A Cross-Platform Comparative Study of User Review Texts from Shared Accommodation and Hotels: Postprint Based on LDA Topic Modeling, Social Network Analysis, and Sentiment Analysis

Authors: Chi Maomao, Pan Meiyu, Wang Weijun

Date: 2023-04-01T00:00:00+00:00

Abstract

[Purpose/Significance] Shared accommodation and hotel booking platforms may simultaneously exhibit substitutability and complementarity. However, current literature has not adequately explored which specific products and services these characteristics are manifested in, necessitating further cross-platform comparative research. [Method/Process] This study selects Ctrip hotel booking platform and Xiaozhu short-term rental platform as experimental subjects, collects 86,635 user review texts related to Beijing listings, and conducts cross-platform comparative analysis of user textual reviews by integrating LDA model, topic social network, and topic sentiment analysis methods. [Results/Conclusion] The research identifies similarities and differences between users of the two platforms regarding review topics, topic social networks, and topic sentiment, thereby explaining the substitutability and complementarity of products and services across the platforms from a micro-level user review perspective. The findings offer important practical implications for platform managers in the development and improvement of accommodation products and services.

Full Text

A Cross-Platform Comparative Study of User Review Texts on Shared Accommodation and Hotel Reservation Platforms: LDA-Based Thematic Social Network and Sentiment Analysis

Chi Maomao[1,2], Pan Meiyu[1], Wang Weijun[3]
[1]School of Information Management, Central China Normal University, Wuhan

430079
[2]E-commerce Research Center of Hubei Province, Central China Normal University, Wuhan 430079
[3]Key Laboratory of Adolescent Cyberpsychology and Behavior, Central China Normal University, Wuhan 430079

Abstract:
[Purpose/Significance] Shared accommodation and hotel reservation platforms may exhibit both substitutability and complementarity, yet current literature lacks investigation into which specific products and services embody these relationships, necessitating further cross-platform comparative research. [Method/Process] This study selects Ctrip' s hotel reservation platform and Xiaozhu' s short-term rental platform as experimental subjects, collecting 86,635 user reviews of relevant listings in Beijing. The analysis integrates LDA modeling, thematic social network analysis, and sentiment analysis to conduct a cross-platform comparative analysis of user text reviews. [Result/Conclusion] The findings reveal similarities and differences between the two platforms in terms of review themes, thematic social networks, and thematic sentiment, explaining the substitutability and complementarity of the two platforms' products and services from the perspective of micro-level user reviews. The results provide important practical guidance for platform managers in developing and improving accommodation products and services.

Keywords: cross-platform comparison; text topic mining; social network analysis; sentiment analysis
Classification Codes: TP391.1; F719.2
DOI: 10.13266/j.issn.0252-3116.2021.02.011

The sharing economy model has rapidly penetrated various sectors of the national economy in recent years, particularly in the tourism industry. Shared accommodation platforms (such as Airbnb, Tujia, and Xiaozhu) have become an important accommodation option for travelers. The development of shared accommodation platforms has impacted hotel reservation platforms and captured part of the hotel market. According to the *China Shared Accommodation Development Report 2019*, the shared accommodation market accounted for 6.1% of the national accommodation industry in 2018, with the market size expected to maintain a 50% growth rate over the next three years. However, unlike traditional hotels, shared accommodation serves a new customer base and stimulates the transformation and upgrading of traditional accommodation industries (including hotels). From 2015 to 2018, China' s shared accommodation development contributed an average annual growth of 2.1 percentage points to the accommodation industry. Therefore, the products and services of the shared accommodation and hotel industries may simultaneously exhibit substitutability and complementarity, but further cross-platform comparative research is needed to identify which specific products and services embody these relationships.

Existing literature on the relationship between shared accommodation and hotel reservation platforms primarily analyzes the impact of shared accommodation on the hotel industry from a macroeconomic perspective, treating shared accommodation (e.g., Airbnb) as a disruptive innovation and examining its effects on traditional hotel pricing and sales. Some studies suggest that shared accommodation and hotels are substitutable. For example, K.L. Xie et al. analyzed data from 86 hotels in Austin and found that Airbnb supply negatively impacted hotel performance, noting that this effect was not influenced by hotel quality attributes (star ratings) [2]. Similarly, T. Dogru et al. found that Airbnb supply negatively affected hotel revenue and occupancy rates in cities like London and Paris over a ten-year period (2008-2017) [6]. However, I. Blal et al. analyzed data from 101 hotels in San Francisco and found both substitutability and complementarity between Airbnb entry and hotel performance, where complementarity was reflected in the fact that Airbnb supply quantity did not affect hotel revenue, while substitutability was evident in the negative correlation between Airbnb' s average price and user satisfaction with hotel performance [1].

