanto acreditamos que ao analisarmos dados relativos ao mesmo contexto (notícias online), somos capazes de fornecer uma perspectiva geral do impacto do anonimato no comportamento do usuário de forma geral. Para a realização de comparações mais minuciosas, realizamos alguns tratamentos que serão descritos durante a apresentação dos resultados.

¹¹O software está disponível em <http://www.liwc.net/>.

4. Resultados e Discussões

Nesta seção, apresentamos os resultados da análise de dados de comentários de notícias coletados do jornal Reuters e do Facebook.

4.1. Características Linguísticas do Conteúdo

Um dos nossos objetivos neste trabalho era comparar características linguísticas de comentários de notícias postadas pelos usuários no ambiente anônimo (Reuters) e não anônimo (Facebook). Para isso, inicialmente medimos a frequência do número de palavras nos comentários em cada um dos cenários e apresentamos o resultado na Figura 1. Percebemos que para ambos os casos, as curvas obtidas são de cauda longa, entretanto, existem muito mais mensagens curtas (até 100 palavras para Reuters e 50 para Facebook) que longas. Entretanto, as mensagens postadas no Facebook tendem a ser ainda mais curtas do que mensagens postadas diretamente no jornal online.

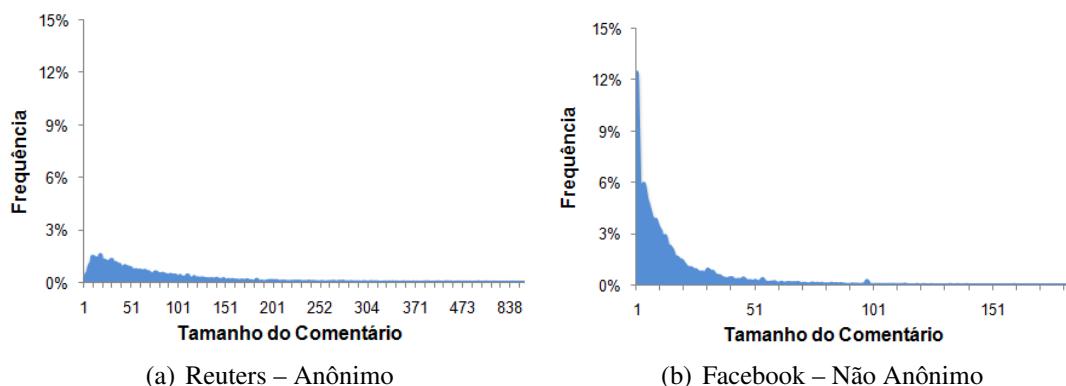


Figura 1. Frequência do número de palavras contidas nos comentários em cada cenário

Em seguida, comparamos e contrastamos amostras de dados destes dois ambientes, conforme detalhamento apresentado em seções anteriores, com a utilização do LIWC [Tausczik and Pennebaker 2010]. Da base de dados apresentada na Tabela 1, nós selecionamos um subconjunto de notícias que possuíam comentários nos dois ambientes: Reuters e Facebook. Para as análises do restante desta subseção, selecionamos uma amostra aleatória de 10 mil comentários inseridos em notícias do jornal Reuters, e 10 mil comentários no Facebook associados a notícias deste subconjunto. As demais análises apresentadas nas subseções seguintes serão efetuadas com todo o conjunto de dados. Em seguida, executamos a ferramenta LIWC e obtivemos o percentual de ocorrência de palavras associadas a cada uma das 85 categorias, para os dois cenários (i.e. anônimo – Reuters, e não anônimo – Facebook).

O LIWC consiste em um conjunto de palavras agrupadas em categorias associadas a quatro classes gerais: *Linguistic Processes* (LP) (e.g., advérbios, pronomes), *Psychological Processes* (PP) (e.g., emoções negativas e positivas), *Personal Concerns* (PC) (e.g., trabalho), e *Spoken Categories* (SC). Além, disso a ferramenta nos fornece taxas de ocorrência de *Punctuation* (PT) (e.g, vírgula, exclamação).

