A data driven approach to guest experiences and satisfaction in sharing

Feifei Xu, Liqing La, Feng Zhen, Tashi Lobsang & Chao Huang

To cite this article: Feifei Xu, Liqing La, Feng Zhen, Tashi Lobsang & Chao Huang (2019) A datadriven approach to guest experiences and satisfaction in sharing, Journal of Travel & Tourism

Marketing, 36:4, 484-496, DOI: 10.1080/10548408.2019.1570420

Accommodation sharing provides unique experiences for tourists. This study collects 33,892 guest reviews of Airbnb. A text analytic approach to deconstruct the nature of sharing experience is explored. The data driven approach extracts the following dimensions contributing to the core experiences of sharing, namely, housing, location, host-guest interaction, facility, overall sensation and others. A regression analysis further demonstrates a strong association between attributes of sharing experiences and guest satisfaction. Factors such as room facilities and home experience convey different semantic space and significantly affect guest satisfaction. Indeed, room facilities and home experience have been extracted as hygiene factor and motivator, confirming the two factor theory.

Keywords: sharing economy; experiences; satisfaction; text analytics; Airbnb

# Introduction

Managing customer experiences and satisfaction is an important part of service management and marketing, which leads to consumers’ repurchasing behaviour (Liang et al., 2015), word of mouth and sales (Fornell, 1992; Halstead and Page,1992; Gundersen et al., 1996; Su, 2004). The emergency of sharing economy platforms in tourism and hospitality industry has revolutionized tourists’ accommodation options in the last few years. As a newly emerged type of accommodation, sharing accommodation has attracted customers with benefits such as low price, different decoration styles, equipped facilities, local experience and home atmosphere but it also faces such challenges as trust and safety issues(Yang et al. 2019). As a pioneer in the sharing industry, Airbnb, has quickly established its leading role as an online platform of accommodation sharing (Guttentag,2013; Kathan et al., 2016). It has become a powerful competitor to the tradition hotels (Zervas et al., 2015; Geiger et al., 2017). Therefore, identifying key factors of guest experiences and satisfaction can help accommodation providers to recognize customer needs and to improve service quality. It may also help hotels to recognize their competitors' competitive advantages.

Although many studies have been conducted on the traditional hotel experiences and satisfaction (Rhee and Yang, 2015a; Callan and Bowman, 2000; Shanka and Taylor, 2004; Qu, Ryan and Chu, 2000; Rhee and Yang, 2015b; Schmitt, 2003; Knutson, 1988), there has been an increasing research interest in the field of accommodation sharing. Previous studies investigating the issues of sharing accommodation have mainly focused on consumers’ motivation and attitude towards it, experiences, satisfaction and the trust among hosts and guest (Gibbs et al., 2017; Guttentag et al., 2016; Möhlmann, 2015; Heo, 2016; Tussyadiah & Pesonen, 2015; Tussyadiah, 2016; Tussyadiah & Zach, 2017). As for consumer experiences and satisfaction, researchers pay attention to either the experience or the antecedents of consumer satisfaction (Tussyadiah & Zach, 2017; Möhlmann, 2015), but fail to focus on the relationship between sharing experiences and customers’ satisfaction. Therefore, there is a research gap to investigate how the factors constructing consumers’ experience contribute to their satisfaction.

Accommodation sharing creates a network and encourages social interaction among people (Geiger et al.,2017), the service provided and consumed can be unique and different from traditional hotel experiences (Tussyadiah, 2016; Poon and Huang 2017). Researchers suggest the unique experiences and cost saving have been considered to be important reasons for tourists to choose sharing accommodation (Guttentag, 2013; Oskam 2016; Tussyadiah and Sigala 2018). However, current research on consumers’ sharing experiences have mainly used a traditional way of collecting data, such as surveys or interviews (Möhlmann,2015; Tussyadiah, 2016; Guttentag et al., 2018). Compared with the data obtained from traditional questionnaire survey, online reviews are more realistic and reflects real time, and can generate new insights (Xiang et al., 2015), therefore, can be used to help the managers and homeowners to improve the service quality. Lipsman (2007) recognizes that more than 87% consumers rely on user generated content when making decisions for booking hotels. It is particularly important for sharing economy. As the nature of sharing is peer to peer, the comments generated by previous users will be used for potential users to help decision making of booking with the hosts. Nowadays, the rapid development of business intelligence and big data analytics made it possible for managers to make timely decisions in respond to market intelligence and competitive analysis (Chen, Chiang, and Storey, 2012; Tussyadiah and Zac, 2017). Therefore, it is critical to understand factors influencing the core elements of sharing accommodation from the perspective of user generated comments. There are emerging research using online reviews in understanding accommodation sharing experiences (Tussyadiah and Zac 2017; Wiles and Crawford, 2017). Yet, these studies have only identified key themes of sharing experiences, without identifying the connection of guest experiences and guest satisfaction. According to Herzberg (1966), the antecedents of guest satisfaction can be divided into two categories, including “hygiene factors” and “motivators”, which play different roles for shaping consumer experience and satisfaction. It has been widely verified and approved in the research of hotel consumer satisfaction (Cadotte and Turgeon, 2016; Chan & Baum, 2007; Kim et al., 2016; Xiang et al., 2017). However, there are few researches extending the two factor theory into the investigation of guest satisfaction of accommodation sharing. Therefore, the unique contribution of this study lies in the initiating application of Herzberg’ two factor theory in the field of accommodation sharing.

