

Neutral or Negative? Sentiment Evaluation in Reviews of Hosting Services

Gustavo Santos*, Marcelo Santos*, Vinícius FS Mota†, Fabricio Benevenuto§, Thiago H Silva*

*DAINF-UTFPR, Curitiba, PR. †DI-UFES, Vitória, ES. §DCC-UFMG, Belo Horizonte, MG.

gustavosantos,marcelos@alunos.utfpr.edu.br,vinicius.mota@inf.ufes.br,fabricio@dcc.ufmg.br,thiagoh@utfpr.edu.br

ABSTRACT

Sharing economy represents activities between people to obtain, provide, or share access to goods and services, coordinated by online services. Airbnb and Couchsurfing are examples of sharing economy, where users offer hosting service in their own houses to the public. In both services, guests can review accommodations. In hosting services of the sharing economy, there is personal contact between those who offer and contract the accommodation, which can affect users' decision to make negative reviews. This is because negative reviews can damage the offered services. To evaluate this issue, we collected reviews from two sharing economy platforms, Airbnb and Couchsurfing, and from one platform of the formal economy that works mostly with hotels, Booking.com, for some cities in the United States and Brazil. We performed a sentiment analysis in the shared texts and found that reviews in the sharing economy tend to be more favorable than those in the formal economy. This can represent a problem in those systems, as an experiment with volunteers performed in this study suggests. In addition, we present some of the main features of these comments, as well as a proposal on how to exploit the results obtained.

CCS CONCEPTS

• **Information systems** → **World Wide Web; Web mining; • Human-centered computing** → *Collaborative and social computing; • Applied computing* → *Law, social and behavioral sciences;*

KEYWORDS

Sentiment Analysis, Reviews, Sharing Economy, Web Mining, Booking, Airbnb, Couchsurfing

ACM Reference format:

Gustavo Santos*, Marcelo Santos*, Vinícius FS Mota†, Fabricio Benevenuto§, Thiago H Silva*. 2018. Neutral or Negative? Sentiment Evaluation in Reviews of Hosting Services. In *Proceedings of Brazilian Symposium on Multimedia and the Web, Salvador-BA, Brazil, October 16–19, 2018 (WebMedia '18)*, 8 pages.

<https://doi.org/10.1145/3243082.3243091>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

WebMedia '18, October 16–19, 2018, Salvador-BA, Brazil

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-5867-5/18/10...\$15.00

<https://doi.org/10.1145/3243082.3243091>

1 INTRODUÇÃO

A economia compartilhada representa atividades colaborativas (pessoa-pessoa) para obter, fornecer, ou compartilhar o acesso a bens e serviços, coordenados por serviços online baseados em uma comunidade de usuários [12]. Plataformas de economia compartilhada vem conquistando uma fatia significativa de mercado em diversos segmentos como o de transporte com, por exemplo, o Uber¹ e o Cabify², e hospedagem com o Airbnb³ e o Couchsurfing⁴.

No mercado de hospedagem, o Airbnb é um serviço que conecta pessoas que possuem um espaço para compartilhar com pessoas que buscam hospedagem e o valor da hospedagem é estipulado pelo anfitrião [28]. Já o Couchsurfing possibilita que pessoas compartilhem seus espaços de maneira semelhante ao Airbnb, porém de forma gratuita, ou seja, os anfitriões não cobram pelo serviço prestado. Esses serviços competem com hotéis, pousadas e até imobiliárias, que sempre foram os dominantes desse mercado, sendo, então, representantes da economia tradicional ou formal.

Tipicamente, os usuários nas plataformas de economia compartilhada são convidados a expressarem suas opiniões a respeito do serviço utilizado. Essas opiniões podem ser capturadas de diversas formas, dentre elas, *ratings* e comentários avaliativos. De fato, essas opiniões, são fundamentais para várias plataformas nesse segmento. Empresas como a Uber, que consideram seus trabalhadores como parceiros temporários em vez de funcionários, exploram as avaliações dos usuários para decidirem se mantêm um determinado parceiro na equipe de colaboradores.

No contexto de hospedagem, as avaliações negativas de usuários que ofertam hospedagem podem impactar na decisão de futuras locações [10]. Além disso, em serviços de hospedagem da economia compartilhada, existe um contato pessoal entre quem oferta e contrata a hospedagem. No Airbnb e no Couchsurfing é comum um hóspede dividir uma acomodação com quem oferta o serviço. Isso possibilita a criação de um vínculo pessoal entre quem oferece e contrata o serviço, o que não tende a acontecer em serviços de hospedagem da economia tradicional.

Com base nessas possibilidades, usuários podem ficar em uma posição desconfortável para realizarem avaliações negativas de serviços ofertados na economia compartilhada, o que pode impactar em uma avaliação adequada do serviço consumido [26]. Sendo assim, uma questão importante de ser investigada é: avaliações realizadas em serviços de hospedagem tendem a ser menos negativas na economia compartilhada? Entender essa questão é crucial não só

¹<http://www.uber.com>.

²<https://www.cabify.com>.

³<https://www.airbnb.com>.

⁴<https://www.couchsurfing.com>.

para ajudar os usuários em suas escolhas, mas também a construção de sistemas de recomendação e ranking de serviços ofertados.

Para avaliar essa questão coletamos comentários avaliativos de duas plataformas da economia compartilhada, Airbnb e Couchsurfing, e um representante da economia formal, o Booking⁵, um popular serviço web de busca por hospedagem em hotéis. Consideramos hospedagens ofertadas em três cidades brasileiras e três cidades dos Estados Unidos. De posse desses comentários, realizamos uma análise de sentimento nos textos compartilhados. Constatamos que os comentários avaliativos na economia compartilhada tendem a ser mais positivos do que os da economia formal. Isso pode afetar a classificação de estabelecimentos feita por usuários, como um experimento realizado neste estudo sugere. Além disso, apresentamos algumas características principais desses comentários. Nosso estudo ainda apresenta uma proposta de como explorar os resultados obtidos para auxiliar na tomada de decisão da escolha de uma hospedagem na economia compartilhada.