Considerando que o nosso propósito é investigar características intrínsecas ao comentário postado nestes ambientes, focamos nos dicionários ou categorias do LIWC que apresentaram um percentual de variação mais significativo. Por exemplo, para a categoria

Money o percentual de ocorrência para os comentários no Facebook foi de 1,73 enquanto para os dados de comentários no jornal Reuters o percentual foi de 0,78. Neste caso, a razão entre os valores é de 2,21 (1,73/0,78), o que indica que a ocorrência de palavras da categoria *Money* é 2,21 vezes maior em comentários de notícias no Facebook (em relação aos dados de comentários do jornal Reuters). Com base no exemplo apresentado, definimos como significativa a classe cuja razão é menor que 0,5 (ocorrência menor ou igual a metade) ou maior que 2 (ocorrência maior ou igual ao dobro). Acreditamos que isso nos forneça um indício mais preciso de categorias que realmente sejam úteis para a diferenciação dos conteúdos analisados neste trabalho. Os resultados obtidos para as classes significativas do LIWC são apresentados na Tabela 3.

Podemos perceber uma grande expressividade de classes associadas à categoria de preocupações pessoais (*Personal Concerns*) da ferramenta utilizada. O comentário postado em notícias de forma não anônima (Facebook), tem presença significativa de palavras relacionadas ao dicionário de dinheiro do LIWC (*Money*), o que pode ser um indício de que este ambiente impulsiona discussões mais profissionais e/ou pessoais. Já para o cenário anônimo, percebemos o ocorrência considerável de palavras associadas aos dicionários de religião e morte (*Religion, Death*) do LIWC.

Para ilustrarmos esta situação, mostramos por meio da Figura 2 visualizações de árvores de palavras¹² para os termos “*Money*”, “*Religion is*” e “*Death*”, usando o nosso conjunto de dados analisado. As referidas visualizações nos apresentam frases que se ramificam a partir de um termo raiz definido. As palavras mais frequentes são apresentadas em fonte maior. É possível notarmos por meio das frases formadas, que se mantém as considerações efetuadas anteriormente. Para a classe “*Money*” (ambiente não anônimo), estão presentes frases como “*money from companies or interest groups aligned against popular opinion and politically correctness.*”, que apresentam teor mais profissional. Já para Reuters frases como “*religion is a human invention. True, but not so concerning the Word God.*” e discussões sobre “*Death penalty*” (classes *Religion* e *Death* respectivamente), reforçam o comportamento mais negativo do usuário para a discussão de temas mais sensíveis quando o mesmo encontra-se anônimo.

Esta análise corrobora resultados apresentados por [Zimbardo 1969] que destacam que o comportamento dos seres humanos é mais negativo em um ambiente menor limitado por normas sociais. Outro aspecto observado está relacionado ao uso intenso de pontuação em comentários não anônimos (Facebook), o que pode indicar que o usuário esteja um pouco mais preocupado com a formalidade neste tipo de ambiente (em contraste com o ambiente anônimo), onde seu perfil e relacionamentos correspondem, em grande parte dos casos, à sua realidade no mundo real.

Tabela 3. Análise das classes significantes do LIWC

Dicion./Classe LIWC	Anôn./Não Anôn. (%)	Anônimo (%)	Não anônimo(%)	Exemplo de Palavras	Categoria LIWC
<i>Money</i>	2,22	1,73	0,78	Audit, cash, owe	
<i>Religion</i>	0,36	0,43	1,19	Altar, church, mosque	<i>Personal Concerns</i>
<i>Death</i>	0,44	0,44	0,99	Bury, coffin, kill	
<i>AllPct</i>	0,48	13,93	28,98	-	<i>Punctuation</i>

