Airbnb or Hotel?

A Comparative Study on the Sentiment of Airbnb Guests in Sydney – Text Analysis Based on Big Data

Zhiyong Li, School of Tourism, Sichuan University, China

Honglin Chen, School of Tourism, Sichuan University, China

Xia Huang, School of Tourism, Sichuan University, China

ABSTRACT

Advances in information technology have hugely influenced the tourism industry. Many tourists can generate and share their travel tips through social media, and people consult online reviews before making travel arrangements because they could access these sources of information easily. Either positive or negative reviews could increase consumer awareness of Airbnb. Using the approach of text mining and sentiment analysis, examining whether guests' emotions are positive or negative, this study investigates the attributes that influence Airbnb consumers' experiences compared with their previous hotel experiences by analysing big data of guests' online reviews. Findings reveal that the factors of guests' positive sentiment are the atmosphere, flexibility, special amenities, and humanized service; the factors of guests' negative sentiment are not value for money, have to clean the room before leaving, sharing amenities and space with strangers, disturbed by hosts' noisy recreational activities, and troubled by hosts' requesting good reviews.

KEYWORDS

Airbnb, Australia, Big Data, Online Tourist Reviews, Sentiment Analysis

INTRODUCTION

The sharing economy, also called the collaborative consumption, based on peer-to-peer activity to share goods and services through Internet, moved by growing shared values of the public and increasing technological advancement of Internet platforms, including services like Couchsurfing, Uber as well as Airbnb (Hamari et al., 2016). Airbnb, founded in 2008, is one of the largest peer-to-peer (P2P) accommodation platforms in the sharing economy, has enjoyed significant worldwide growth in more than 81,000 cities and 191 countries. It allows ordinary people to rent residences such as an entire apartment or a private room to tourists, exists to create a world where anyone can belong anywhere and provides healthy travel that is local, authentic, diverse, inclusive and sustainable (Airbnb, 2019). Given its popularity in the tourism industry, researchers have begun undertaking studies on the Airbnb phenomenon, mainly focusing on the issues about Airbnb's impacts on the traditional hospitality sector, attributes comparison between Airbnb and hotels and Airbnb guests' motivations for using the service (Birinci, Berezina, & Cobanoglu, 2018; Guttentag, 2015; Guttentag & Smith, 2017; Neeser,

2015; Tussyadiah, 2015; Yannopoulou, 2013), however, it seldom involves the comparative study of tourists' emotional characteristics between Airbnb and hotels based on the sentiment analysis.

With the rapid development of information technologies, online social media and we-media, a massive new source of data called user-generated content (UGC) have come into being and been shared (Kaplan & Haenlein, 2010). Most of the UGC data at present are the travel photos, travel vlogs and tourist reviews of scenic spots, accommodations, caterings and the overall destinations, through which tourists could know about the tourism information more veritably and objectively. Several recent studies explore the issue of online reviews, or electronic word-of-mouth, focusing mainly on matters such as motivations of, and social dynamics between, users and contributors of review sites (Hennig-Thurau, Gwinner, Walsh, & Gremler, 2004). Yet, previous studies primarily rely on surveys, personal interviews, and other communication-based methods, but not made full use of this UGC (Xiang, Schwartz, Gerdes, & Uysal, 2015). The present sentiment analysis based on travelers' online reviews has mainly restricted the processing of unstructured text especially Airbnb guests' comments needs to be further studied and developed.

As a result, based on the theory of sentiment analysis, this study uses online reviews of world tourists to Australia and selects LIWC dictionary and text analysis instrument to compare the sentiment characteristics of Airbnb users and their previous hotel experience, to investigate the influencing factors of their positive and negative reviews. This study is an attempt to explore the sentiment of tourists based on the tourism big data. For tourists, it is helpful for them to obtain useful information from the huge amount of information before making reservations. For the managers and hosts of the Airbnb platform, it would be useful to perceive tourists' sentiment demands so that they can promote the quality of the accommodation products and services provided during their stay.