In this research, we chose London, one of the most active users of airbnb.com as a case and 33892 pieces of reviews were collected as sample by means of web crawl tool. Text analytic approach is used to analyze the influencing factors of accommodation sharing experiences, and in order to further explore which are the critical factors that influence tourists’ satisfaction on sharing, a regression analysis was also used to explore a data driven approach to customer satisfaction. The research objectives are therefore established:

1. to identify the core elements of sharing experiences in a tourism accommodation context;
2. to investigate the relationship between the experiential factors and consumer satisfaction in the context of sharing accommodation .

The paper begins with an introduction before considering relevant literature on the attributes of experiences and satisfaction; user-generated comments and customer experiences. Thereafter, specific text analytics approach will be discussed in the methodology part. Subsequently, the main results of the finding will be presented with the aim of demonstrating a data driven approach to experiences and satisfaction of accommodation sharing. The final part of the paper includes a discussion and conclusion, and recommendation for future research.

# Literature Review

## The attributes of experiences and satisfaction

Experiences play an important role in creating service competitiveness. Pine and Gilmore (2011) argue that experience is the fourth new engine to bring economic benefits to businesses after commodities, goods and services. Managing customer experiences can turn customers into supporters; therefore, can contribute to long-term competitiveness and profits, and lead to good word of mouth and repurchase (Smith and Wheeler, 2002).

Studies demonstrated that different hotel attributes influence satisfaction and post-purchase behaviour to a varying degree (Tussyadiah and Zach, 2017). Herzberg’s two factor theory has been used to explain the different hotel attributes that contribute to satisfaction (Xiang et al., 2015). That is, first, meeting the needs of customers cannot motivate enthusiasm, only those which are called “motivators” can lead to high satisfaction; second, the absent of hygiene factors will cause strong dissatisfaction, but the possess of them does not necessarily lead to satisfaction (Herzberg, 1966). The factors that influence guest experiences and satisfaction are categorized into different aspects by researchers, including the hotel's attractiveness (i.e., facilities, location, price, and environment) and the service interaction with the guests (Callan and Bowman, 2000; Shanka and Taylor, 2004; Rhee and Yang, 2015).

Researchers suggest different factors that may affect hotel customer satisfaction, however, literature on accommodation sharing is still emerging. Sharing is to offer platforms that bring individuals together to distribute their excess capacity of property resources among each other (Tussyadiah and Zach, 2017). Guests could have opportunities to obtain authentic experience and the sense of belonging (Möhlmann, 2015; Liang et al., 2015). This experience is completely different from the traditional standardized hotel staying, so it is valuable to identifying what factors contribute to the experiences and satisfaction in sharing accommodation.

As Airbnb is the largest accommodation sharing platform at present, several researchers use Airbnb as case studies. Researchers suggest tourists use Airbnb not only because of its economic benefits (Gibbs et al.,2017), but also of its authentic experiences (Guttentag,2013). Tourists have the opportunity to participate in the daily life of the locals and understand the local culture. Through questionnaire, Möhlmann (2015) points out that the determinants serving users’ self-benefit (i.e., utility, trust, cost savings, familiarity, service quality and community belonging) affect the satisfaction of accommodation sharing and the likelihood of choosing a sharing option. Heo (2016) argues Service Advantage Logic (SDL) proposed by Vargo and Lusch (2004) may apply to accommodation sharing to explain its popularity. SDL argues that the current economic paradigm has changed from simply enterprises creating value to value co-creating with consumers, emphasizing the importance of social interaction as the driving force for value co-creation. Some researchers believe that consumers like sharing economy might be related to their consumption values, showing a shift from “you are what you own” to “you are what you share” (Rifkin, 2014; Kathan, Matzler and Veider,2016). The perspectives of various research are to construct a theoretical framework of customer satisfaction under sharing accommodation based on the existing theory, for example, the theory of ‘The Logic of Collective Action’ (Olson, 1965; Möhlmann, 2015).

Previous studies suggest social exchange theory and self-determination theory can be used to explain sharing activities (Möhlmann, 2015; Geiger et al. 2017). Indeed, guests’ intentions of sharing are determined by the satisfaction and benefits gained by themselves. Based on self-determination theory, the attributes of guest satisfaction in sharing accommodation can be divided into both internal elements (such as the pursuit of sustainability, entertainment and curiosity) and external elements (cost saving, social needs like establishing new relationships, home experience, amenities and location) (Tussyadiah, 2016). However, which factors contribute to satisfaction and dissatisfaction is not clear. Herzberg’s two factor theory has been successfully applied in the studies of hotel guest experiences to identify the key elements of experiences (Xiang et al., 2015), therefore, it would be useful to identify key elements and its association with satisfaction in the sharing context.

## User-generated comments and customer experience research

Recently, user-generated comments have become a popular way of researching hotel customer experiences. User-generated comments are defined as personally generated reviews published on corporate or third-party websites (Mudambi and Schuff, 2010). The comment itself is a rich and spontaneous expression and evaluation on the travel experiences and the authentic and informative information can provide a more comprehensive view of the unique tourist experiences; and the tourists' online reviews are cost-effective and easy to obtain. As a result, it is more economical and effective to collect feedback of tourists and control quality. Tourists’ ratings and reviews have become an important source of research in tourism experiences (Liu et al., 2017; Rhee and Yang, 2015). The processing of text data has developed from manual coding to machine learning, and the volume of data has run to “big data”. Pan, Maclaurin and Crotts (2007) conduct a text analysis of the 40 travel blogs in Charleston, USA and concludes that the essence of reviews is a reflection of the travel experience. Zhou et al. (2014) identify 23 key attributes from the internet reviews that underpin customer satisfaction including apartment facilities, location, service, staff and so on, with the traditional text processing methods of manual coding. Xiang et al. (2015) proves the usefulness of online reviews in the study of hotel guests’ experience and satisfaction through text mining of customer reviews.