O restante do trabalho é organizado da seguinte forma. A Seção 2 apresenta alguns dos principais trabalhos relacionados. A Seção 3 apresenta as plataformas avaliadas neste estudo. A Seção 4 discute o conceito de análise de sentimento e como realizamos essa tarefa. A Seção 5 descreve as bases de dados estudadas. A Seção 6 discute os resultados obtidos com relação à polaridade de sentimento nos sistemas analisados, bem como algumas das principais características relacionadas ao conteúdo e outros fatores. A Seção 7 apresenta os resultados obtidos com um estudo feito com usuários para verificar os resultados obtidos. A Seção 8 discute uma nova forma de avaliação de estabelecimentos de hospedagem da economia compartilhada. Por fim, a Seção 9 conclui o estudo e apresenta apontamentos futuros.

2 TRABALHOS RELACIONADOS

O grande volume de dados disponibilizados por usuários nas mídias sociais habilitou estudos em diversas temáticas, desde a tentativa de previsão de bolsa de valores [6] e terremotos [22] ao entendimento de preferências de gênero [16] e diferenças culturais [24].

Dentre os vários tipos de dados que podem ser compartilhados em mídia social, podemos citar os comentários avaliativos. Nessa direção, Pang et al. [19] verificam que a importância das avaliações e opiniões de outros usuários é cada vez mais essencial na escolha do consumidor, impactando diretamente no faturamento de um produto ou serviço.

Fradkin et al. [10] estudaram comentários avaliativos e as avaliações dos locais no Airbnb e verificaram que, em média, 72% dos usuários escrevem pelo menos um comentário sobre o local que ficaram. Os autores verificaram ainda que 94% das avaliações de estrelas variam entre 4 e 5, esse número é importante para mostrar que, de fato, verificar apenas as estrelas de um local para a tomada de decisão pode não ser a melhor estratégia. Além disso, ressaltam a importância que comentários avaliativos representam para as acomodações.

Fu et al. [11] mineraram comentários avaliativos sobre imóveis, dentre outros dados, visando desenvolver métodos para classificar imóveis de acordo com o valor de mercado. Isso pode fornecer apoio

à tomada de decisões para compradores de imóveis e, portanto, pode desempenhar um papel importante no mercado imobiliário.

Existem várias estratégias que podem ser empregadas ao estudar comentários textuais, dentre eles podemos citar a análise de sentimento. Essa estratégia já avançou consideravelmente nos últimos anos [3, 4, 21] e vem sendo aplicada em diversos contextos [15]. Por exemplo, Alves [2] utilizou análise de sentimentos para a construção de um sistema que faz compra e venda de ações. Para isso a autora utilizou textos provenientes do *Twitter* e o histórico de preços e volume de negociações na BOVESPA.

Oliveira e Bermejo [18] aplicaram a análise de sentimentos em mídias sociais no contexto de gestão social e política. Segundo os autores, o panorama de mídias tradicionais mudou muito nos últimos anos, pois antes predominavam as mídias tradicionais (por exemplo, jornais, revistas e TV) e recentemente estão sendo complementadas ou substituídas por meios interativos de comunicação social [25]. Com isso, a análise de sentimentos das mídias sociais pode ser de grande utilidade sobre o cenário político atual [18].

Leung [13] utiliza análise de sentimentos em comentários avaliativos de produtos. Leung mostra que através da análise de sentimentos, é possível utilizar o resultado positivo ou negativo para apoiar um produto ou dar direcionamento de melhorias.

Existem várias ferramentas para a análise de sentimento, cada uma com características particulares. Nessa direção, Abbasi et al. [1] e Ribeiro et al. [4, 21] fizeram um *benchmarking* em várias dessas ferramentas. Os autores mostram a alta precisão do SentiStrength em textos provenientes de mídias sociais, inclusive comentários.

No melhor do nosso conhecimento, nosso trabalho se diferencia de todos os trabalhos relacionados anteriores, pois o nosso foco é entender se comentários avaliativos em plataformas de hospedagem online da economia compartilhada tendem a ser menos negativos do que os seus concorrentes da economia tradicional. Além disso, realizamos testes com usuários no problema considerado, estudamos algumas das principais características dos comentários analisados, bem como discutimos as implicações para o projeto de novas funcionalidades para plataformas da economia compartilhada.

3 PLATAFORMAS CONSIDERADAS

Consideramos neste estudo três plataformas online de hospedagem. A Booking é a representante tradicional, onde hotéis são comumente negociados. Nessa plataforma, o pagamento em valores monetários é sempre demandado e a negociação dos serviços é tipicamente feita por funcionários, ou seja, sem ter um contato pessoal com o dono do negócio. Além disso, consideramos uma plataforma que não cobra pelas hospedagens: o Couchsurfing. Nessa plataforma a negociação é comumente feita pelo dono do estabelecimento e o contato pessoal entre hóspede e hospedeiro é elevado. Por fim, consideramos ainda um representante intermediário: o Airbnb. No Airbnb o pagamento pela hospedagem é exigido, mas o contato pessoal entre hospedeiro e hóspede tende a ser elevado, muitas das vezes a acomodação é compartilhada com o hospedeiro. A seguir mais detalhes das plataformas.

3.1 Airbnb

O Airbnb, fundado em 2008, é uma plataforma de aluguel de acomodações particulares ao redor do mundo. Presente em 190 países e

⁵<https://www.booking.com>.

em mais de 34 mil cidades, conta atualmente com mais de 2 milhões de acomodações. O Airbnb empodera economicamente milhões de pessoas no mundo inteiro para abrirem e capitalizarem seus espaços, se tornando empreendedores da área de hospedagem. Essa plataforma vem ajudando muitos viajantes a economizar, por se tratar de uma opção, no geral, mais barata que os hotéis. Além disso, outros fatores que podem ajudar a explicar o sucesso dessa plataforma são experiências diferenciadas e mais pessoais.

3.2 CouchSurfing

O CouchSurfing (CS), criado em 2004, é uma plataforma de hospedagem com base na Web, similar ao Airbnb. Já atendeu a mais de 11 milhões de viajantes em mais 150.000 cidades ao redor do mundo. O diferencial desse serviço é que as hospedagens são gratuitas e, tipicamente, o contato pessoal entre o hospedeiro e o hóspede é maior do que no Airbnb e Booking. Enquanto no Airbnb os usuários podem dividir a acomodação com os anfitriões, no CS isso acontece em praticamente todas as vezes.