REVIEW OF LITERATURE

Comparisons Between Airbnb and Hotel

Various studies have attempted to investigate Airbnb's impacts on traditional hotel accommodation industry. And some researchers have compared Airbnb's attributes with hotels (Belarmino, Whalen, Koh, & Bowen, 2017; Guttentag & Smith, 2017; Zhang, Cui, Cheng, Zhang & Li, 2019). For example, given the small sample capacity, Belarmino et al. (2017) made a preliminary comparison between Airbnb and hotel to see how significant the connection of tourist and hosts is. By using guest reviews, Zhang et al. (2019) found some nuanced attributes between Airbnb and hotels: Airbnb's distinctive attributes compared with hotels include pets, atmosphere, flexibility, value for money and so on. The similar attributes are cleanliness, location, services as well as amenities. These research not only certifies prominent attributes of Airbnb that have been found in preceding studies (e.g. Cheng & Jin, 2019) but also supplement researches with the content that to what extent Airbnb is considered as a good alternative for hotels.

Specifically, a few researchers have made a comparison of users' experiences between Airbnb and Hotel, on the one hand, some studies considered the "social communications" and "authentic experience" as the central point of the Airbnb costumers' experience (Festila & Müller, 2017; Yannopoulou, 2013). For instance, Yannopoulou (2013) studied two brands of Airbnb and Couchsurfing, finally found four themes, including the environment of privacy, humanism, the interaction between guests and hosts which is meaningful as well as authenticity. On the other hand, "cheaper price" (Guttentag & Smith, 2017), "convenient location" (Tussyadiah & Zach, 2016), "home amenities" (Guttentag, 2015) and "cleanliness" (Bridges and Vásquez2016) are of critical importance. However, there is very little literature on the emotional characteristics of Airbnb and hotel's customers from the perspective of sentiment analysis. Besides, there is also a lack of in-depth research on the influencing factors of sentiment comparing Airbnb's guests with traditional hotels.

Big Data and Online Reviews

Tourism has been recognized as the most popular sector in online engagement, traditional approaches are insufficient to academic research because of the volume, velocity, and variety of big data (McAbee et al., 2017), so that the use of big data is rapidly emerging in the tourism research (Fuchs, Höpken, & Lexhagen, 2014). There is a growing literature on social media analysis that combines web crawling, computational linguistics and machine learning to collect and analyze the big data for business purposes such as tracking trending topics as well as identifying opinions about products (Fan & Gordon, 2014). Online consumer review comments, vary significantly in terms of their linguistic characteristics as well as semantic features (Xiang, Du, Ma & Fan, 2017), widely considered big data source that naturally reflects consumer evaluation of products and experiences, often reflecting consumer's sentiments such as happiness, anxiety, sadness and some others (O'Leary, 2011), have been acknowledged and used by researchers to understand a range of research problems in hospitality and tourism (e.g., Abbie-Gayle & Barbara, 2017; Brochado et al., 2017; He et al., 2013; Kozinets, 2010; Schuckert, Liu, & Law, 2015b; Tussyadiah & Zach, 2017; Xiang et al., 2015). Belarmino et al. (2017) compared relevant online reviews to investigate guests' motivations behind staying in P2P accommodations rather than hotels; XXX., however, the tourism and hospitality field has few relevant articles based on big data from a sentiment perspective using online reviews (Fuchs et al., 2014; Xiang et al., 2015).

Sentiment Analysis

The sentiment is an emotional psychological reaction produced in the human perception process, which is situational and unstable, changing with the external environment. After many times of emotional changing, people will form a more rational and organized attitude. From this level, we can understand that tourism commentary expressed by written language is the expression of sentiment.

There are two main orientations in the current sentiment theory, including the categorical approach and dimensional approach. Categorical orientation is enlightened by Darwin's evolutionism, which holds that emotions can be divided into basic emotions (happiness, sadness, anger, disgust, fear and surprise) and complex emotions. Dimensional orientation holds that emotions are not composed of basic emotions and complex emotions, but depend on the influence of three vector dimensions: valence, arousal and power (Ekman and Friesen,1971; Russell, 1980).