However, most research on sharing accommodation experiences have been done via a traditional research method, where questionnaire or interviews (Guttentag, 2013; Tussyadiah and Pesonen, 2015; Gibbs et al.,2017; Guttentag et al., 2018) were used with only a few exemptions, such as Tussyadiah and Zach (2017) who contracted key contents and themes from online reviews. After comparing reviews on sharing accommodation with traditional hotels, they suggest clean rooms and comfortable beds are important to both type of customers, but sharing accommodation users pay more attention to social motives, such as the hosts. And through a qualitative content analysis and interpretative phenomenological analysis of 910 reviews on a network hospitality website in the USA, Wiles and Crawford (2017) identify themes such as verbal communication and the sense of feeling at home are important parts of sharing experiences. However, neither research has explored which factors might contribute to the high level of guest satisfaction or dissatisfaction.

# Methodology

##  Case selection

This paper takes London as the case, of which the main reasons are as follows:

Airbnb is the world's largest tourism accommodation sharing platform where users can publish their rooms online, search for room information and make online bookings. In recent years, Airbnb has developed rapidly. Since its establishment in 2008, Airbnb has become the world's third largest start-up company with more than 3 million listings in the world and covers more than 65,000 cities in 191 countries (About us-Airbnb). This figure has exceeded the Hilton, Starwood and any other global chain hotel group and its market valuation has reached 24 billion dollars (Wang & Nicolau, 2017).

In addition, the system of Airbnb only allows those who have stayed and paid the fees to write reviews, avoiding the situation that the host asks people to write comments for themselves. It emphasizes the user's authentic comments. UGC (User Generated Comments) has the characteristics of authenticity, initiative, and thus has a high research value, from which we can learn the user's real feelings (Guo et al., 2017).

London, one of the world's leading tourist destinations, is the main gateway to the UK and other parts of Europe, one of the active areas of Airbnb users. And by June 2015, London had more than 14,000 Airbnb rooms (Quattrone et al., 2016). Tourism is a key factor in the future of the London’s economy, accounting for 8% of London's GDP, employing about 276,000 people (PricewaterhouseCoopers, 2016).

##  Data collection

This study utilizes a web crawler to collect the guest reviews to form a database, gathering 33,892 pieces of reviews associated with rating scores of 11,023 listings. As a commonly used network data, crawler is often used to study network reviews to explore human behaviour and customer satisfaction (Bai et al.,2014; Xiang et al. 2015), gathering 33,892 pieces of reviews. Data collection spans from October to November 2016 and the user comments were published between January 2014 and September 2016. Data was analysed by KH Coder Software (Higuchi, 2015) following pre-processing, lexical analysis, and visualization process.

## Data processing and text mining

This study analyses the nature of experience of sharing accommodation by extracting important attributes from the tourists reviews with the approach of text mining.

The first step of text analysis is to use Stanford POS Tagger to pre-process data (Tussyadiah and Zach, 2017), a Java implementation of part of speech tagging method (Toutanova, Klein, Manning, and Singer, 2003). The pre-processing includes sentence segmentation, mark (i.e. a text segmentation marker), remove stop words (i.e. meaningless words like "a", "an", "the"), POS tagging (the classification of words having similar grammatical properties, such as nouns, verbs, adjectives, etc.), and lemmatization (putting together words with different morphological changes). The pre-processed Airbnb data consists of 2,562,486 markers and 59,656 word types. After removing stop words, 1,092,008 markers and 58,019 word types left for analysis.

In order to identify the main words and subjects that represent the tourists' experience, the pre-processed data are analysed by lexical analysis, association statistics and data visualization. To identify the important words used in the review, the high-frequency words are extracted based on the term frequency (TF). According to the word frequency of each part of speech, the term frequency list is generated (see Table 1). Because of the word limit in the paper, only 100 high-frequency words are listed in Table 1.

INSERT TABLE 1 NEAR HERE.

Next, multidimensional scaling analysis was used to cluster the words discussed in the reviews, which formed the subjects of sharing accommodation experiences. As suggested by Tussyadiah and Zach (2017), the Kruskal algorithm, with Jaccard Distance can be used. The Jaccard coefficient (Romesburg, 1984) is a statistical measure of the similarity between finite sample sets (i.e., words), which is defined as the intersection of the sample sets divided by the union. The Jaccard coefficient of a word for A and B is:

The above text analysis, including word frequency and multi-dimension analysis, reflects the information of the review data as fully as possible and shows the semantic structure of tourists’ experience expressed of sharing accommodation intuitively and clearly.

Finally, to understand which attributes of experiences contribute to the key elements of satisfaction, this research screens the key attributes related to guest satisfaction through the following operations, and then examine the relationships of these attributes and guest satisfaction.