São disponibilizadas várias funcionalidades como perfis pessoais detalhados, um sistema de verificação de identidade, um sistema de certificação pessoal, bem como um sistema de referências pessoais para aumentar a segurança e a confiança entre membros. Talvez pelo fato de não envolver valores monetários, o perfil do usuário tem um grande valor nessa plataforma. Candidatos a uma hospedagem com um perfil mal avaliado ou duvidoso podem ter maior dificuldade em encontrar uma acomodação do que usuários com um perfil impecável.

3.3 Booking

A Booking, fundada em 1996, é hoje uma das maiores empresas de hospedagens do mundo. Conta com mais de 15.000 funcionários em 198 escritórios em 70 países pelo mundo. Esse serviço de hospedagem visa conectar os viajantes a várias opções de acomodação, incluindo desde pequenas pousadas administradas por famílias a grandes hotéis 5 estrelas. O site também oferece opções de apartamentos e casa particulares, de forma similar ao Airbnb. Apesar do crescimento desse tipo de alternativa, a maioria das acomodações disponíveis, cerca de 68,5%, são quartos de hotéis. Do restante, 8,5% são opções de alugueis temporários ("*Vacation Rental Rooms*", onde várias opções do estilo Airbnb são disponíveis) e 23% são opções de hospedagem únicas ("*Unique Categories of Places to Stay*"), onde é possível encontrar hospedagens inusitadas [7].

Essa empresa representa um dos sites mais populares de hospedagem da economia formal. Todo dia, mais de 1.550.000 diárias são reservadas na plataforma da Booking [8].

4 ANÁLISE DE SENTIMENTOS

A análise de sentimentos tem como objetivo extrair opiniões, sentimentos e emoções em diferentes canais de comunicação, principalmente no formato textual [3, 17], mas também em outros formatos, como em imagens [9]. Particularmente, a identificação de sentimentos em textos vem se tornando uma importante ferramenta para a análise de dados de mídias sociais, habilitando diversos novos serviços [3, 14]. Por exemplo, empresas podem obter opiniões sobre a aceitação de um novo produto.

Os métodos atuais para detecção de sentimentos em sentenças podem ser divididos em dois temas: Baseado em aprendizado de máquina e os métodos léxicos. Métodos baseados em aprendizado de máquina geralmente dependem de base de dados rotulados para treinar os classificadores [20], o que pode ser considerado uma desvantagem devido ao custo de obtenção de dados rotulados. Por outro lado, métodos léxicos utilizam listas, dicionários de palavras associadas e sentimentos específicos. Apesar de não dependerem de dados rotulados para treinamento, a eficiência dos métodos léxicos está diretamente ligada ao vocabulário utilizado, para os diversos contextos existentes. Métodos híbridos também são possíveis.

Existem várias ferramentas disponíveis para a análise de sentimentos, dentre elas a SentiStrength [27]. Essa ferramenta utiliza um dicionário léxico rotulado por humanos que foi aprimorado por aprendizado de máquina. O SentiStrength classifica o sentimento do conteúdo analisado em positivo ou negativo em uma escala de -4 (fortemente negativo) a $+4$ (fortemente positivo), 0 indica um sentimento neutro. Essa ferramenta foi escolhida neste estudo por apresentar melhores resultados na análise de sentimentos para textos representando comentários avaliativos na Web [1, 21].

Neste trabalho, foi utilizado a análise de sentimentos em textos curtos em comentários avaliativos sobre hospedagens. Temos como objetivo descobrir qual a polaridade, negativa, positiva ou neutra, do conteúdo dos comentários avaliativos.

5 BASE DE DADOS

Para a realização deste estudo, nós coletamos de forma automática comentários avaliativos dos sistemas considerados, Airbnb, Couchsurfing e Booking para as cidades de Curitiba, Rio de Janeiro e São Paulo, no Brasil, e as cidades de Boston, Las Vegas e Nova Iorque, nos Estados Unidos. Escolhemos essas cidades por serem populares entre turistas e por oferecerem tipos distintos de atrações, possivelmente atraindo diferentes perfis de turistas. Por exemplo, Curitiba é a terceira cidade mais procurada por estrangeiros para turismo de negócios no Brasil⁶.

Para cada serviço, procuramos por hospedagens nas cidades selecionadas. Em seguida, coletamos todos os comentários feitos pelos usuários em todas as opções de hospedagem disponíveis. Com isso, nossa base é composta por um identificador de estabelecimento, um identificador do usuário que fez um comentário avaliativo, bem como o comentário realizado por esse usuário sobre o estabelecimento. Se a plataforma permite a resposta de comentários, nós também coletamos essas respostas, no entanto elas foram descartadas nas análises realizadas.

A Figura 1 ilustra um comentário avaliativo feito no Booking. Nessa figura o anfitrião respondeu a um comentário avaliativo. Além disso, nesse exemplo, a palavra "ameiii" não existe formalmente na língua portuguesa, no entanto as pessoas são livres para utilizarem esse tipo de construção e isso não é raro de ser encontrado em textos compartilhados em mídias sociais. A ferramenta para a análise de sentimentos escolhida entende e trata esses casos. No exemplo, a palavra *amei* foi considerada intensificada e isso reflete na avaliação de sentimento final do comentário.

⁶<http://www.brasil.gov.br/editoria/turismo/2017/08/curitiba-e-a-3-cidade-mais-procurada-por-estrangeiros-para-turismo-de-negocios>.

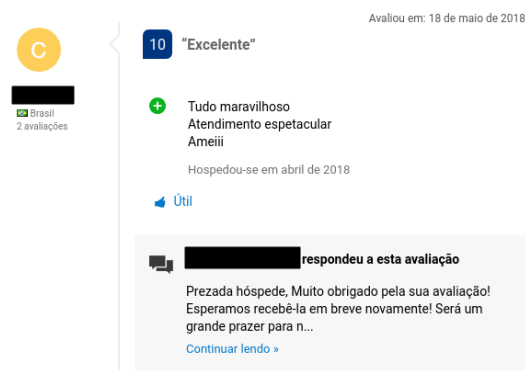


Figura 1: Ilustração de um comentário avaliativo em um serviço de hospedagem online.