Sentiment analysis is a type of text mining that measures people's sentiments on whether they are positive, negative or neutral toward a particular topic through extracting and analyzing people's subjective information but not objective information (Ma, Cheng, & Hsiao, 2018; Zhou, Yang, Bao, & Huang, 2016), for that objective reviews usually do not contain those words that can be classified as positive or negative in a sentiment dictionary. When Compared with traditional questionnaires, using sentiment analysis to evaluate customer satisfaction has become a more objective approach in the marketing and consumer behavior field (Das & Chen, 2007; Mady, 2011). Therefore, it is emerging as an automated process of examining semantic relationships and meaning in reviews.

At present, mature analysis techniques include semantic localization and machine learning. Semantic localization method is to classify words by measuring the distance between positive and negative words, while machine learning method is to mine ideas and analyze emotions through prior supervisory training tests, which includes Naive Bayesian, Support Vector Machine and so on (Pang and Lee, 2008). Sentiment analysis can be explored at the word, sentence, paragraph and document levels. Relatively less research has focused on sentence-level analysis since it is more difficult to exemplify from words compared with paragraphs and documents (Schmunk et al., 2014). Data collection, cleaning, mining and evaluating of the data results are the major steps in most of the sentiment analysis in tourism (Hippner & Rentzmann, 2006; Schmunk et al., 2014). There are three considerations in these steps: 1) the data collection source should be distinguished, researchers have typically studied two types of online content for analysis including professional tourism platforms such as TripAdvisor and social media like micro-blog and Facebook, these online text are usually

limited, for instance, micro-blog in China allows no more than 140 characters in length; 2) in data cleaning, some meaningless words, repetitive words and non-text content should be filtered out initially; 3) in the data mining process, three sentiment analysis methods can be selected, that is machine learning (supervised and unsupervised machine learning), dictionary-based (rely on the use of comprehensive sentiment lexicons and sets of fine-tuned rules created either by humans or by machine) and hybrid approaches.

Although the study of judging consumer satisfaction through the evaluation of tourists' emotions is becoming a new hot spot in recent years, there is still a lack of sentiment analysis based on big data to compare the hotel and Airbnb users, so that here we choose dictionary-based sentiment analysis method on account that it can more flexibly interpret Airbnb users' comments on the Internet, and more suitable for the analysis of tourism phenomena.

RESEARCH METHODS AND DATA

Research Area and Data Collecting

This paper focused on the tourism city of Sydney in Australia for two reasons. One is that Sydney is the largest city in Australia with spectacular natural scenic spots and rich cultural tourism resources located on the east coast. The other is that the Sydney has seen an explosive increase in the number of Airbnb listings (about 23,615) by December 2016, which is ranked the fourth most popular destination for Airbnb in the world (Stuart, 2017), therefore, it has important research value to mine the potential influencing factors behind.

This study used the data available on Inside Airbnb website in Sydney. With the help of a web crawler tool, the researchers collected 181263 online reviews in total. To directly investigate the dissimilarity between Airbnb guests' and the hotel's in preferences and sentiment characteristics, only reviews with the keyword 'hotel' in English were selected, which commendably reflect Airbnb guests' direct comparison with their previous hotel experiences because these relevant reviews were derived from the same respondents. Therefore, 2938 online reviews of 331997 words posted between 24 July 2010 and 4 December 2016 were retained for analysis.

Processing Methods

The first necessary step was to use the TextBlob to process textual data and calculate the emotional scores of the guests' review. TextBlob is a Python library which provides a simple API for diving into common natural language processing (NLP) tasks (Loria, 2018). It has various features such as noun phrase extraction, classification (naive Bayes, decision tree), spelling correction, tokenization and sentiment analysis. TextBlob can break reviews can into words or sentences and the sentiment property returns a named tuple of the form Sentiment(polarity, subjectivity). The polarity score is a float within the range [-1.0, 1.0], if the score of one review was positive, it would be categorized as a positive review and vice versa. The subjectivity ranges between 0.0 and 1.0 where 0.0 is very objective and 1.0 is very subjective. In the tourism field, "positive" and "negative" may respectively mean "satisfied" and "unsatisfied", but further research to link sentiment with satisfaction still under explored.