¹²<https://www.jasondavies.com/wordtree/>

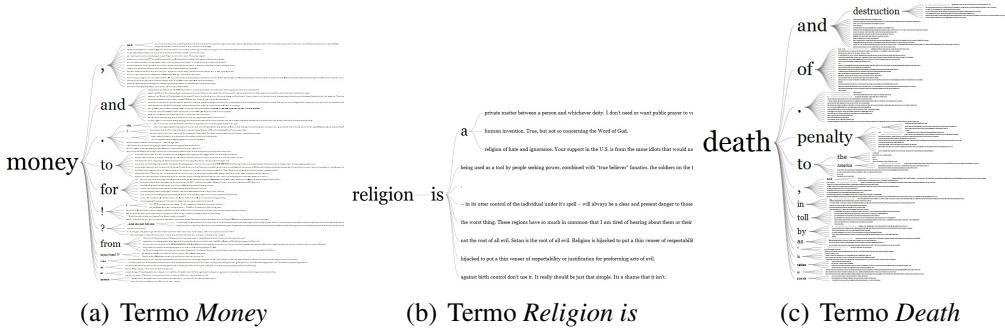


Figura 2. Árvore de palavras

4.2. Polaridade dos Comentários

Depois de analisarmos características linguísticas de comentários postados em notícias no ambiente do jornal Reuters, onde consideramos o anonimato do usuário, e na rede social Facebook, onde consideramos que o identificador do usuário esteja fortemente associado à sua identidade real, partimos para a medição da polaridade de todas as manchetes de notícias e comentários coletados. O nosso objetivo nesta etapa do trabalho era analisar a polaridade da notícia e do comentário em relação ao anonimato do usuário ao inserir postagens neste contexto. Ou seja, como são as reações dos usuários às manchetes com diferentes sentimentos associados a elas, quando estão ou não anônimos? Será que a polaridade do sentimento dos comentários dos usuários inseridos nestes ambientes (e.g. anônimo – Reuters, não anônimo – Facebook) refletem a polaridade das manchetes (notícias)?

Em primeiro lugar, na Figura 3, apresentamos por meio de boxplots (um para cada cenário), a relação entre o número de comentários e manchetes com determinada força de polaridade, conforme classificação do nosso conjunto de dados. A parte inferior das barras corresponde ao primeiro quartil, a parte interior da barra (ponto) representa a mediana, e a parte superior corresponde ao terceiro quartil. Note que as notícias publicadas no Facebook são muito mais comentadas do que as notícias veiculadas no jornal Reuters. Podemos observar ainda que notícias com manchetes negativas são, de forma geral, mais propensas a receberem comentários do que notícias com força de polaridade positiva. Além disso, percebemos que no ambiente anônimo (Reuters) as notícias tendem a receber mais comentários quando o valor de força da polaridade afasta-se do neutro (0), enquanto para o cenário não anônimo (Facebook) as notícias classificadas com polaridade neutra são as que recebem maior número de comentários (mediana de 130 comentários).

Nossos resultados sugerem que no ambiente onde é conferido ao usuário certo grau de anonimato uma manchete tem mais chance de ser mais bem sucedida (atrair mais comentários) se a polaridade do sentimento expressa em seu título é extrema, tanto para o lado positivo quanto para o lado negativo, além de indicar que manchetes com polaridade neutra são geralmente menos atraentes. Entretanto, isso não se mantém para o ambiente onde o usuário não está anônimo, o que pode indicar que o anonimato impacta o comportamento do usuário neste contexto.

A fim de verificarmos de forma mais aprofundada esta hipótese de que o comportamento do usuário varia considerando o fato de que ele esteja anônimo ou não, medimos a polaridade de cada um dos comentários do nosso conjunto de dados e traçamos um

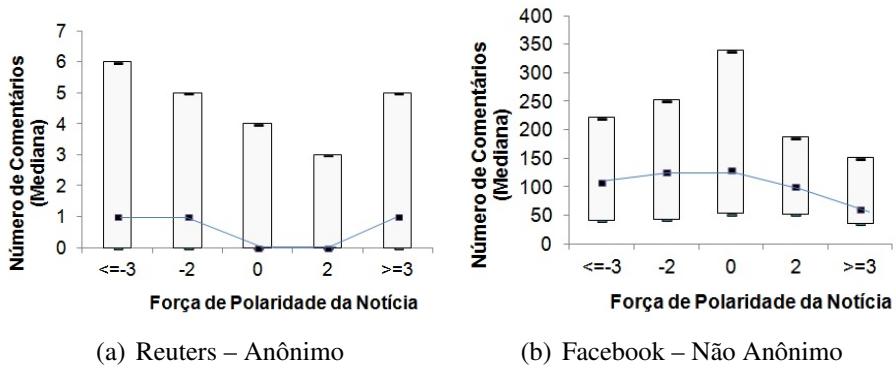


Figura 3. Boxplot para o número de comentários em função do valor da força da polaridade da notícia nos dois cenários