To further mine user evaluation information on tourism reservation platforms, relevant literature has conducted text analysis of user reviews on hotel reservation and shared accommodation platforms from a micro perspective. Scholars have identified thematic dimensions that users focus on for hotel reservation platforms. For instance, Zhao Xuefeng et al. selected the top 10 reviews from 150 hotels in Beijing, Shanghai, and other cities on Ctrip, using the DBSCSN clustering method to summarize five main dimensions of user concern: hardware provision, service quality, environment, food, and value for money [3]. Wu Weifang et al. collected hotel review data from Las Vegas, using K-means clustering to further identify location, internet provision, and cleanliness as three thematic dimensions [7]. Other research has found that hotel staff service has a greater impact on users than other feature dimensions [8]. Additional literature has attempted to mine the thematic dimensions that shared accommodation platform users focus on. For example, M.M. Cheng et al. collected review texts from Sydney on Airbnb, using word co-occurrence analysis to find that user themes mainly include accommodation facilities, host interaction, and location [4]. Lu Changbao et al. divided Airbnb user review themes into eight dimensions based on high-frequency feature words and their linguistic context, proposing two additional themes: home-like quality and authenticity (cultural atmosphere) [5].

In summary, current research on the relationship between hotel reservation and shared accommodation platforms focuses on two aspects: macroeconomic impact (i.e., shared accommodation as a disruptive innovation affecting traditional hotels) and micro-level user reviews (i.e., mining user review themes on platforms). However, three gaps remain: First, existing literature on the relationship between hotel reservation and shared accommodation platforms primarily analyzes the impact of shared accommodation on traditional hotels from a macroeconomic perspective, lacking systematic micro-level comparative research from a user review perspective. Second, current text review studies on tourism platforms are limited to single platforms, separately conducting theme mining for Ctrip' s hotel reservation platform and Airbnb' s shared accommodation platform, without cross-platform comparative analysis of user review texts. Third, current tourism literature mainly relies on theme mining for text review analysis, lacking detailed analysis of thematic association relationships and sentiment analysis results, making it difficult to present differences in hotel quality attributes at a granular level.

Motivated by these practical and theoretical gaps, this study collects 86,635 user review texts from hotel reservation (Ctrip) and shared accommodation (Xiaozhu) platforms in Beijing, integrating LDA theme modeling, social network analysis (SNA), and sentiment analysis methods for cross-platform user review text analysis. The findings reveal similarities and differences between the two platforms regarding user review themes, thematic social networks, and thematic sentiment tendencies. The results provide important theoretical guidance and practical insights for accommodation platform managers to develop and improve products and services.

1 Related Research

1.1 Research on the Relationship Between Hotel Reservation and Shared Accommodation Platforms Current literature on the relationship between hotel reservation and shared accommodation platforms concentrates on analyzing the impact of shared accommodation on hotels from a macroeconomic perspective, treating shared accommodation as a disruptive innovation and examining its effects on traditional hotel pricing and sales, while proposing that substitutability or complementarity may exist between shared accommodation and traditional hotels. Related studies have found that Airbnb entry negatively impacts local hotel performance, indicating substitutability between shared accommodation and hotels. However, some literature has found both substitutability and complementarity between Airbnb entry and hotel performance. To further mine user evaluation information on tourism reservation platforms, relevant literature has conducted text analysis of user reviews on hotel reservation and shared accommodation platforms from a micro perspective.

To summarize, current research on the relationship between hotel reservation and shared accommodation platforms mainly focuses on macroeconomic impact (i.e., shared accommodation as disruptive innovation affecting traditional hotels) and micro-level user reviews (i.e., mining user review themes on platforms). However, three shortcomings remain: First, current literature lacks systematic micro-level comparative research from a user review perspective. Second, existing text review studies on tourism platforms are limited to single platforms. Third, current tourism literature mainly relies on theme mining for text review analysis, lacking detailed analysis of thematic association relationships and sentiment analysis results.