Second, to analyze tourists' preference and understand what features customers are mainly concerned about, the researchers utilize the ROST CM6 instrument to carry out high-frequency word analysis. ROST Content Mining is a tool of content analysis, developed by professor Shenyang of an information management institute in Wuhan university in China. This system can not only assist various disciplines in scientific research, but also carry out text and content analysis from a large number of data. At present, it can be used to analyze papers, micro-blogs, blogs, forums, web pages, e-mails, chat records and other content files composed of text. The software has many powerful functions such as word classification, word frequency, word type and word frequency, emotional analysis from

simple to complex, citation and citation analysis. The frequency of words is called word frequency, moreover, before using this tool to analyze the frequency of words, we should deal with the data of online reviews in advance, that is delete the pictures, emoticons, wrongly written character, unknown adverbs, singular and plural numbers in the comments for modification and replacement at first, and then filter out meaningless prepositions and articles.

Third, to further analyze tourists' preference and impact factors of the positive and negative review, each of the classified reviews, filtering the neutral reviews, went through the LIWC (Linguistic Inquiry and Word Count) program. The way that the LIWC program works is reading a given text and counting the percentage of words that reflect different emotions, thinking styles, social concerns, and even parts of speech. And LIWC2015 is the most recent evolution in which the heart is default LIWC2015 Dictionary, composed of almost 6,400 words, word stems, and select emoticons (Pennebaker, Booth, Boyd, & Francis, 2015). For each text file, approximately 90 output variables are written as one line of data to an output file. This data record includes the file name and word count, 4 summary language variables (analytical thinking, clout, authenticity, and emotional tone), 3 general descriptor categories (words per sentence, percent of target words captured by the dictionary, and percent of words in the text that are longer than six letters), 21 standard linguistic dimensions (e.g., prepositions, conjunctions, negations, etc.), 41 word categories tapping psychological constructs (e.g., affective processes, social processes, drives, etc.), 6 personal concern categories (e.g., work, home, money, etc.), 5 informal language markers (assents, fillers, swear words, netspeak and nonfluencies), and 12 punctuation categories (colons, question marks, etc.). This study applied LIWC2015 in two steps: (1) selecting the categories the researchers want to analyze and (2) color-coding the words in the dictionary from selected categories.

RESULTS AND DISCUSSIONS

After the process of the above methods, this study extracts 2872 positive, 10 neutral and 56 negative online reviews respectively, of which positive sentiments are the main ones, accounting for 97.75% of the total number of online comments.

Next, we extract the top 150 high-frequency words of the Airbnb guests' online comments, as shown in Table 1. As we can see, the top ten words are "hotel", "stay", "great", "place", "apartment", "Sydney", "location", "room", "host" and "walking", their words frequency are all more than 1000, of which the highest is "hotel" of 3160 in total, hence we could conclude that guests care about "location", "amenities" and "host". First, guests not only emphasize place's convenience such as near the station (ranking No.25 of 778 words frequency), airport (ranking No.33 of 684 words frequency), restaurant (ranking No.54 of 464 words frequency) and shops (ranking No.64 of 373 words frequency) and CBD (ranking No.102 of 220 words frequency), but also prefer Airbnb's superiority that they can easily have a fantastic landscape of beautiful beach (ranking No.42 of 568 words frequency) and cultural harbor (ranking No.97 of 231 words frequency). Second, "hotel", "apartment", "room" are bout Airbnb listings' quality and type because they're the basic element, on the other hand, they put emphasize on houses' specific amenities, for instance, "bed" ranks No.17 of 922 words frequency, "kitchen" ranks No.54 of 464 words frequency, "bathroom" ranks No.57 of 435 words frequency. Third, the high ranking of "stay", "great" and "host" indicate that Airbnb guest regard the great experience of connecting to hosts during their stay as exactly important, the atmosphere is of vital importance because the word "home", "feel", "comfortable", "nice", "lovely", "helpful", "friendly", "wonderful" and "happy" well reflected the feelings or emotions that tourists have when staying in Airbnb listings and communicating with hosts, this fully illustrates the important value of sentiment analysis based on LIWC sentiment lexicon of the third step of processing method.