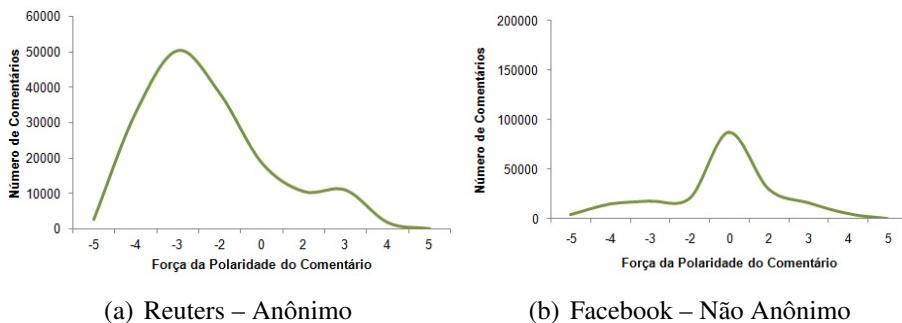


Figura 4. Número de comentários por valor de força de polaridade dos comentários nos dois cenários

histograma de comentários por valor de força da polaridade para cada um dos cenários, conforme representado na Figura 4. Nosso objetivo nesta etapa é entendermos qual a polaridade esperada para comentários inseridos sobre um mesmo grupo de notícias, considerando o usuário nestes dois ambientes (anônimo e não anônimo).

Podemos observar que, para o jornal Reuters em particular, a grande maioria dos comentários possui força de polaridade negativa. Em resumo, 75% dos comentários possuem polaridade negativa, 11% são neutros, e 14% têm força de polaridade positiva. Já para o Facebook, percebemos que essas proporções são distribuídas de forma distinta. Especificamente para este cenário, 74% dos comentários possuem polaridade neutra, 22% são positiva e apenas 2% têm a força de polaridade negativa.

Os resultados identificados, considerando o ambiente onde o usuário encontra-se anônimo (Reuters), corroboram o conhecimento construído por trabalhos realizados anteriormente [D’Costa 2013, Moosa 2014]. Por meio de uma análise mais ampla do ambiente online, os autores afirmam que a seção de comentários é geralmente “casa ou local para hostilidade”. D’Costa (2013) explica que, quando pessoas de diferentes culturas e origens interagem, eles podem ler algo que desafia sua visão de mundo e, em seguida, a lógica e a razão desaparecem. Tudo o que importa é que eles permitam que você saiba o quanto errado você, o tópico, a perspectiva ou o assunto são. E de repente, a seção de comentários tornou-se um discurso retórico. De maneira distinta podemos observar que, para o cenário onde o usuário não está anônimo (Facebook), o número de comentários negativos inseridos pelo usuário é bem inferior, o que evidencia que o comportamento do

usuário é variável de acordo com o ambiente e impactado por fatores como anonimato.

A fim de verificar se o valor de força da polaridade do comentário em ambos os cenários é correlacionado com (ou consequência de) o valor da força de polaridade da manchete, elaboramos as Tabelas 4 e 5 que apresentam matrizes de confusão entre a polaridade da manchete (título) e a polaridade do comentário postado naquela referida notícia, para o Reuters (ambiente anônimo) e o Facebook (ambiente não anônimo), respectivamente. Note que para o jornal Reuters, no qual consideramos o usuário anônimo, os comentários com polaridade negativa são publicados de forma independente da polaridade da notícia (manchete). Ou seja, a percentagem de comentários negativos é significantemente mais elevada para todas as classes de polaridades de notícias. Isso indica que, mesmo que o título esteja resumindo uma história de polaridade positiva, os comentários são de polaridade negativa em mais de 70% dos casos. É curioso notarmos que as notícias neutras são as que recebem o maior percentual de comentários negativos. As notícias de polaridade positiva, por sua vez, receberam mais comentários positivos, ainda que poucos.

Para o Facebook, onde consideramos o não anonimato do usuário, podemos notar uma situação mais uniforme em relação à situação do jornal Reuters em termos de proporção, com grande expressividade de comentários neutros em todas as classes de polaridade das notícias. Tal situação indica que para uma notícia negativa o usuário tende a postar mais comentários negativos (41,12% dos casos), e de forma similar ocorre para as notícias neutras que recebem comentários neutros em mais de 55% dos casos. Já as notícias positivas recebem praticamente a mesma proporção de comentários neutros e positivos (cerca de 39,2%). Finalmente, é importante salientar que, em nosso conjunto de dados, mais de 85% dos usuários postaram em média, em ambos os cenários, apenas um comentário por notícia, o que descarta a possibilidade de que este resultado seja influenciado por usuários que inserem muitos comentários em uma determinada (ou única) notícia.