Tabela 1: Estatísticas básicas dos dados para cada plataforma considerada.

	# de acomodações	# de comentários
Booking	880	648.030
Airbnb	6.332	115.760
CouchSurfing	963	8.589

Entre outubro de 2016 a março de 2017 foram considerados 648.030 comentários avaliativos do Booking, 115.760 do Airbnb e 8.589 do CouchSurfing. Os comentários podem aparecer em diversas línguas, no entanto concentramos em comentários escritos na língua inglesa e portuguesa, esses valores representam os comentários após essa filtragem.

A Tabela 1 sumariza os dados coletados para cada plataforma. Já a Tabela 2 ilustra exemplos de comentários da nossa base e a força da polaridade do sentimento associada, calculada utilizando a ferramenta *SentiStrength*.

6 SENTIMENTOS EM COMENTÁRIOS

Nesta seção são realizadas as análises relacionadas aos sentimentos em comentários.

6.1 Avaliação agregada

A Figura 2 mostra a distribuição dos sentimentos entre todas as plataformas, Airbnb, Booking e Couchsurfing (CS). Os dados considerados são agregados, ou seja, contêm todas as cidades sem separação. O eixo X representa a polaridade do sentimento identificado para um determinado comentário e o eixo Y representa a porcentagem de comentários atribuídos a cada polaridade.

Como se pode observar, as plataformas da economia compartilhada apresentam mais comentários com polaridade positiva do que na plataforma da economia formal. O resultado apresentado na Figura 2 sugere que o tipo de economia pode afetar o sentimento do consumidor em uma avaliação. Fradkin et al. [10] apresentam alguns motivos que podem explicar o maior positivismo, ou a falta de negatividade em um comentário avaliativo feito no Airbnb, dentre eles: a interação pessoal entre anfitrião e hóspede que muitas vezes

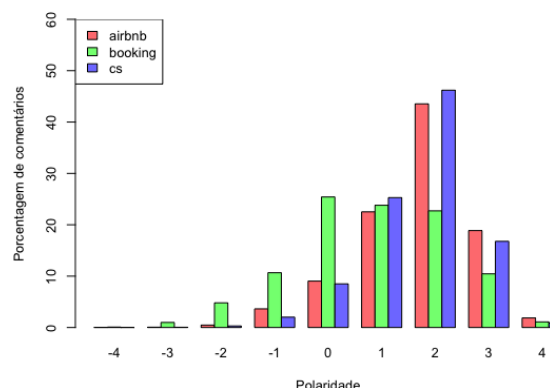


Figura 2: Distribuição dos sentimentos entre todas as plataformas. Nesta figura os dados estão agregados contendo todas as cidades sem separação.

tende a ocorrer; receio de receber uma réplica do anfitrião com relação ao comentário feito; receio de prejudicar o anfitrião com uma avaliação ruim que pode afastar outros possíveis hóspedes.

Como o Couchsurfing se trata de um negócio com relações semelhantes ao Airbnb, esses mesmos motivos poderiam também serem utilizados para explicar os resultados encontrados. Talvez pelo fato de ser gratuito, alguns desses motivos podem ainda serem potencializados, ajudando a explicar a porcentagem menor de comentários negativos com relação ao Airbnb, especialmente o motivo de receio em receber uma réplica do anfitrião. No Couchsurfing o perfil do usuário é muito importante para conseguir uma hospedagem. Usuários podem ser motivados a não realizarem avaliações ruins dos seus hospedeiros para evitarem receber alguma avaliação ruim deles.

Ainda em [10] os autores verificam que no Airbnb quando um consumidor não fica satisfeito com sua experiência ele tende a não escrever um comentário em vez de avaliar negativamente o local. O que também pode ser explicado pelos pontos supracitados.

No Booking como os anfitriões são empresas, tipicamente hotéis e muitas vezes grandes propriedades, essa sensação de proximidade entre anfitrião e consumidor do serviço pode não ocorrer na mesma frequência do que nos outros serviços analisados. Isso pode ajudar a explicar a maior negatividade nos comentários, quando comparado ao Airbnb e ao Couchsurfing.

6.2 Avaliação por cidades

Para verificar se existe alguma relação entre a cidade/país do local avaliado, analisamos cada cidade considerada separadamente. A Figura 3 mostra a distribuição dos sentimentos entre todas as plataformas, considerando todas as cidades estudadas separadamente. O eixo X representa a polaridade do sentimento identificado para um determinado comentário e o eixo Y representa a porcentagem de comentários atribuídos a cada polaridade.

Com o auxílio da Figura 3, é possível notar que o país, ou mesmo a cidade isoladamente, não parece ser um fator de influência relevante

Tabela 2: Exemplos de comentários e a força associada da polaridade do sentimento. Nomes mencionados foram substituídos por W, X, Y e Z para preservar a privacidade dos usuários.

Polaridade de sentimento	Comentário
-4	As acomodações do quarto são terríveis, velhas e desconfortáveis. O sofá de plástico da sala é horrível. Não foi noticiado no Booking que a piscina estava em reforma (escolhemos o hotel por causa da piscina) e a informação que o custo da academia é extra também nos pegou de surpresa. Jamais ficaremos nesse hotel novamente.
-3	Experiência desagradável. O hotel é horrível, não gostamos e não pretendemos voltar mais!
-2	Por erro do hotel não foi providenciada cama extra para meu filho de 4 anos; eu e meu marido tivemos que unir 2 camas “viúvo” para ele dormir conosco, pois quando solicitamos à recepção que o fizesse, nos disseram não ser possível.
-1	A sauna não estava funcionando (se diz em reforma há mais de 5 meses) e a piscina poderia ser maior, além de aquecida.
0	Apartamento como descrito nas fotos. Próximo ao metrô e de um supermercado que possui tudo.
1	O W foi sempre atencioso conosco, a X que nos recepcionou no apartamento também sempre prestativa.
2	O Apartamento é exatamente como está nas fotos. Bem confortável e limpo. Está localizado numa ótima região. A portaria é muito simpática e o Y muito prestativo.
3	Gostei muito do apartamento. Tudo muito bem arrumado, limpo, organizado. Bem localizado, próximo de supermercados, ônibus, pontos turísticos como a Lapa.
4	Foi tudo maravilhoso! Apartamento bem localizado e muito bom! Z respondeu a todas as minhas dúvidas, foi bem prestativo!