1.2 Theme Mining Research: LDA-Based Thematic Social Networks and Sentiment Analysis Theme mining methods can effectively identify text themes and mine user online opinions. Currently, there are two main types of theme mining: traditional topic clustering models that rely on text similarity [9-10], and probabilistic topic models such as LDA [12-13]. Probabilistic topic models can efficiently mine topic information contained in large document collections. Since LDA has no strict limitation on text length [14], it has been widely applied in recent years to identify user review themes and hotspots. For example, in hotspot mining on Weibo, relevant literature has used LDA model results for classification inference and theme mining [15]. In e-commerce online reviews, research has used LDA to mine key factors affecting user satisfaction in the fresh food e-commerce industry, such as packaging, freshness, and logistics [17]. In tourism online reviews, literature has used LDA to thematically cluster online reviews of "Lushan tourism" and propose product/service improvement suggestions for different demand types [13].

LDA is a three-layer Bayesian probability model and a document topic generation model. User review texts include a collection of topics that users pay attention to with certain probabilities, while each topic includes a collection of words with certain probabilities. Both review-to-topic and topic-to-word relationships follow multinomial distributions, as shown in formulas (1) and (2):

$$z \sim \mathrm{Multi}(z|(\theta_m)) \quad (1)$$

where z represents the random topic variable generated in the m-th review, and θ_m represents the multinomial distribution parameters for the m-th review in an $M \times K$ matrix, with M representing the number of reviews and K representing the number of topics.

$$w \sim \mathrm{Multi}(w|(\theta_k)) \quad (2)$$

where w represents the random word variable generated by the k-th topic, and θ_k represents the multinomial distribution parameters for the k-th topic in a $K \times V$ matrix, with V representing the number of words.

Although probabilistic topic models can help identify text themes, theme mining alone cannot finely identify relationships between themes or sentiment tendencies toward themes. Building on theme mining, some literature has used co-occurrence relationships between theme words to mine word-word relationship networks, applying LDA-SNA text mining models to information retrieval [18] and news hotspot identification [14], with effectiveness validated. In addition to identifying user attention information feature dimension relationships, user sentiment tendencies also help further understand user attitudes toward themes. Therefore, combining theme mining and sentiment analysis will help identify user response attitudes and attention priorities toward events.

In summary, different from previous text processing methods of "theme identification-co-word analysis" [19] and "theme identification-sentiment analysis" [20], this study introduces social network analysis (SNA) on the basis of theme identification to construct theme-theme and theme-feature word social networks, and uses sentiment analysis methods to quantify user sentiment tendencies toward themes. SNA provides support for analyzing the importance of theme words and associated words, with word co-occurrence also reflecting relationships [21]. This method can identify core themes and social network relationships between themes, facilitating the mining of relationships among user review themes. Sentiment analysis identifies subjective emotions, opinions, and attitudes from text data using natural language processing techniques [22]. Sentiment analysis methods mainly include machine learning-based sentiment classification and sentiment dictionary-based semantic analysis [23]. Accommodation platform review texts are short texts with typically non-standard sentence expressions. Since machine learning is relatively insensitive to emotional symbols, machine learning-based sentiment analysis requires large training samples and manual intervention, resulting in relatively low efficiency [20]. Therefore, this study uses sentiment dictionary-based sentiment analysis for cross-platform user reviews. Specifically, by establishing connections between themes and sentiment attitudes, this study focuses on mining sentiment tendencies of each theme in user text reviews to analyze thematic sentiment differences between hotel reservation and shared accommodation platforms.

2 Research Methods

2.1 Data Sources This study's data comes from online reviews on hotel reservation and shared accommodation platforms from November 2015 to November 2019, with Beijing selected as the data collection location. As China' s political and cultural center, Beijing attracts travelers from around the world, with rich and comprehensive online user reviews that authentically reflect the current state of China' s traditional hotel and shared accommodation industries. Hotel platform review data was crawled from Ctrip, the industry benchmark for online hotel reservations in China, which maintains the top position in the online accommodation booking market with strong competitiveness. Due to the large number of hotel reviews, this study crawled hotels on the homepage of each district, ultimately obtaining valid data from 70 hotels and 55,761 user text reviews. The shared accommodation platform review data was crawled from Xiaozhu, a rising star in China' s online short-term rental industry that has won numerous users with its humanized brand services. The platform currently has over 800,000 listings globally, covering more than 710 cities. This study crawled 5,534 listings from Xiaozhu, with final valid data from 2,635 listings and 30,874 user text reviews after excluding listings without reviews. In total, this study collected 86,635 text reviews.