		Comentários		
		Negativo	Neutro	Positivo
Manchete (Títulos)	Negativa	71,30%	12,71%	16,00%
	Neutra	76,55%	10,46%	12,99%
	Positiva	70,43%	11,37%	18,19%

Tabela 4. Matriz de confusão entre a polaridade da manchete e a polaridade do comentário postado na notícia – Reuters

		Comentários		
		Negativo	Neutro	Positivo
Manchete (Títulos)	Negativa	41,12%	34,53%	24,35%
	Neutra	20,61%	55,07%	24,32%
	Positiva	21,56%	39,25%	39,20%

Tabela 5. Matriz de confusão entre a polaridade da manchete e a polaridade do comentário postado na notícia – Facebook

5. Conclusões e Trabalhos Futuros

Nos dias de hoje as notícias online têm atraído uma quantidade significativa de usuários na Internet. Essas notícias são publicadas de diferentes formas como por exemplo, na

página do próprio jornal ou disseminadas por meio das redes sociais. A interação do usuário nestes ambientes por meio de comentários tem sido cada vez mais presente, e tem possibilitado a discussão sobre diversos assuntos.

Neste trabalho investigamos o impacto do anonimato no comportamento do usuário por meio da análise de mais de 360 mil comentários inseridos em notícias do jornal Reuters e em notícias postadas no Facebook. Endereçamos o ambiente do jornal Reuters com certo grau de anonimato considerando que neste contexto o usuário pode utilizar um identificador qualquer (fraco) que não está associado, necessariamente, à sua identidade no mundo físico. Já no Facebook consideramos que o usuário não está anônimo uma vez que espera-se que seu identificador e círculos de amizade representem sua identidade e relacionamentos reais do mundo *off-line* (identificador forte). Nossos resultados revelam aspectos interessantes sobre o impacto do anonimato no comportamento do usuário considerando o ambiente de notícias online.

Primeiramente, caracterizamos o conteúdo postado nestes dois cenários com objetivo de identificarmos características linguísticas que distinguam o comentário postado pelo usuário considerando o fato de que o mesmo esteja anônimo ou não. Para um subconjunto de notícias online comentadas nos dois ambientes percebemos que o comentário postado pelos usuários anônimos possui características linguísticas particulares. Neste estudo descobrimos que comentários postados no Facebook tendem a ser mais curtos, entretanto notícias publicadas neste ambiente são mais comentadas em relação às notícias veiculadas no próprio jornal. Além disso, mostramos presença significativa de palavras relacionadas à classe de preocupações pessoais do dicionário do LIWC, além do diferente uso de pontuações em contraste com comentários postados no ambiente não anônimo. Essas características poderiam ser utilizadas por exemplo, na construção de algoritmos que identifiquem e façam a separação de um conteúdo anônimo de um não anônimo. Esta estratégia já tem sido explorada de forma mais genérica em trabalhos que treinam classificadores para identificação de conteúdo anônimo [Correa et al. 2015], mas poderia ser também explorada para cenários mais específicos.

De forma geral, nossos resultados geram indicadores de que o anonimato pode ser um fator de impacto no comportamento dos usuários quando analisamos comentários postados pelos mesmos no contexto de notícias online. Tal situação corrobora estudos realizados na psicologia social que mostram que o anonimato influencia fortemente o comportamento do usuário – online e *off-line*. Seres humanos se tornam mais agressivos ou violentos (negativos) em um ambiente menos limitado por normas sociais [Zimbardo 1969], além de que este contexto proporciona desinibição do usuário [Pinsonneault and Heppel 1997, Suler 2004]. Alguns trabalhos argumentam que a falta de prestação de contas no ambiente anônimo estimula o discurso negativo [Wang et al. 2014]. Acreditamos que o não anonimato do usuário talvez o reprema quanto a assumir determinadas escolhas.

Neste trabalho damos apenas um primeiro passo no entendimento deste universo, entretanto entendemos que possam haver outros fatores, além do anonimato, que podem influenciar o comportamento dos usuários nestes ambientes, como por exemplo, o fato de que no Facebook o usuário saiba quem é a sua audiência (normalmente sua redes de amigos), podendo assim evitar falar de coisas ruins para não passar uma imagem negativa para seu meio social. Em contrapartida, no jornal Reuters, a audiência esperada é composta por outros leitores que o usuário não necessariamente conhece, e ainda que estes

outros leitores o conheçam não quer dizer que consigam identificá-lo. Ainda assim, esse trabalho contribui revelando várias observações interessantes sobre o impacto do anonimato no comportamento do usuário neste contexto, fornecendo características que podem ser utilizadas para a concepção de novos sistemas, e finalmente, para compreendermos a sociedade em que vivemos.