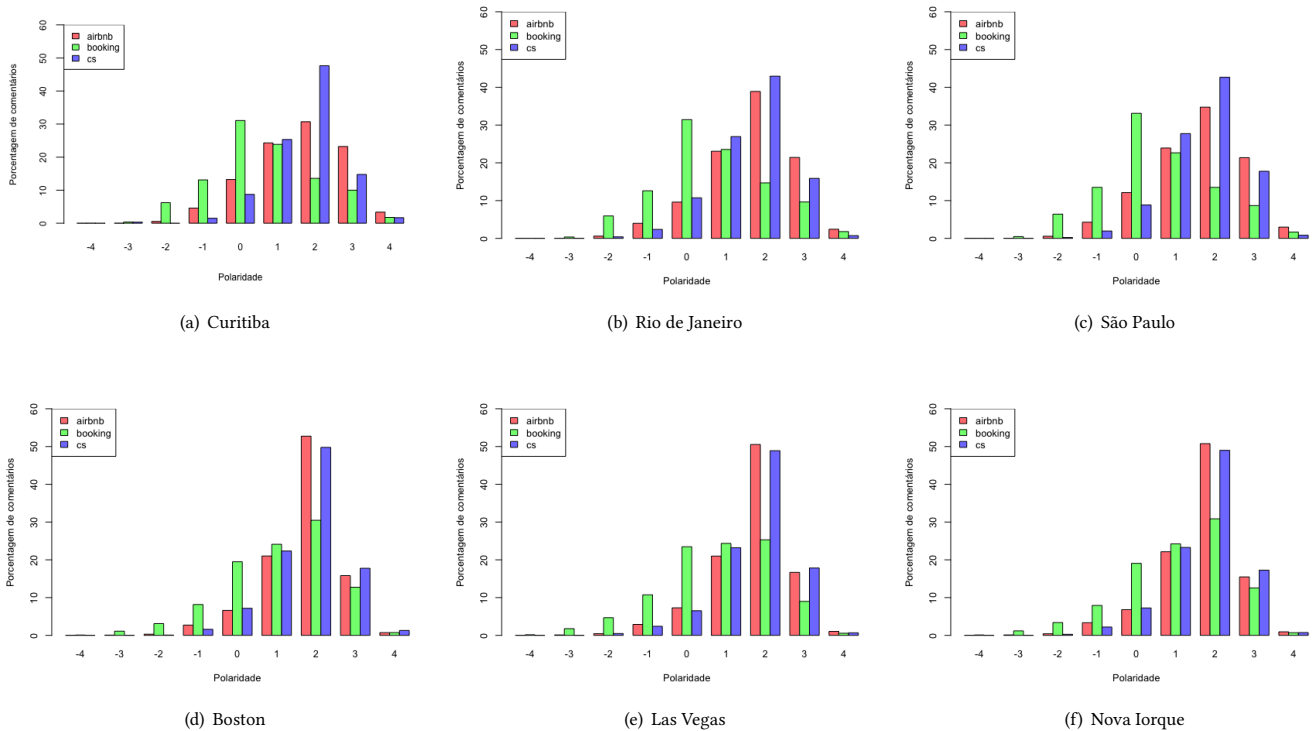


Figura 3: Distribuição dos sentimentos entre todas as plataformas, considerando todas as cidades estudadas separadamente.

com relação à polaridade dos comentários nas plataformas. Ou seja, uma maior positividade em relação às hospedagens da economia compartilhada também é vista quando separamos os resultados por cidades dos dois países analisados. É possível observar ainda que a tendência para essa análise desagregada é muito similar à observada para a análise agregada, incluindo a menor porcentagem de comentários negativos observados no Couchsurfing em relação ao Airbnb.

Os resultados obtidos indicam que na economia compartilhada a presença de comentários negativos é bem reduzida em relação ao que é observado para a economia formal. Isso pode dificultar a percepção dos usuários a respeito da qualidade de um determinado local. Como avaliações negativas tendem a ser mais escassas em comentários avaliativos na economia compartilhada, a opinião neutra pode passar a ter uma importância maior. É como se a escala de polaridade começasse próximo do neutro, ou seja, representando as opiniões mais negativas expressadas pelos usuários. Isso sugere

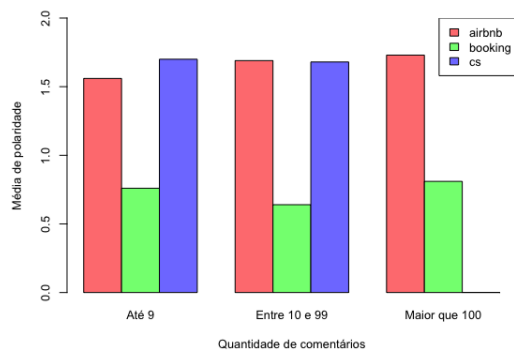


Figura 4: Relação entre número de comentários de um estabelecimento por polaridade de sentimentos.

que avaliações neutras devem ser levadas em conta no momento da decisão de escolha. Essas avaliações talvez podem fazer a diferença para a classificação e tomada de decisão com relação a algum local para hospedagem.

6.3 Sentimento por número de comentários

Uma pergunta que pode surgir neste momento é se a popularidade de um estabelecimento pode influenciar no sentimento expressado pelos usuários nas plataformas analisadas. Para medir a intuição de popularidade consideramos o número de comentários que um determinado estabelecimento recebeu.

A Figura 4 mostra a relação entre número de comentários de um estabelecimento por polaridade de sentimentos. Os tipos de acomodações foram agrupados de acordo com o número de comentários, até 9 comentários (pouco popular), de 10 a 99 comentários (razoavelmente popular), acima de 100 comentários (muito popular). Observe que nenhuma acomodação do CouchSurfing recebeu mais do que 100 comentários.