2.2 Review Text Theme Mining Model The research process consists of five steps: (1) Collect review data from Ctrip and Xiaozhu to form a cross-platform review text database; (2) Preprocess text data, including text segmentation, stop word removal, and part-of-speech tagging; (3) Use LDA topic modeling to cluster text reviews and mine review themes; (4) Construct social networks based on relationships between different themes and relationships between feature words within themes, and conduct sentiment dictionary-based sentiment analysis for each theme; (5) Compare LDA clustering, social network analysis, and sentiment analysis results between hotel reservation and shared accommodation platforms. The specific process is shown in Figure 1 [Figure 1: see original paper].

This section focuses on introducing the main methods used in the research process. First, data preprocessing involves segmenting review content. To improve segmentation efficiency and ensure accuracy and completeness, this study uses a combination of automatic segmentation and manual processing. The stop word dictionary was constructed by integrating and deduplicating the Harbin Institute of Technology stop word list (767 words) and Baidu stop word list (1,395 words). The Python Jieba package was then used to complete text segmentation. Second, LDA topic modeling: Since LDA theme extraction effectiveness is directly related to theme number determination, prior estimation of the number of themes in the dataset is needed before determining the optimal theme number [24]. This study combines relevant literature' s rule of thumb [4-5, 7] to estimate 3-8 themes for LDA, then calculates Coherence values for 3-8 themes for both hotel and shared accommodation reviews to determine the optimal theme number using Coherence scores. After determining the optimal theme number, the Python visualization tool LDAvis is used to visualize feature words under themes. To ensure clear boundaries between themes, feature words that are ambiguous and appear in multiple themes (e.g., "house," "room") are removed, and the top 8 relatively frequent words are selected as theme representatives. Theme names are further confirmed based on semantic relationships of feature words. For themes with high overlap with previous literature, this study combines tourism management literature to code and summarize themes, while for themes with significant differences, one research group determines theme names based on feature word lists, and another group confirms the final theme names.

Third, thematic social network analysis: This section first organizes LDA feature words, using feature words under the same theme as theme feature identifiers to construct a theme-theme external co-occurrence matrix (see Table 1). Off-diagonal elements in the co-word matrix represent the frequency of two keywords appearing together in the same review, while diagonal elements represent the frequency of the word appearing in all reviews. Second, to reveal relationships between feature words within a single theme, an internal co-occurrence matrix is constructed based on feature word-feature word co-occurrence relationships (see Table 2). Finally, Ucinet and Netdraw software are used to visualize thematic social network results for Ctrip and Xiaozhu platforms. Fourth, thematic sentiment analysis: Based on LDA clustering results, sentiment analysis is conducted for each theme using the HowNet dictionary (8,936 words) and manual annotation (989 words) to extract sentiment words. Following reference [20], review text sentiment polarity is divided into three categories: positive, neutral, and negative. Considering the complexity and focus differences of consumer reviews, a single review may evaluate multiple themes simultaneously, resulting in one-to-one, many-to-one, or other matching situations between themes and sentiment words. Therefore, following W. Duan et al.' s method for hotel reviews [8], all reviews are separated into single sentences based on punctuation to match [theme feature word, sentiment word] patterns to confirm user sentiment tendencies for each theme. For example, in the hotel review "The subway station is right next door, very convenient for travel, the staff is very enthusiastic and proactively asks about cleanliness," the analysis process is shown in Figure 2 [Figure 2: see original paper].

3 Empirical Results

3.1 LDA Theme Mining Results Text classification themes are established based on theme word clustering, with thematic coherence scores calculated for different theme numbers. Since optimal theme number determination requires prior estimation, combined with tourism management literature [7], this study estimated 3-8 themes for both platforms and conducted experimental iterations to achieve optimal clustering results. As LDA models extract many feature words, and excessive theme feature words are difficult to use directly in practical analysis [3, 25], this study selected the top 8 relatively frequent words as theme representatives for theme feature identification and summarization. Experimental results show that when the theme number for Ctrip is 7, the coherence score is highest (Coherence Score = 0.421). LDA results indicate seven major themes in platform user text reviews: facility cleanliness, transportation convenience, room hardware, interaction experience, overall feeling, family service, and hotel hardware. When Xiaozhu' s theme number is 6, the coherence score is highest (Coherence Score = 0.420). LDA results show six major themes: facility cleanliness, transportation convenience, room hardware, interaction experience, overall feeling, and bedding supplies. Among these, five themes—facility cleanliness, transportation convenience, room hardware, interaction experience, and overall feeling—are common to both platforms. Additionally, family service and hotel hardware are characteristic themes of Ctrip, while bedding supplies is a characteristic theme of Xiaozhu. Detailed LDA results are shown in Table 3 .