Theoretically, the ways people use words and expressions can provide abundant information about their faiths, dreads, thinking styles, social connections, and their own characteristics (Dunphy, Stone & Smith, 1965; Gottschalk & Glaser, 1969; Weintraub, 1989). According to all of the categories of

Table 1. Top 150 high frequency vocabulary of tourists' online reviews

No.	words	words frequency	No.	words	words frequency	No.	words	words frequency
1	hotel	3160	51	beautiful	486	101	work	221
2	stay	2828	52	back	485	102	CBD	220
3	great	2162	53	love	471	103	extremely	219
4	place	2063	54	restaurant	464	104	central	216
5	apartment	2002	55	need	463	105	nearby	211
6	Sydney	1654	56	kitchen	441	106	park	210
7	location	1586	57	bathroom	435	107	new	205
8	room	1469	58	accommodation	431	108	balcony	203
9	host	1258	59	enjoy	422	109	price	195
10	walking	1148	60	highly	404	110	flat	194
11	clean	1062	61	wonderful	402	111	modern	193
12	house	1047	62	use	399	112	warm	192
13	home	1030	63	fantastic	395	113	food	187
14	feel	962	64	shops	373	114	living	186
15	recommend	941	65	after	372	115	parking	181
16	everything	929	66	check	370	116	WIFI	173
17	bed	922	67	private	366	117	kind	171
18	really	873	68	trip	362	118	amenity	170
19	well	866	69	short	359	119	fresh	168
20	me	866	70	booking	349	120	water	167
21	staying	859	71	most	343	121	quality	162
22	comfortable	831	72	best	342	122	communication	161
23	nice	805	73	provide	338	123	unit	157
24	time	804	74	looking	335	124	fridge	155
25	station	778	75	street	332	125	early	153
26	perfect	723	76	star	304	126	far	152
27	thanks	719	77	people	302	127	equipment	151
28	they	717	78	friends	300	128	large	150
29	good	720	79	family	297	129	public	149
30	lovely	700	80	bus	294	130	couple	147
31	experience	699	81	excellent	285	131	TV	146
32	her	686	82	breakfast	277	132	road	146
33	airport	684	83	studio	272	133	guest	146
34	when	680	84	weekend	269	134	tidy	144
35	easy	665	85	arrival	266	135	room	141
36	better	640	86	happy	264	136	information	135
37	helpful	637	87	small	262	137	appreciate	133
38	city	620	88	super	257	138	old	131
39	close	611	89	convenient	255	139	awesome	129
40	space	579	90	neighborhood	252	140	describe	128
41	quiet	577	91	local	252	141	facility	128
42	beach	568	92	visit	251	142	extra	127
43	Airbnb	565	93	value	244	143	service	123
44	area	563	94	enough	244	144	floor	122
45	minutes	561	95	distance	242	145	safe	120
46	your	522	96	towel	233	146	darling	118
47	coffee	509	97	harbor	231	147	free	117
48	first	505	98	shower	224	148	noise	115
49	view	493	99	question	224	149	bar	114
50	friendly	489	100	access	223	150	business	113