To get the attributes related to antecedents of guest satisfaction, the high-frequency words were then cleaned artificially. The process was conducted by 1 researcher, and checked by two others. In the first 200 high-frequency words, some irrelevant verbs and adjectives are deleted (i.e. “love”, “like”, “beautiful”, “wonderful”, “fantastic”). They may be meaningful to describe travel experiences, for example, adjectives can express visitors' emotional attitudes toward lodging, but don’t provide specific information about antecedents of guest satisfaction. After recoding, the remaining 80 terms are closely related to the second question. Each of the terms has a score of term frequency (tf) inverse document frequency (idf), which is used to evaluate the importance of a term to a text file (Blei & Lafferty, 2006).

Due to the large amount of the eighty terms, factor analysis was used to reduce the dimensions and to study the internal relationship between the numerous variables to achieve the purpose of data concentration (Xiang et al. 2015). It is still appropriate to use a factor analysis with the non-metric variables because the variables in focus are interrelated (Hair et al., 2009; Xiang et al. 2015)

Then, after checking the data match the underlying assumptions of regression analysis, a multi-linear regression analysis was made using the extracted components (PCi) as independent variables and the house ratings as the dependent variable to examine the relationship between the attributes of guests’ experiences and customer satisfaction. PCi scores are normalized between 0 and +1, and house ratings between 0 and 100, with the ordinary least squares (OLS) solution fitting a linear model.

# Results

The section includes the main findings of this research, first, we gave a description of collected data (rating scores and experience-related words in text reviews). Second, the results of statistical analysis (factor analysis and regression analysis) were presented to answer the research question.

Table 2 shows the distribution of rating scores of Airbnb listings in London. As can be seen, Airbnb listings with ratings scores from 90 to 100 constitute 68.2% of all Airbnb listings, indicating most listings of hosts of inner London get a positive evaluation from the hosts. The average satisfaction scores of these listings is 91.48 with a standard deviation of 10.06. The average number of reviews per listing is 3.07 .

INSERT TABLE 2 NEAR HERE.

Figure 1 shows the distribution of term frequency (TF). The mean of TF in reviews is 18.82 (i.e., words appear about 19 times on average) with a standard deviation of 221.70. As shown, the distribution conforms to the long tail theory, that is, about 96% of the words appear less than 20 times in the review. These indicate that high-frequency keywords (i.e., most frequently discussed words) contain about 73% of the word types. High frequency words constructs the core of reviews, including basic and crucial aspects for the accommodation sharing experience. Long tail words, referring to the words with low frequencies, represents other relatively less important aspects which some guests pay attention to.

Table 1 shows the appearance frequency of the top 100 words, which were used to extract key guest experiential factors with its TF-IDF value to explore their relationship with guest satisfaction.

These high frequency words reflects the terms people most frequently discussed in the sharing experience, which can be classified as the following themes utilizing the method of multi-dimension scaling analysis (Figure 2): (1) Housing factors. The theme has the highest frequency, with a total of 151,527 times. The most frequently mentioned term is the house itself (apartment, house, room, place, flat). In particular, the cleanliness of apartments appeared, indicating the importance of cleanliness which is similar to hotel experiences. (2)Location occurs the second most popular category, only next to housing factors, with a total of 108,848 times (See Table 1). The city, the region(area), the location(street), the transportation (station, bus, tube), accessibility(convenient) and the nearby facilities (pub, shop, restaurant) all have an influence on the experience, and the tourists are particularly interested in surrounding restaurants. (3)The third mostly mentioned category is apartment facility. Apart from bed, bathroom, towels and other necessary facilities, embedded facilities still contain kitchen facilities specific to shared housing. In addition, the space and quietness of apartment also matters in tourists’ reviews. (4) In the host-guest interaction factors, the tourist experience is related to the owner’s friendliness, communication, timeliness and flexibility. The high-frequency words in these factors are positive evaluation vocabulary, such as warm, friendly, helpful, etc., from which it can be seen that tourists have a positive tendency on evaluating the hosts on Airbnb, generally agreeing that the hosts are friendly, easy to communicate and making good interaction with the guests. The finding supports the conclusion reached in previous studies that Airbnb’ appeal to tourists as it offers guests the chance to interact with the hosts and have a more “local” experience (Guttentag,2013; Möhlmann, 2015; Liang et al., 2015; Tussyadiah, 2016; Tussyadiah and Zach, 2017; Wiles and Crawford, 2017).(5) Overall sensation refers to the overall evaluation of accommodation experiences, the words mainly are positive emotional vocabulary, such as beautiful, great, fantastic, excellent, amazing, love, etc., indicating that tourists’ experience of the accommodation sharing tends to be positive and highly satisfied in general. (6) Other categories of factors involve picture, breakfast, key, show and provide. The access of breakfast and key, the picture hosts showed on Airbnb and other factors may also affect the tourist experience.

INSERT FIGURE 1&2 NEAR HERE

In order to further examine the underlying semantic structure from the dataset (Xiang et al; 2015) and make it easier to construct a regression model with guest satisfaction scores, a factor analysis has been used to reduce the terms. Table 3 shows the results of factor analysis after Varimax with Kaiser normalization rotation. It can be seen that this step extracts more specific attributes of experience (PCi) related to guest satisfaction from the 6 categories of the last section.

INSERT TABLE 3 NEAR HERE

The result of the KMO(Kaiser–Meyer–Olkin) measure of sampling adequacy for the overall data set reveals that it is acceptable (0.631). Bartlett’s test of sphericity also reveals that it is highly significant (approx. chi-square: 5384.852, df: 1770, Sig.: 0.000). It shows that the data is suitable to factor analysis. Using the Very Simple Structure criterion (Revelle & Rocklin, 1979) for the data suggests that eight components are optimal. Combined with screen plots and eigenvalue, 8 principle components consisting 43 terms were extracted from 80 terms explaining 26.22% of all variance.