Como trabalhos futuros, para fortalecimento de nossos indicadores, pretendemos analisar de forma mais aprofundada a relação do anonimato com outros possíveis fatores de influência do comportamento do usuário, inclusive, levando em consideração outros contextos em que notícias são comentadas de forma anônima ou não. Além disso, pretendemos explorar o impacto do anonimato no comportamento dos usuários de forma mais qualitativa. Entendemos, por exemplo, que uma análise dos traços de personalidade dos usuários talvez possa revelar aspectos interessantes que nos permitam compreender de forma mais aprofundada os fatores de impacto do comportamento do usuário diante desta situação. Por fim, acreditamos que a metodologia que apresentamos neste trabalho pode facilitar ou apoiar diferentes esforços de investigação neste domínio impulsionando assim a produção de conhecimento.

Agradecimentos

Este trabalho foi parcialmente financiado pelo projeto FAPEMIG - PRONEX - MASWeb, Modelos, algoritmos e sistemas para a Web (número do processo APQ - 01400-14), e por projetos individuais do CNPq, CAPES e Fapemig. Finalmente, os autores gostariam de agradecer Leandro Araújo pelas inspiradoras discussões no estágio inicial desse trabalho.

Referências

- Allan, S. (2006). *Online News: Journalism And The Internet: Journalism and the Internet*. McGraw-Hill Education. McGraw-Hill Education.
- Araújo, M., Diniz, J. P., Bastos, L., Soares, E., Júnior, M., Ferreira, M., Ribeiro, F., and Benevenuto, F. (2016). ifeel 2.0: A multilingual benchmarking system for sentence-level sentiment analysis. In *Proceedings of the International AAAI Conference on Web-Blogs and Social Media*, Cologne, Germany.
- Araújo, M., Gonçalves, P., Benevenuto, F., and Cha, M. (2014). ifeel: A system that compares and combines sentiment analysis methods. In *WWW (Companion Volume)*. International World Wide Web Conference (WWW'14).
- Buhl, F., Günther, E., and Quandt, T. (2016). Observing the dynamics of the online news ecosystem: News diffusion processes among german news sites. *Journalism Studies*, pages 1–26.
- Cho, D. and Kwon, K. H. (2015). The impacts of identity verification and disclosure of social cues on flaming in online user comments. *Computers in Human Behavior*, 51:363–372.
- Correa, D., Silva, L. A., Mondal, M., Benevenuto, F., and Gummadi, K. P. (2015). The many shades of anonymity: Characterizing anonymous social media content. In *Ninth International AAAI Conference on Web and Social Media(ICWSM)*.
- D’Costa, K. (2013). Don’t read the comments! (Why do we read the online comments when we know they’ll be bad?). *Scientific American*.