O resultado, apresentado pela Figura 4, sugere que a popularidade de um estabelecimento parece não ter influência significativa na opinião dos usuários, pois os resultados para diferentes grupos de acomodações de acordo com a sua popularidade não variaram consideravelmente. Isso indica que os resultados observados na Seção 6.2 não são afetados pela popularidade do local.

6.4 Conteúdo dos comentários negativos

Nesta seção investigamos os assuntos mais abordados pelos usuários nos comentários negativos. Concentramos em comentários negativos porque a nossa hipótese é de que, além de mais raros, eles tendem a não ser muito informativos nas plataformas de hospedagem da economia compartilhada. Para realizar essa análise utilizamos a técnica *Latent Dirichlet Allocation* (LDA), uma técnica popular para modelagem de tópicos em conteúdo textual [5].

A modelagem de tópicos é um método para classificação não supervisionado de documentos semelhante ao agrupamento em dados numéricos. No nosso contexto, documentos são comentários avaliativos, contendo palavras diversas. Particularmente, o LDA trata cada documento como uma mistura de tópicos. Isso permite

que os documentos "se sobreponham" uns aos outros em termos de conteúdo, em vez de serem separados em grupos distintos, de forma a espelhar o uso típico da linguagem natural. Por exemplo, em um modelo de dois tópicos, poderíamos dizer que o Comentário 1 é 85% tópico X e 15% tópico Y, enquanto Comentário 2 é 40% tópico X e 60% tópico Y. Além disso, o LDA trata cada tópico como uma mistura de palavras. Assim, em um modelo de dois tópicos sobre acomodações, por exemplo, com um tópico para "quarto" e outro para "café da manhã". As palavras mais comuns no Tópico Quarto podem ser "cama", "lençol" e "barulho". Enquanto o Tópico Café da manhã pode ser composto de palavras como "comida", "café" e "suco". É importante ressaltar que as palavras podem ser compartilhadas entre os tópicos; uma palavra como "ambiente" pode aparecer em ambos os tópicos [5, 23].

Antes da identificação de tópicos, realizamos um pré-processamento nos comentários. Removemos URLs, caracteres especiais, espaços em branco desnecessários, pontuações, números, **stopwords** e transformamos a palavra na sua raiz (*stemming*). Após esse processo, com o auxílio do LDA, identificamos 10 tópicos para os comentários negativos (polaridade -4 a -1) no Airbnb e Booking. A Tabela 3 apresenta as sete palavras que melhor descrevem cada um dos tópicos. Podemos notar que todos os tópicos para o Booking são negativos. Por exemplo, o Tópico 1 é relacionado a reclamações com relação a problemas estruturais de hotéis em geral, já o Tópico 2 é especificamente sobre o banheiro. No entanto, ao analisar os tópicos para o Airbnb, podemos identificar vários tópicos que sugerem sentimentos positivos, todos marcados em negrito e com "*" na tabela. Por exemplo, o Tópico 6 sugere estar relacionado com a acomodação de forma geral, onde comentários indicam que os usuários aprovaram a estadia.

Os resultados para os comentários em inglês seguem um padrão similar ao apresentado para português.

7 EXPERIMENTO COM USUÁRIOS

A análise de comentários sobre os estabelecimentos antes da tomada de decisão é uma tarefa que é comumente realizada por usuários de sistemas de hospedagem online, como os estudados neste trabalho. No entanto, com base nos resultados da Seção 6, surge a pergunta: Com base somente nos comentários, os usuários conseguem ranquear adequadamente as acomodações?

Para avaliar essa questão, realizamos um experimento com usuários. Recrutamos 30 voluntários com um perfil diverso. Incluem nesse grupo representantes de várias faixas etárias, de adolescentes a adultos, possuindo diversos níveis de escolaridade, desde ensino superior incompleto à pós-graduação em andamento.

Entre fevereiro a abril de 2018, esses voluntários foram solicitados a responderem um questionário. Nesse questionário, os voluntários precisavam avaliar um determinado local apenas a partir dos comentários avaliativos feitos por outros usuários. Para isso, primeiramente, escolhemos quatro acomodações listadas no Airbnb. Essas acomodações foram escolhidas com o auxílio da média da polaridade, conforme apresentado anteriormente. Levando em consideração a média de polaridade para cada estabelecimento do Airbnb, a base de locais foi dividida em quartis de acordo com a polaridade. Após esse passo, um local foi escolhido aleatoriamente de cada quartil. Para cada quartil atribuímos os seguintes rótulos,

Tabela 3: Dez tópicos latentes dos comentários negativos observados no Booking e Airbnb.

Tópicos para o Booking									
1	2	3	4	5	6	7	8	9	10
hotel	banheiro	quarto	nao	pouco	ruim	mal	falta	barulho	quarto
nada	limpeza	cheiro	dia	estacionamento	condicionado	atendimento	chuveiro	rua	cama
elevador	box	velho	check	ser	wifi	cafe	toalha	noit	roupa
todo	porta	corredor	reserva	ter	internet	manha	tudo	andar	travesseiro
velho	sujo	mofo	hora	poderia	fraco	restaurant	piscina	outro	casal
reforma	vazamento	pequeno	pessoa	caro	barulhento	equip	banho	lado	solteiro
problema	desejar	apartamento	poi	achei	antigo	problema	frigobar	durant	pequena

Tópicos para o Airbnb									
1	2	3	4	5	6**	7	8**	9**	10**
apart	certeza	apartamento	breve	pena	tudo	bem	voltar	estadia	recomendo
veri	apartamento	excel	espero	todo	recomendo	sujo	voltaremo	recomendo	algun
local	sempr	problema	simpl	vale	excel	cama	rio	novament	valeu
alugar	pena	piscina	ficar	tudo	novo	certo	apartamento	medo	obrigada
obrigado	banheiro	reali	hora	aconcheg	vale	pode	casa	otimo	nao
realment	dia	best	voltarei	pessoa	total	never	banheiro	bom	expectativa
nao	limpeza	ficar	perfeito	otima	funcionando	anunciado	cama	atencioso	tudo

Tabela 4: Resultados do experimento com usuários, contendo a média de nota atribuída pelos voluntários para cada local.