LIWC2015 dictionary,first of all, the category of "affective processes" includes positive emotion words (such as nice, love, and sweet) and negative emotion (anxiety, sadness and anger) words (such as annoyed, worried and hurt), and through this dictionary's filter, in the column of positive emotion, the biggest number is 35.29, which means that the proportion of positive emotion words in the total online reviews is 35.29%, for example, review No.42 states that "fantastic location and great value when compared to nearby Manly hotels, I would definitely come back again", in this sentence, "fantastic", "great value" and "definitely" are colored red by LIWC sentiment lexicon, that is, these words belong to affective processes words. In review No.40 "very welcoming host, spacious and clean bedroom, comfy bed, good wifi...", and the red color words are "welcoming" and "good". Review No.97 said: "Stuart was the perfect host, really helpful, couldn't have been better. Best visit to bondi ever. Will never use a hotel again." This review's affective processes words are "perfect", "helpful", "better" and "best". And in review No.446 "Thiago's place is in a great location and very modern and nice. The bed is amazingly comfortable - easily as good as any 5-star hotel I've been in. Communication was quick and easy, and Thiago was very helpful and friendly. I would gladly stay here again", the sentiment words are "great", "nice", amazingly", "comfortable", "easy", "helpful" and "friendly". Also, No.951 said: "The room is clean, nice and cozy. 15minutes walk from Central station. Hazel and Gareth are very nice, they even gave me AC on my second day. I ordered food from restaurant downstairs and got 30% discount (works like in-room dining at the hotel). Amenities like towel, bedsheet, tissue, wifi, TV and xbox available in the room. Overall, I am very satisfied with my experience staying here." The colored word in this review is "satisfied", consequently, we can infer that why the guest is pleased is that the quality of house is good, the attitude and service of host is great, the location is very convenient to eating and the amenities are well-considered. In summary, factors affecting the positive emotions of tourists include the atmosphere (e.g. home feeling, authentic feeling and friendly feeling), flexibility (e.g. earlier check-in and late check-out), special amenities (e.g. kitchen) and humanized service (e.g. allowing to bring pets in, offering coffee and local food, providing travel tips and advice, lending bicycle or car, etc.). These factors are unique compared to their previous hotel experience.

Moreover, in the column of negative emotion, the highest proportion of negative emotion words is 12.5%, including 3.57% of "anxiety", 3.23% of "anger" and 12.5% of "sadness". For instance, the No.355 described that "just felt not comfortable to stay in there. Have to wipe the floor by ourselves, it is ridiculous! Did you hear that before when you stay in hotel, the staff ask you wipe the wet floor after shower? Same in here, right?" We can see that the factor of guest's negative emotion is cleaning the place before check-out. In addition, the No.105 review of "the apartment was smaller than expected and there was no air-condition, We found the apartment quite expensive compared to a hotel with breakfast included" and the No.89 review of "this is a budget hotel, not an apartment, the room and furnishings were old and worn. Dirty carpet and very poor wifi" both reflected that the lower-than-expected housing quality will bring to dissatisfaction. Review of No.383 wrote:" accommodation trap, not suitable for any serious guest. Poor check-in process, very noise hosts. Had to leave 12:00 Pm and stay are hotel accommodation, bed was incredibly uncomfortable. This is serious, the positive comments below must be not genuine. Do not stay!" This remark strongly expresses the guest's complaints and distrust, demonstrating that the poor service of host seriously affects customer sentiment. On the whole, factors affecting the negative emotions of tourists include that not value for money, have to clean the room before leaving, sharing amenities and space with strangers, disturbed by hosts' noisy recreational activities such as party, troubled by hosts requesting good reviews and so on.

CONCLUSION AND RECOMMENDATIONS

Either positive or negative reviews could increase consumer awareness of Airbnb. Using the approach of text mining and sentiment analysis, examining whether guests' emotion is positive or negative,

this study identified the attributes that influence Airbnb consumers' experiences compared with their previous hotel experiences by analysing big data of guests' online reviews, findings reveal that the factors of guests' positive sentiment are the atmosphere, flexibility, special amenities and humanized service; the factors of guests' negative sentiment are not value for money, have to clean the room before leaving, sharing amenities and space with strangers, disturbed by hosts' noisy recreational activities and troubled by hosts requesting good reviews.

The study's theoretical contribution lies in demonstrating how big data can be used and visually interpreted in tourism and hospitality studies, it not only makes a useful supplement to the sharing platform Airbnb in the tourism accommodation industry, but also provides guests' direct experience comparisons between Airbnb and hotel through sentiment analysis, reveals the main factors affecting the positive and negative emotions of Airbnb guests, laying the foundation for further theoretical discussions.

The major contribution of this study is practical. It is of great value to use massive UGC such as tourist reviews for that it not only helps tourists make better reservations but also provides important managerial guidance for Airbnb developers, operators and hosts to grasp guests' sentiment preference and strengthen their competitive advantages in tourism marketing and management field.