Xiang et al. (2015) and Sánchez‐Franco et al.(2016) suggest, unlike factor analysis based upon metric data, factors obtained from this analysis represent the common semantic spaces in customer reviews. Based on this, factors were named by analysing the terms in the specific factor and the sentiment examples (positive or negative) of each factor is presented in Table 4. PC1 is the “traffic & accessibility” factor, containing 6 related terms, that is, “tube”, “station”, “close”, “convenient”, “bus” and “distance” and the high frequency words “close”, “convenient” convey a positive sentiment space of this factor. PC2 contains 5 terms (“bathroom”, “shower”, “bed”, etc.), indicating in this semantic space guests tend to talk or care about the apartment facilities, so it was named “apartment facilities”. After extracting the texts related to “bathroom”, “shower” and “bed”, word frequency count was conducted and terms such as “small”, “problem” were found to be the high frequency words, indicating the negative meaning of the factor (refer to the sample of reviews in Table 5). PC3 was named “interaction and communication”, because it reflected the communication and encounters between guests and hosts, such as “check (in/out)”, “responsive”, and “recommend”. PC4 “environment and location” refers to the location of the room and if the environment is quiet or noise, as well as sanitation and cleanliness. High frequency words “noise” indicates the negative semantic space. PC5 represents guests’ emotional experience such as the feelings of “home” and “warm”. PC6 contains “bar”, “pub”, etc., suggesting that guests usually want to get some other added services around the house which is linked with their need for entertainment. PC7 refers to the “friendliness and helpfulness of host”, also points out when guests mention the host, they are more likely to evaluate or describe them like this. These three factors (PC5, PC6, PC7) are identified to convey a positive sentiment space. PC8 represents other specific attributes of the house such as space (“small”, “space”) and comfort, which includes both the positive and negative sentiment.

INSERT TABLE 4 NEAR HERE

The factors and listed terms are the representing attributes of contributing to the guest satisfactory according to the tfidf scores. It also conveys a new sight of the novel dimensions of antecedents of guest satisfaction, because the methods of tfidf algorithm and factor analysis deeply observed the internal structure of the original review data and represented the knowledge the data itself produced.

Next, a regression analysis was conducted. Results can be found in Table 5.

INSERT TABLE 5 NEAR HERE

Table 4 shows the ANOVA results using house ratings as the dependent variable and PCi as independent variables. Five factors including PC1 (Traffic & accessibility), PC2(Apartment facilities), PC4(Noise & location), PC5(Home experience) and PC7(Friendliness of host) are significant (at the *p*=.05 level), while other factors(PC3,6,8) are insignificant. And note that although these factors are not significant, we can’t draw the conclusion that there is no relationship between these factors and house ratings (Sánchez‐Franco et al., 2016).And among the significant factors, PC1, 5& 7 have a positive effect with PC2 &4 having a negative effect on the house ratings.

# Discussion

In order to understand the relationship of guest experience and their satisfaction, text analytics and statistical analysis approach were utilized in this research to analyse the rating scores and text reviews on Airbnb.com. The results of multi dimension analysis and factor analysis on text reviews provide the understanding of how guests tell about their lodging experience in sharing economy platform and reveal eight key factors of guest experience expressed online, including traffic & accessibility, apartment facilities, interaction & communication, noise & location, home experience, nearby facilities, friendliness of host and space & comfort. The result of regression analysis indicates that five out of these eight factors have a significant effect on the guest satisfaction. Among them, the absolute value of standardized coefficients of PC2 (Apartment facilities -.308) is the largest, suggesting that the core product or facilities provided by sharing houses such as bed, bathroom, kitchen, are always the most powerful antecedents to the guest satisfaction. The coefficient of the factor loadings of related words are all positive (Tab. 3), suggesting that the higher mentioning of these terms (related to apartment facilities) is associated with a lower level of guest satisfaction. Another negative factor is PC4 (Noise & location), this reveals that a higher house rating is associated with the terms such as “noise”, “location” not being mentioned in the context.

According to the co-efficiency of positive factors (Table4), the effect of PC5 (Home experience) is the highest followed by PC1 (Traffic & accessibility) and PC7 (Friendliness of host). It is interesting to see that home experience is the second important influencing factor (after PC2-Apartment facilities) of the house ratings according to the standardized coefficients. We can see the distinction of antecedents of guest satisfaction in the context of sharing accommodation in contrary to the hotel product, that is, in the sharing accommodation, guests have the opportunity to experience the feeling of staying at another home in the destination (Möhlmann, 2015), and this proves to be a distinctive and important contributor to guest satisfaction. As for PC1 (Traffic & accessibility), another positive factor, the results suggest that a higher level of guest satisfaction is associated with the more appearance of related terms (i.e., tube, station, close) in the reviews. The sentiment space represented by PC 7 is the guests’ description of the host attitude such as friendly and helpful. The positive coefficient reveals that the high mentioning of these terms is along with a high level of satisfaction.