Local	Média das Notas (I.C. 95%)
Representante de Q1	4, 8 (±0, 15)
Representante de Q2	4 (±0, 31)
Representante de Q3	4, 6 (±0, 24)
Representante de Q4	3, 4 (±0, 4)

de acordo com o intervalo de polaridade que eles representam: Q1 (quartil 1); Q2 (quartil 2); Q3 (quartil 3); e Q4 (quartil 4). Note que Q1 representam os melhores valores e Q4 os piores.

Coletamos todos os comentários disponíveis para cada acomodação selecionada. Após um pequeno texto que serve para contextualizar o usuário, ele deveria ler os comentários do local e então classificá-lo, atribuindo uma nota em uma escala entre 0 e 5, sendo 0 ruim e 5 excelente. No formulário a ordem de apresentação dos locais foi embaralhada. Isso foi feito para diminuir algum viés que a ordem dos locais poderia trazer na percepção dos respondentes.

A Tabela 4 sumariza as 30 respostas fornecidas pelos voluntários.

Os resultados do experimento com usuários evidenciam que todos os estabelecimentos foram considerados bons ou muito bons pela maioria dos usuários. Por exemplo, o Local representante da classe Q4, apesar de estar no quartil dos locais com piores média de sentimentos ficou com uma avaliação média/boa na visão dos voluntários.

Isso corrobora com o resultado observado anteriormente, ou seja, na economia compartilhada, que inclui o Airbnb, os comentários tentem a ser mais positivos, o que pode dificultar uma avaliação por humanos. Isso sugere que locais com polaridade média neutra podem não ser boas opções de acomodações, é como se o neutro fosse a polaridade negativa nesse caso. É importante notar ainda que o local representante da classe Q3 ficou com uma média de nota superior ao local representante da classe Q2. Esse experimento reforça a sugestão de que a percepção de qualidade através de comentários no Airbnb pode ser difícil.

8 IMPLICAÇÕES PARA O PROJETO DE NOVAS FUNCIONALIDADES

Após a análise dos resultados apresentados nas Seções 6 e 7, é necessário refletir sobre as suas possíveis implicações. A maior positividade do Airbnb e do Couchsurfing pode ser prejudicial para os consumidores, afinal, anfitriões ruins podem não ficar evidenciados. Más experiências não compartilhadas acabam dificultando a escolha e podem levar o usuário da ferramenta a preferir um local bom devido à ambiguidade dos sentimentos presentes nos comentários.

Em alguns serviços de hospedagem, como o Airbnb, os usuários podem, além de realizar comentários avaliativos sobre a sua experiência com o serviço obtido, dar uma nota em formas de estrelas, que pode variar de 1 (mais baixa) a 5 (mais alta). Porém, como foi observado em um estudo anterior, 94% das notas variam entre 4, 5 e 5 estrelas [10]. Isso faz com que essa classificação de estrelas não seja uma métrica suficiente para que uma análise sobre a qualidade de um local seja feita de maneira adequada.

Com isso, neste trabalho propomos uma nova forma de avaliação de estabelecimentos de hospedagem da economia compartilhada, levando em consideração a polaridade média e a quantidade de comentários de um determinado local. Para isso, sugerimos a seguinte equação:

$$score = \log C + (4 + P)^2, \quad (1)$$

onde P representa a média da polaridade dos comentários do local e C representa a quantidade de comentários do local.

Assim, um valor de $score$ mais alto tende a ser atribuído a locais com sentimentos médios mais positivos nos comentários avaliativos. A equação ainda leva em conta o número de avaliações realizadas, onde quanto mais opiniões é melhor. Ao levar em conta a quantidade de comentários, é possível distinguir se os locais possuem um número de comentários suficiente para uma avaliação.

Para cada local do Airbnb coletado, foi atribuído um $score$ a partir da equação supracitada. A Tabela 5 mostra alguns exemplos de como cada variável impactou no $score$. Nesses exemplos, os valores

Tabela 5: Exemplos de cálculo do score baseado nos comentários e polaridade.

# Comentários	Polaridade média dos sentimentos	score
2	4	64,69
8	3	51,08
3	3	50,10
19	2.8333	49,64
40	2.0938	40,82
26	1.913	38,22
2	1.5	30,94
11	0.2	20,04
13	-1	11,56
3	-1	10,10
3	-2	5,09

do score variaram entre 5,09 (pior score) e 64,69 (melhor score). Acreditamos que esse score, ou alguma variação nesse sentido, pode ser útil para, por exemplo, uma nova forma de ranqueamento de locais. Assim como já existem ranqueamentos do menor para o maior preço, poderia ser criado um ranqueamento de acordo com o score calculado. Isso talvez possa ajudar os usuários a tomarem melhores decisões na escolha de uma hospedagem nas plataformas da economia compartilhada. Para os locais estudados na Seção 7 (mostrados na Tabela 4) o score ranquearia os locais da seguinte maneira: 1° Q1 (score=51,07); 2° Q2 (score=40,66); 3° Q3 (score=34,91); 4° Q4 (score=14,04). Uma avaliação qualitativa com usuários para verificar se essa abordagem pode ajudar a melhorar a experiência do usuário está fora do escopo deste trabalho, no entanto, é importante de ser realizada.

9 CONCLUSÃO

Por meio da avaliação de comentários avaliativos em duas plataformas de hospedagem da economia compartilhada e um da economia formal, encontramos evidências de que os comentários avaliativos tendem a ser mais positivos nas plataformas da economia compartilhada. Isso corrobora com a hipótese de que isso pode acontecer devido ao contato pessoal que ocorre entre quem oferece e contrata o serviço. Esse resultado pode ser prejudicial para os consumidores, pois anfitriões ruins podem não ficar evidenciados. Sabendo disso, nós apresentamos uma possível forma de tratar esse problema. Nós acreditamos que, a partir desses resultados, novas abordagens para solucionar o problema possam ser inspiradas.

Como trabalhos futuros pretendemos estudar padrões entre o sentimento em comentários avaliativos com outros atributos, como sexo e horário. Além disso, realizar uma avaliação qualitativa para investigar se novas formas de ranqueamento de locais explorando os insights obtidos neste trabalho, como a abordagem apresentada, podem trazer uma melhor experiência para os usuários.