Notwithstanding this study's contributions, this study is not without limitations. First, the online reviews were restricted to the people who are Internet surfers and capable of using Airbnb app to make tour bookings. Carrying out some in-depth interviews offline could be more comprehensive. Second, this study was limited to a collection of data from only one city, i.e., Sydney. Therefore, different data sets of Airbnb collected in other cities in Australia would help to generalize the findings. Third, researchers could make comparisons to Xiaozhu in China as well as the cultural varieties leading to the differences in reservation preference in the future. Also, a specific lexicon in the tourism and hospitality industry needs to be conducted.

REFERENCES

Abbie-Gayle, J., & Barbara, N. (2017). Airbnb – An exploration of value co-creation experiences in Jamaica. *International Journal of Contemporary Hospitality Management*, *29*(9), 2361–2376. doi:10.1108/IJCHM-08-2016-0482

Airbnb. (2018). *About us*. Retrieved from: https://press.airbnb.com/about-us/

Belarmino, A., Whalen, E., Koh, Y., & Bowen, J. T. (2019). Comparing guests' key attributes of peer-to-peer accommodations and hotels: Mixed-methods approach. *Current Issues in Tourism*, *22*(1), 1–7. doi:10.1080/13683500.2017.1293623

Birinci, H., Berezina, K., & Cobanoglu, C. (2018). Comparing customer perceptions of hotel and peer-to-peer accommodation advantages and disadvantages. *International Journal of Contemporary Hospitality Management*, *30*(2), 1190–1210. doi:10.1108/IJCHM-09-2016-0506

Bridges, J., & Vásquez, C. (2018). If nearly all Airbnb reviews are positive, does that make them meaningless? *Current Issues in Tourism*, *21*(18), 2057–2075. doi:10.1080/13683500.2016.1267113

Brochado, A., Troilo, M., & Aditya, S. (2017). Airbnb customer experience: Evidence of convergence across three countries. *Annals of Tourism Research*, *63*, 210–212. doi:10.1016/j.annals.2017.01.001

Das, S. R., & Chen, M. Y. (2007). Yahoo! For amazon: Sentiment extraction from small talk on the web. *Management Science*, *53*(9), 1375–1388. doi:10.1287/mnsc.1070.0704

Dunphy, D. C., Stone, P. J., & Smith, M. S. (1965). The general inquirer: Further developments in a computer system for content analysis of verbal data in the social sciences. *Behavioral Science*, *10*(4), 468. PMID:5838381

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124–129. doi:10.1037/h0030377 PMID:5542557

Fan, W., & Gordon, M. D. (2014). The power of social media analytics. *Communications of the ACM*, *57*(6), 74–81. doi:10.1145/2602574

Fuchs, M., Höpken, W., & Lexhagen, M. (2014). Big data analytics for knowledge generation in tourism destinations–A case from Sweden. *Journal of Destination Marketing & Management*, *3*(4), 198–209. doi:10.1016/j.jdmm.2014.08.002

Gottschalk, L. A., & Gleser, G. C. (1969). *The measurement of psychological states through the content analysis of verbal behavior*. University of California Press.

Guttentag, D. (2015). Airbnb: Disruptive innovation and the rise of an informal tourism accommodation sector. *Current Issues in Tourism*, *18*(12), 1192–1217. doi:10.1080/13683500.2013.827159

Guttentag, D. A., & Smith, S. L. (2017). Assessing Airbnb as a disruptive innovation relative to hotels: Substitution and comparative performance expectations. *International Journal of Hospitality Management*, *64*, 1–10. doi:10.1016/j.ijhm.2017.02.003

Hamari, J., Sjöklint, M., & Ukkonen, A. (2016). The sharing economy: Why people participate in collaborative consumption. *Journal of the Association for Information Science and Technology*, *67*(9), 2047–2059. doi:10.1002/asi.23552

He, W., Zha, S., & Li, L. (2013). Social media competitive analysis and text mining: A case study in the pizza industry. *International Journal of Information Management*, *33*(3), 464–472. doi:10.1016/j.ijinfomgt.2013.01.001

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the internet? *Journal of Interactive Marketing*, *18*(1), 38–52. doi:10.1002/dir.10073

Hippner, H., & Rentzmann, R. (2006). Text mining. *Informatik-Spektrum*, *29*(4), 287–290. doi:10.1007/s00287-006-0091-y

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, *53*(1), 59–68. doi:10.1016/j.bushor.2009.09.003

Kozinets, R. V. (2010). *Netnography: Doing ethnographic research online*. London: Sage.