Above all, the core facilities (PC2-Apartment facilities) and the experience of feeling at home (PC5) are more important to the guest satisfaction than the location, environment and the attitude of hosts. The results suggest no matter in the experiences of sharing or staying in a hotel, the apartment facilities are the core products and are most concerned by guests (Qu, Ryan and Chu, 2000; Shanka and Taylor, 2004). In addition, the novel experience of feeling at another “home” has been identified in the guests’ satisfaction of accommodation sharing, supporting pervious research on sharing experiences (Wiles and Crawford, 2017).

The effect of the two factors on customer satisfaction in sharing accommodation can be explained by the Two-Factor Theory of Herzberg (Herzberg, 1966).

Based on two factor theory (Herzburg, 1966), the existence of hygiene factors may not cause too high customer satisfaction, but their absence or poor quality will necessarily lead to customer dissatisfaction. The attributes with significantly negative correlation coefficients including apartment facilities, interaction& communication, noise & location were identified as hygiene factors, because negative correlation represents the mention of these terms is associated with the low satisfaction scores. The semantic words and sample of reviewers’ semantic expressions of the attributes in Table 5 were presented to validating and understanding the meaning represented by negative coefficients. Also, the negative correlation coefficients exist in Xiang et al.(2015)’ study of hotel consumer satisfaction based on big data analytics. The apartment facilities, which have a dominant and significant negative impact on customer satisfaction, are the hygiene factors in sharing accommodation. After checking the reviews involving apartment facilities, the related semantic space tends to be the complaint of the functions of the specific facilities such as broken lock, cold room, dirty bedding, small bathroom etc*.* Therefore, the mentioning of facilities is associated with a low rating score. It is the low quality of facilities that result in guests' dissatisfaction. The largest coefficients of facilities also reveals that if hygiene factors do not meet the expectations or needs of tourists, it is quite difficult to improve customer satisfaction, although there are motivators.

The existence of motivators has a positive contribution to customer satisfaction. In the results, home experience is the unique enjoyment of feeling at home and gives customers emotional and psychological satisfaction, therefore, contributing to the high ratings of the house (related to high satisfaction).

Another critical and positive factor of customer satisfaction is PC7(friendliness of host), which also validates the importance of interaction between hosts and guests in the experience of sharing accommodation concluded by others (Guttentag,2013; Möhlmann, 2015; Tussyadiah, 2016; Tussyadiah and Zach, 2017; Wiles and Crawford, 2017). Home experience, friendly attitude and good interaction with hosts, which is often accompanied by each other are identified to be the important roles in shaping a satisfying accommodation experience. This can be seen from the following reviews with a high rating scores.

“*The host was great and if we had a question responded promptly with an answer.*”

“*We had a very nice stay. Very handy location for public transportation. Everything went very smoothly. Prompt, helpful communication and the place was very nice.*”

 “*We had a very enjoyable 6-day stay in the apartment. Anthony was very prompt in his responses and so we didn't have any issues whatsoever during our stay. Overall, great find for us during our trip to London.*”

The importance of host-guest interaction verifies the significant value of the service dominant logic in the tourism and hospitality industries. Vargo and Lusch(2004) believe that customers are co-creators of value, and traditional form of value creation is within the enterprise (such as hotel facilities, decoration, disability facilities, luggage and breakfast services), and then exchange product or service with customers in the market. Now, more emphasis is placed on the value co-creation of enterprises and customers. Customers no longer simply accept the value created by enterprises, customers participate in the value creation of enterprise products, and share the value created together. In this regard, compared with traditional hotels, accommodation sharing encourages customers to cooperate with the host (communicate, share ideas, and co-organize recreational activities). In the process, the host and guest enjoy unique experiences, social interaction and interpersonal relationships, which are the added value of the co-creation process. The core values of shared experience are the core values of shared housing core competitiveness. Therefore, it can been seen that effective host-guest interaction, especially the good attitude of hosts play the crucial role in the process of co-creation value, contributing to guest satisfaction and core competitiveness advantage of accommodation sharing.

# Conclusions

Although some researches have paid great attention to analysing the consumer experience or the satisfaction of accommodation (Möhlmann, 2015; Tussyadiah, 2016; Tussyadiah and Zach, 2017), to the best of our knowledge so far, there have been no research combining the guest experience and their satisfaction to explore how the customer experience affect their satisfaction in the sharing context. Recognizing the research gap, in this paper, we use a data driven approach of text analytic method to explore guest satisfaction of sharing in the perspective of sharing experiences. This research reveals the critical factors of guest satisfaction of sharing include apartment facilities, home experience, environment and location, traffic and accessibility, friendliness of hosts (ranked according to the importance). And these factors are identified to play different roles in affecting customer satisfaction as hygiene factors (apartment facilities) and motivators (home experience and host attitude) separately based on two-factor theory (Herzberg, 1966). In other words, the poor quality of hygiene factors (in this case, poor apartment facilities) will cause the decline of guest satisfaction, and the existence of motivators, such as the home feeling experience and the friendly attitude of the hosts, will lead to high level of guest satisfaction.

This research contributes to the existing literature in several aspects. First, this study combines several research methods, including big data collection and processing, text mining and statistical analysis. This research is based on user generated content on the internet, which is the real-time and realistic expressions of consumers about their lodging experience. It takes the data as drivers to analyse and solve problems with the information and patterns reflected by the review data, provides an in-depth understanding of the internal structure of data (tourist experiences expressed online).