AGRADECIMENTOS

Este trabalho foi parcialmente apoiado pelo projeto CNPq-URBCOMP (processo 403260/2016-7), CAPES e Fundação Araucária. Os autores também gostariam de agradecer a todos os voluntários pela valiosa ajuda neste estudo.

REFERÊNCIAS

- [1] Ahmed Abbasi, Ammar Hassan, and Milan Dhar. 2014. Benchmarking Twitter Sentiment Analysis Tools. In *Proc. of LREC*, Vol. 14. 26–31.

- [2] Deborah Silva Alves. 2016. *Uso de técnicas de Computação Social para tomada de decisão de compra e venda de ações no mercado brasileiro de bolsa de valores*. Ph.D. Dissertation. Universidade de Brasília.
- [3] Matheus Araújo, Pollyanna Gonçalves, and Fabricio Benevenuto. 2013. Measuring sentiments in online social networks. In *Proc. of Webmedia*. ACM, Salvador, BA, Brazil, 97–104.
- [4] Matheus Araújo, João P. Diniz, Lucas Bastos, Elias Soares, Manoel Júnior, Miller Ferreira, Filipe Ribeiro, and Fabricio Benevenuto. 2016. iFeel 2.0: A Multilingual Benchmarking System for Sentence-Level Sentiment Analysis. In *Proc. of ICWSM*. Cologne, Germany.
- [5] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research* 3, Jan (2003), 993–1022.
- [6] Johan Bollen, Huina Mao, and Xiaojun Zeng. 2011. Twitter mood predicts the stock market. *Journal of Computational Science* 2, 1 (2011), 1–8.
- [7] Booking.com. 2015. *Need a Bed? Why Booking.com is the Most Diverse and Popular Accommodation Platform in the World*. Booking.com. <https://goo.gl/BSWAQH>.
- [8] Booking.com. 2018. *About Booking.com*. Booking.com. <https://www.booking.com/content/about.en-gb.html>.
- [9] Victor Campos, Brendan Jou, and Xavier Giró i Nieto. 2017. From pixels to sentiment: Fine-tuning CNNs for visual sentiment prediction. *Image and Vision Computing* 65 (2017), 15–22. Multimodal Sentiment Analysis and Mining in the Wild Image and Vision Computing.
- [10] Andrey Fradkin, Elena Grewal, Dave Holtz, and Matthew Pearson. 2015. Bias and reciprocity in online reviews: Evidence from field experiments on airbnb. In *Proc. of ACM SIGecom*. ACM, 641–641.
- [11] Y. Fu, Y. Ge, Y. Zheng, Z. Yao, Y. Liu, H. Xiong, and J. Yuan. 2014. Sparse Real Estate Ranking with Online User Reviews and Offline Moving Behaviors. In *Proc. of KDD*. 120–129.
- [12] Juho Hamari, Mimmi Sjöklint, and Antti Ukkonen. 2016. The sharing economy: Why people participate in collaborative consumption. *Journal of the Association for Information Science and Technology* 67, 9 (2016), 2047–2059.
- [13] Cane WK Leung. 2009. Sentiment analysis of product reviews. In *Encyclopedia of Data Warehousing and Mining, Second Edition*. IGI Global, 1794–1799.
- [14] Bing Liu. 2010. Sentiment Analysis and Subjectivity. *Handbook of natural language processing* 2 (2010), 627–666.
- [15] Bing Liu. 2012. Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies* 5, 1 (2012), 1–167.
- [16] Willi Muller, Thiago H. Silva, Jussara M. Almeida, and Antonio A.F. Loureiro. 2016. Study of Gender Preferences for Locations Around the World Using Social Media Data. In *Proc. of Webmedia*. ACM, Teresina, Piauí, Brazil, 295–302.
- [17] Ramanathan Narayanan, Bing Liu, and Alok Choudhary. 2009. Sentiment analysis of conditional sentences. In *Proc. of EMNLP*. Association for Computational Linguistics, Singapore, Singapore, 180–189.
- [18] Daniel José Silva Oliveira and Paulo Henrique de Souza Bermejo. 2017. Mídias Sociais e Administração Pública: Análise do sentimento social perante a atuação do governo federal brasileiro. *Organizações & Sociedade* 24, 82 (2017).
- [19] Bo Pang, Lillian Lee, et al. 2008. Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval* 2, 1–2 (2008), 1–135.
- [20] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up?: sentiment classification using machine learning techniques. In *Proc. of EMNLP*. Association for Computational Linguistics, Philadelphia, PA, USA, 79–86.
- [21] Filipe N. Ribeiro, Matheus Araújo, Pollyanna Gonçalves, Marcos André Gonçalves, and Fabricio Benevenuto. 2016. SentiBench - a benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Science* 5, 1 (2016), 1–29.
- [22] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. 2010. Earthquake shakes Twitter users: real-time event detection by social sensors. In *Proc. of WWW*. Raleigh, USA, 851–860.
- [23] Julia Silge and David Robinson. 2017. *Text Mining with R: A tidy approach*. "O'Reilly Media, Inc."
- [24] Thiago H. Silva, Pedro O.S. Vaz de Melo, Jussara M. Almeida, Mirco Musolesi, and Antonio A.F. Loureiro. 2017. A large-scale study of cultural differences using urban data about eating and drinking preferences. *Information Systems* 72, Supplement C (2017), 95–116.
- [25] Pawel Sobkowicz, Michael Kaschesky, and Guillaume Bouchard. 2012. Opinion mining in social media: Modeling, simulating, and forecasting political opinions in the web. *Government Information Quarterly* 29, 4 (2012), 470–479.
- [26] Berley A Sparks and Victoria Browning. 2011. The impact of online reviews on hotel booking intentions and perception of trust. *"Tourism Management"* (2011).
- [27] Mike Thelwall. 2013. Heart and soul: Sentiment strength detection in the social web with sentistrength. *Proc. of CyberEmotions* (2013), 1–14.
- [28] Georgios Zervas, Davide Proserpio, and John W Byers. 2014. The rise of the sharing economy: Estimating the impact of Airbnb on the hotel industry. *Journal of Marketing Research* (2014).