3.2 Thematic Social Network Analysis Results This section further constructs social networks based on LDA results to explore relationships between review text themes on both platforms. Using UCINET6 software, co-occurrence networks are examined to investigate relationships between themes and between feature words within single themes. Social network visualization results are shown in Figure 3 [Figure 3: see original paper].

First, node size in Figure 3 is proportional to theme frequency—larger nodes indicate more user attention. In terms of theme-theme relationships, Ctrip' s seven themes show relatively large proportions for: (1) overall feeling theme with feature words like price, environment, and grade; (2) facility cleanliness theme with feature words like facilities, hardware, and sanitary conditions; and (3) hotel hardware theme with feature words like breakfast, fruit, and restaurant. Among Xiaozhu' s six themes, larger proportions appear for: (1) interaction experience theme with feature words like "big brother," "big sister," and "beauty" ; (2) transportation convenience theme with feature words like bus, subway, and line number; and (3) overall feeling theme with feature words like feeling, overall arrangement.

Second, the thickness of connecting lines between themes is proportional to the frequency of co-occurrence between corresponding theme nodes. In Ctrip' s social network, theme-theme line thickness shows little variation, indicating hotel users' attention is relatively evenly distributed across themes without obvious preference. In Xiaozhu' s social network, the connections among interaction experience, transportation convenience, and overall feeling are particularly strong, indicating these three themes receive concentrated attention from platform users.

Finally, analyzing relationships between feature words within single themes reveals that transportation location, interaction experience, and overall feeling show little difference between platforms, while other themes show significant user differences. The transportation convenience theme has the tightest internal social network, with feature words like bus, subway station, and walking appearing together frequently, indicating both platforms' users focus on accessibility. In the interaction experience theme, Xiaozhu' s social network nodes centered on hosts show less dense connections, while Ctrip' s nodes show more connections. In the overall feeling theme, Ctrip users focus on value for money, while Xiaozhu users focus on style and decoration. In the room hardware theme, Ctrip users focus on air conditioning and heating provision, while Xiaozhu users also pay attention to microwave ovens, washing machines, and projectors. In the facility cleanliness theme, Xiaozhu users, beyond Ctrip users' concern dimensions, also focus on kitchenware cleanliness. In the family service theme, Ctrip users' comments center on children with fewer node connections. In the hotel hardware theme, Ctrip users center on breakfast with more node connections, indicating breakfast, dishes, and restaurant are key considerations. In the bedding supplies theme, sheets are the central node, with quilts and duvet covers also being important simultaneous concerns for Xiaozhu users.

3.3 Thematic Sentiment Analysis Results Thematic sentiment analysis mines sentiment tendencies of each LDA theme to further identify differences in user review texts between platforms. This section primarily uses sentiment polarity to determine positive, negative, and neutral ternary sentiment attitudes [20], calculating sentiment scores shown in Table 4 . After obtaining sentiment scores for each theme, weighted average sentiment intensity is calculated based on overall theme proportions, with results sorted by platform theme proportion shown in Figure 4 [Figure 4: see original paper].

Aggregating sentiment scores across themes, Ctrip' s positive, negative, and neutral sentiment intensities are 0.76, 0.06, and 0.18, respectively, while Xiaozhu' s are 0.82, 0.05, and 0.11. In comparison, Ctrip user reviews show lower positive sentiment scores, while Xiaozhu users' positive emotions vary more significantly across themes. Both platforms show almost no difference in overall negative sentiment scores and variation. Among positive reviews, overall feeling, interaction experience, and transportation convenience are the three highest-scoring themes on both platforms. For negative reviews, Ctrip's negative comments mainly concentrate on room hardware, family service, and hotel hardware themes, while Xiaozhu' s negative comments focus on bedding supplies, room hardware, and facility cleanliness.