Loria, S. (2018). Textblob Documentation. Technical report.

Ma, E., Cheng, M., & Hsiao, A. (2018). Sentiment analysis – a review and agenda for future research in hospitality contexts. *International Journal of Contemporary Hospitality Management*, *30*(11), 3287–3308. doi:10.1108/IJCHM-10-2017-0704

Mady, T. T. (2011). Sentiment toward marketing: Should we care about consumer alienation and readiness to use technology? *Journal of Consumer Behaviour*, *10*(4), 192–204. doi:10.1002/cb.329

McAbee, S. T., Landis, R. S., & Burke, M. I. (2017). Inductive reasoning: The promise of big data. *Human Resource Management Review*, *27*(2), 277–290. doi:10.1016/j.hrmr.2016.08.005

Neeser, D., Peitz, M., & Stuhler, J. (2015). *Does Airbnb hurt hotel business: Evidence from the Nordic countries.* Universidad Carlos III de Madrid.

O'Leary, D. E. (2011). The use of social media in the supply chain: Survey and extensions. *Intelligent Systems in Accounting, Finance & Management*, *18*(2-3), 121–144. doi:10.1002/isaf.327

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends? Information Retrieval*, *2*(1–2), 1–135. doi:10.1561/1500000011

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015*. Austin, TX: University of Texas at Austin.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161-1178.

Schmunk, S., Höpken, W., Fuchs, M., & Lexhagen, M. (2014). Sentiment analysis: Extracting decision-relevant knowledge from UGC. In Information and Communication Technologies in Tourism 2014. Springer.

Schuckert, M., Liu, X., & Law, R. (2015). Hospitality and tourism online reviews: Recent trends and future directions. *Journal of Travel & Tourism Marketing*, *32*(5), 608–621. doi:10.1080/10548408.2014.933154

Stuart, R. (2017). *Sydney Is Airbnb's Australian Boomtown, but Not Everyone Is Celebrating the Website's Success*. Retrieved from https://www.abc.net.au/news/2017-01-31/airbnb-booming-in-sydney-but-it-could-be-pushing-up-rents/8223900

Tussyadiah, I. P., & Zach, F. (2017). Identifying salient attributes of peer-to-peer accommodation experience. *Journal of Travel & Tourism Marketing*, *34*(5), 636–652. doi:10.1080/10548408.2016.1209153

Weintraub, W. (1989). *Verbal behavior in everyday life*. Springer.

Xiang, Z., Schwartz, Z., Gerdes, J. H. Jr, & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *International Journal of Hospitality Management*, *44*, 120–130. doi:10.1016/j.ijhm.2014.10.013

Yannopoulou, N. (2013). User-generated brands and social media: Couchsurfing and Airbnb. *Contemporary Management Research*, *9*(1), 85–90. doi:10.7903/cmr.11116

Zhang, G., Cui, R., Cheng, M., Zhang, Q., & Li, Z. (2019). A comparison of key attributes between peer-to-peer accommodations and hotels using online reviews. *Current Issues in Tourism*, 1–8.

Zhou, J., Yang, Y., Bao, X., & Huang, B. (2016). Combining user-based and global lexicon features for sentiment analysis in twitter. *International Joint Conference on Neural Networks*.

Zhiyong Li is a professor as well as the Dean of Tourism School at Sichuan University. His research interests center on tourism marketing, hotel management, and tourist behavior.

Honglin Chen is a postgraduate student in Tourism School at Sichuan University. Her research interests include tourist behavior and sentiment analysis.

Xia Huang is a postgraduate student in Tourism School at Sichuan University. Her research interest is sharing economy.