Second, the paper contributes to our understanding of customer satisfaction in the field of sharing economy (i.e. Möhlmann, 2015; Owyang, 2015; Xiang et al., 2015; Tussydiah, 2016; Sánchez‐Franco, 2016; Tussyadiah and Zach, 2017) by initiating the investigation of satisfaction from the perspective of customer experiences. The findings identified eight factors that guests really care about in their experiences of sharing and discuss their relationship with satisfaction. Further results revealed the two different groups of factors as hygiene factors and motivators. In addition, we demonstrated the usefulness of two factor theory (Herzberg, 1966) in exploring antecedents of customer satisfaction in sharing accommodation as well. For hygiene factors, the findings suggest the product is always the core element, a poor quality of which will lead to a decrease of guest satisfaction. If a guest is disappointed with the absent of hygiene factors, it is difficult for him to get interested in other aspects of this experience and have a high evaluation. This provides an advice to accommodation sharing to pay attention to facilities and make sure guests can receive reliable assistance and timely problem solving. For the motivators, it is suggested that home experience, friendly host attitude and good host guest interaction are the important incentive to achieve high level of guest satisfaction, loyalty and good word of mouth. This supports the findings of previous studies that Airbnb’ distinguished competitive advantage is that it offers guests the chance to interact with the hosts and have a more “local” experience (Guttentag,2013; Möhlmann, 2015; Liang et al., 2015; Tussyadiah, 2016; Tussyadiah and Zach, 2017; Wiles and Crawford, 2017). Therefore, this study extends the two factor theory into the consumer satisfaction of accommodation sharing. This enriches the research on customer experience and satisfaction in the emerging literature of sharing economy and tourism research.

# Limitations and future research

As many other research, this study has its own limitations. First, people’s online reviews do not fully reflect their accommodation experience because of self-selection bias (Li and Hitt, 2008; Xiang et al., 2015) and Airbnb data also seem to have very few low score reviews. Second, the sample represented only covers the stay in London, in the future, other countries and other areas may be explored to compare a cross-cultural influence on experiences. Future research can consider the impact of other variables on customer satisfaction. For example, as price affects tourists’ decisions and people only choose and review places that they can afford, scholars could compare different elements of tourists' experience and satisfaction with different price range. In the future, a combination of different methods based on big data analysis and small data of questionnaire or interviews can be conducted to compensate the triangulation of research.

# References

About us. Airbnb(2016).Retrieved on April 27, 2017 from https://zh.airbnb.com/about/about-us

Bai, Q., Xiong, G., Zhao, Y., & He, L. (2014). Analysis and detection of bogus behavior in web crawler measurement ☆. *Procedia Computer Science,* *31*, 1084-1091.

Cadotte, E. R., & Turgeon, N. (2016). Key factors in guest satisfaction:. *Cornell Hotel & Restaurant Administration Quarterly*, 28(4), 44-51.

Callan, D. R. J., & Bowman, L. (2000). Selecting a hotel and determining salient quality attributes: a preliminary study of mature british travellers. *International Journal of Tourism Research,* *2*(2), 97-118.

Chan, K. L. J. C., & Baum, T. (2007). Researching consumer satisfaction: an extension of herzberg's motivator and hygiene factor theory. *Journal of Travel & Tourism Marketing，*(1), 71.

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). *Business intelligence and analytics: from big data to big impact*. Society for Information Management and The Management Information Systems Research Center.

Fruchterman, T.M.J., and E.M. Reingold.(1991). Graph drawing by force-directed placement. S*oftware - Practice and Experience*, 21(11), 1129-1164.

Geiger, A., Horbel, C., & Germelmann, C. C. (2017). “give and take”: how notions of sharing and context determine free peer-to-peer accommodation decisions. *Journal of Travel & Tourism Marketing*, 1-11.

Gibbs, C., Guttentag, D., Gretzel, U., Morton, J., & Goodwill, A. (2017). Pricing in the sharing economy: a hedonic pricing model applied to airbnb listings. *Journal of Travel & Tourism Marketing*(15), 1-11.

Guo, Y., Barnes, S. J., & Jia, Q. (2017). Mining meaning from online ratings and reviews: tourist satisfaction analysis using latent dirichletallocation. *Tourism Management,* *59*, 467-483.

Guttentag, D. (2015). Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector. *Current Issues in Tourism,* *18*(12), 1192-1217.

Guttentag, D., Smith, S., Potwarka, L., & Havitz, M. (2016). Why tourists choose airbnb: a motivation-based segmentation study. *Journal of Travel Research,* *57*(1), 004728751769698.

Heo, C. Y. (2016). Sharing economy and prospects in tourism research. *Annals of Tourism Research,58*, 166-170.

Herzberg, F.(1966). *Work and the Nature of Man*. World Publishing, Cleveland.

Kathan, W., Matzler, K., & Veider, V. (2016). The sharing economy: your business model's friend or foe?. *Business Horizons,* *59*(6), 663-672.

Kim, B., Kim, S., & Heo, C. Y. (2016). Analysis of satisfiers and dissatisfiers in online hotel reviews on social media. *International Journal of Contemporary Hospitality Management*, 28(9).

Knutson, B. J. (1988). Frequent travelers: making them happy and bringing them back \*. *Cornell Hotel & Restaurant Administration Quarterly,29*(1), 82-87.

Liang, L. J., Choi, H. C., & Joppe, M. (2015). Understanding repurchase intention of airbnb consumers: perceived authenticity, electronic word-of-mouth, and price sensitivity. J*ournal of Travel & Tourism Marketing*, 1-17.

Lipsman, A.(2007). Online consumer-generated reviews have significant impact on offline purchase behavior. comScore.