- Diakopoulos, N. and Naaman, M. (2011). Towards quality discourse in online news comments. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, pages 133–142. ACM.
- Domingo, D., Quandt, T., Heinonen, A., Paulussen, S., Singer, J. B., and Vujnovic, M. (2008). Participatory journalism practices in the media and beyond: An international comparative study of initiatives in online newspapers. *Journalism practice*, 2(3):326–342.
- Elgesem, D., Feinerer, I., and Steskal, L. (2016). Bloggers' responses to the snowden affair: Combining automated and manual methods in the analysis of news blogging. *Computer Supported Cooperative Work (CSCW)*, pages 1–25.
- Ellison, N. B., Steinfield, C., and Lampe, C. (2007). The benefits of facebook "friends": social capital and college students' use of online social network sites. *Journal of Computer-Mediated Communication*, 12(4):1143–1168.
- Emmer, M., Vowe, G., and Wolling, J. (2011). *Bürger onlin*. Konstanz: UVK.
- Gerhards, J. and Schäfer, M. S. (2010). Is the internet a better public sphere? comparing old and new media in the us and germany. *New media & society*.
- Gonçalves, P., Araújo, M., Benevenuto, F., and Cha, M. (2013). Comparing and combining sentiment analysis methods. In *Proceedings of the 1st ACM Conference on Online Social Networks (COSN'13)*.
- Hayes, M., van Stolk-Cooke, K., and Muench, F. (2016). Corrigendum to understanding facebook use and the psychological effects of use across generations comput. hum. behav. 49 (2015) 507-511. *Computers in Human Behavior*, 56(C):377.
- Jøsang, A. and Pope, S. (2005). User centric identity management. In *AusCERT Asia Pacific Information Technology Security Conference*, page 77. Citeseer.
- Kothari, A., Darwish, K., Mourad, A., and Taei, A. (2013). Detecting comments on news articles in microblogs. In *Proceedings of the 7th International AAAI Conference on Web-Blogs and Social Media*.
- Marder, B., Houghton, D., Joinson, A., Shankar, A., and Bull, E. (2016). Understanding the psychological process of avoidance-based self-regulation on facebook. *Cyberpsychology, Behavior, and Social Networking*.
- Mitchelstein, E. and Boczkowski, P. J. (2009). Between tradition and change: A review of recent research on online news production. *Journalism*, 10(5):562–586.
- Moore, M. J., Nakano, T., Enomoto, A., and Suda, T. (2012). Anonymity and roles associated with aggressive posts in an online forum. *Computers in Human Behavior*, 28(3):861–867.
- Moosa, T. (2014). Comment sections are poison: handle with care or remove them. *The Guardian*.
- Nielsen, C. E. (2014). Coproduction or cohabitation: Are anonymous online comments on news-paper websites shaping news content? *New Media & Society*, 16(3):470–487.
- Pinsonneault, A. and Heppel, N. (1997). Anonymity in group support systems research: A new conceptualization, measure, and contingency framework. *Journal of Management Information Systems*, pages 89–108.

- Purcell, K., Rainie, L., Mitchell, A., Rosenstiel, T., and Olmstead, K. (2010). Understanding the participatory news consumer, pew internet & american life project. Technical report, Research Report.
- Reis, J., Benevenuto, F., Vaz de Melo, P., Prates, R., Kwak, H., and An, J. (2015). Breaking the news: First impressions matter on online news. In *Proceedings of the 9th International AAAI Conference on Web-Blogs and Social Media*, Oxford, UK.
- Reis, J., Goncalves, P., Miranda, L., Prates, R., and Benevenuto, F. (2014a). A percepção dos usuários sobre a utilização de análise de sentimentos como apoio a seleção de notícias. In *Proceedings of the XIII Brazilian Symposium on Human Factors in Computing Systems*. Sociedade Brasileira de Computação.
- Reis, J., Gonçalves, P., Vaz de Melo, P. O., Prates, R., and Benevenuto, F. (2014b). Magnet news: You choose the polarity of what you read. *Proceedings of the International Conference on Web and Social Media (ICWSM)*.
- Santana, A. D. (2014). Virtuous or vitriolic: The effect of anonymity on civility in online newspaper reader comment boards. *Journalism Practice*, 8(1):18–33.
- Silva, L., Modal, M., Correa, D., Benevenuto, F., and Weber, I. (2016). Analyzing the targets of hate in online social media. In *Proceedings of the International AAAI Conference on Web-Blogs and Social Media*, Cologne, Germany.
- Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & behavior*, 7(3):321–326.
- Tatar, A., Antoniadis, P., De Amorim, M. D., and Fdida, S. (2014). From popularity prediction to ranking online news. *Social Network Analysis and Mining*, 4(1):1–12.
- Tausczik, Y. R. and Pennebaker, J. W. (2010). The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology*, 29(1):24–54.
- Tewksbury, D. (2005). The Seeds of Audience Fragmentation: Specialization in the Use of Online News Sites. *Journal of Broadcasting & Electronic Media*, 49(3):332–348.
- Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., and Kappas, A. (2010). Sentiment strength detection in short informal text. *Journal of the American Society for Information Science and Technology*, 61(12):2544–2558.
- Wang, G., Wang, B., Wang, T., Nika, A., Zheng, H., and Zhao, B. Y. (2014). Whispers in the dark: analysis of an anonymous social network. In *Proceedings of the 2014 Conference on Internet Measurement Conference*, pages 137–150. ACM.
- Zimbardo, P. G. (1969). The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. In *Nebraska symposium on motivation*. University of Nebraska press.