Matzler, K., Veider, V., & Kathan, W. (2015). Adapting to the sharing economy. *MIT Sloan Management Review*, 71–77.

Meyer, C., & Schwager, A. (2007). Understanding customer experience. *Harvard Business Review,85*(2), 116.

Möhlmann, M. (2015). Collaborative consumption: determinants of satisfaction and the likelihood of using a sharing economy option again. *Journal of Consumer Behaviour,* *14*(3), 193-207.

Mudambi, S. M., & Schuff, D. (2010). What makes a helpful online review? a study of customer reviews on amazon.com. *Social Science Electronic Publishing,* *34*(1), 185-200.

Owyang, J. S. A. G.(2015). Sharing is the new buying: How to win in the collaborative economy. *Sb Business Weekly*.

Pan, B., T. Maclaurin & J. C. Crotts.(2007). Travel blogs and the implications for destination marketing. *Journal of Travel Research*, 46, 35-45.

Pine, B. J. & Gillmore. J. H. (2011). *The Experience Economy*. Boston: Harvard Business Review Press.

Poon, K. Y.; Huang,W. (2017). Past experience, traveler personality and tripographics on intention to use Airbnb. *International Journal of contemporary Hospitality Management*, 29(9), 2425-2443.

PricewaterhouseCoopers.(2016). Demand and capacity for hotels and conference centres in londons: Report to the Greater London Authority. London: the Greater London Authority.

Quattrone, G., Proserpio, D., Quercia, D., Capra, L., & Musolesi, M. (2016). Who benefits from the sharing economy of airbnb?. *Social Science Electronic Publishing*, 1385-1394.

Quinby, D., and M. Gasdia. 2014. “Share This! Private Accommodation and the Rise of the New Gen Renters.” Report. PhoCusWright.

Rhee, H. T., & Yang, S. (2015a). Does hotel attribute importance differ by hotel? Focusing on hotel star-classifications and customers’ overall ratings. *Computers in Human Behavior*, 50, 576-587.

Rhee, H. T., & Yang, S. B. (2015b). How does hotel attribute importance vary among different travelers? An exploratory case study based on a conjoint analysis. *Electronic Markets*, 25(3), 211-226.

Rifkin, J. (2014). *The zero marginal cost society: The internet of things, the collaborative commons and the eclipse of capitalism*. New York: Palgrave Macmillan.

Rhee, H. T., & Yang, S. (2015a). Does hotel attribute importance differ by hotel? Focusing on hotel star-classifications and customers’ overall ratings. *Computers in Human Behavior*, 50, 576-587.

Romesburg, H. C.(1984). *Cluster Analysis for Researchers*. Belmont, CA: Lifetime Learning Publications.

Sánchez‐Franco, M. J., Navarro‐García, A., & Rondán‐Cataluña, F. J. (2016). Online customer service reviews in urban hotels: a data mining approach. *Psychology & Marketing*, 33(12), 1174-1186.

Schmitt, B. H. (2003). *Customer Experience Management Revolutionary Approach:* John Wiley & Sons.

Shanka, T., & Taylor, R. (2004). An Investigation into the Perceived Importance of Service and Facility Attributes to Hotel Satisfaction. *Journal of Quality Assurance in Hospitality & Tourism*, 4(3-4), 119-134.

The Greater London Authority.(2014). The value and volume of tourism, London Boroughs (2011). Retrieved from <http://www.lda.gov.uk/server/show/ConWebDoc.2577.2016.11>.

Toutanova, K., D. Klein, C. Manning, and Y. Singer.(2003). Feature-rich part-of-speech tagging with a cyclic dependency network. *Proceedings of HLT-NAACL*, 252-259.

Tussyadiah, I. P., & Pesonen, J. (2015). Impacts of Peer-to-Peer Accommodation Use on Travel Patterns. *Journal of Travel Research*.

Tussyadiah, I. P. (2016). Factors of satisfaction and intention to use peer-to-peer accommodation. *International Journal of Hospitality Management*, 55, 70-80.

Tussyadiah, I.P., Zach, F. (2017). Identifying Salient Attributes of Peer-to-Peer Accommodation Experience. *Journal of Travel & Tourism Marketing*, 34(5), 636-652

Tussyadiah, I. P., & Sigala, M. (2018). Shareable tourism: tourism marketing in the sharing economy. *Journal of Travel & Tourism Marketing*, 35(1), 1-4.

Vargo, S. L. & R. F. Lusch. (2004). Evolving to a New Dominant Logic for Marketing. *Journal of Marketing*, 68, 1-17.

Wiles, A., & Crawford, A. (2017) Network hospitality in the share economy: Understanding guest experiences and the impact of sharing on lodging, *International Journal of Contemporary Hospitality Management*, 29(9),2444-2463.

World Travel Market (WTM) London. (2014). “WTM Global Trends Report: The Latest Trends from around the World.” http://www.wtmlondon.com/files/wtm\_global\_trends\_2014.pdf (accessed January 1, 2015).

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

Yang, Y., Tan, K. P., & Li, X. R. (2019). Antecedents and consequences of home-sharing stays: Evidence from a nationwide household tourism survey. *Tourism Management*, 70, 15-28.

Zhou, L., Ye, S., Pearce, P. L., & Wu, M. (2014). Refreshing hotel satisfaction studies by reconfiguring customer review data. *International Journal of Hospitality Management*, 38, 1-10.