4 Discussion

4.1 Main Findings Based on user review text datasets from Ctrip and Xiaozhu, this study mines user comments on hotel and shared accommodation platforms through text analysis. First, LDA modeling identifies user review themes on both platforms. Social network analysis visualizes theme-theme and theme-feature word co-occurrence networks to analyze differences in internal networks between the two platforms. Finally, thematic sentiment analysis reveals specific differences in user sentiment tendencies. The main findings are: (1) The two platforms share five common themes: facility cleanliness, transportation convenience, room hardware, interaction experience, and overall feeling. Common themes reflect shared content of platform user concerns during accommodation experiences. For example, convenient geographical location and transportation are important for both platforms. Related literature has found similar themes—for hotel platforms, Zhao Xuefeng et al. summarized hardware provision, hotel service, food, value for money, and environment [3]; Zhang Meng et al. found that hotel facilities, service level, and surrounding environment significantly impact online booking [26]. For shared accommodation platforms, M.M. Cheng et al. identified accommodation facilities, host interaction, and location [4]; R. Yan et al. found that service, amenities, and overall feeling influence switching intentions [27]. Cross-platform analysis of common themes reveals that substitutability between hotel and shared accommodation platforms is embodied in these five common themes (i.e., users have similar demands for these products/services, making them substitutable). Additionally, through LDA clustering, this study further discovers complementarity between platforms (i.e., users have different demands for products/services, making them complementary). For example, Ctrip' s characteristic themes include family service (with feature words like parking, children, elderly) and hotel hardware, while Xiaozhu' s characteristic theme is bedding supplies (with feature words like sheets, pillows, quilts). (2) LDA-based social network analysis reveals that even when users focus on the same theme, differences exist in feature word associations within themes. For instance, in the interaction experience theme, Ctrip's social network nodes show denser connections than Xiaozhu' s. Previous accommodation review text processing mostly relied on word co-occurrence degrees, neglecting theme-theme relationships [4, 19]. This study introduces theme-theme social networks and finds that Ctrip users' review themes are relatively evenly distributed, while Xiaozhu users concentrate on interaction experience, transportation convenience, and overall feeling. (3) LDA-based sentiment analysis shows that user comments on both platforms are generally positive across themes. Overall feeling, interaction experience, and transportation convenience are the three highest-scoring positive themes on both platforms, indicating similar demands for quality services and products and suggesting substitutability. However, Ctrip' s negative comments concentrate on room hardware, family service, and hotel hardware, while Xiaozhu' s negative comments focus on bedding supplies, room hardware, and facility cleanliness. Negative sentiment results show important differences in product/service themes between platforms, indicating complementarity.

4.2 Research Contributions This study makes three theoretical contributions. First, it identifies differences in user review themes between hotel reservation and shared accommodation platforms through cross-platform comparative text analysis. Previous literature focused on single platforms without cross-platform comparison [4, 7]. Second, this study compares user review themes between platforms from a micro-level text perspective, while previous literature mainly discussed shared accommodation' s impact on traditional hotels from a macroeconomic perspective [1, 6]. Third, different from previous "theme identification-co-word analysis" [19] and "theme identification-sentiment analysis" [20] methods, this study constructs theme-theme and theme-feature word social networks on the basis of theme identification, matching theme words with sentiment words to quantify user sentiment tendencies toward themes. This method integrates LDA, SNA, and sentiment analysis to finely differentiate user review themes between platforms.

4.3 Practical Implications The findings offer practical implications for managers. First, by comparing user review themes on Ctrip and Xiaozhu, this study summarizes five important common themes and identifies characteristic themes for each platform. This enables platform managers to optimize products and services around themes users care about—for example, shared accommodation managers should focus on improving bedding supplies. Second, social network analysis shows that Ctrip users' reviews involve almost all themes simultaneously, while Xiaozhu users' reviews mainly involve interaction experience, overall feeling, and transportation convenience. Different platform managers can achieve more significant returns with less investment—for instance, shared accommodation hosts should highlight user interaction experiences, while hotel managers need to balance product and service quality to create good overall user experiences, potentially learning from shared accommodation' s advantages to provide personalized experiences beyond standardization. Finally, for tourism accommodation platform managers, designing tailored products/services based on user expectations and travel modes, and addressing negative information in reviews, are important ways to maintain competitiveness. Sentiment analysis shows that Ctrip' s room hardware, family service, and hotel hardware have lower sentiment scores, while Xiaozhu' s bedding supplies, room hardware, and facility cleanliness have lower scores. Both platforms should currently focus on products and services related to these themes to maximize user satisfaction with minimal investment. Specifically, hotel managers should recognize the importance of buffet and family services, leveraging their advantages to meet user needs. Shared accommodation managers should emphasize room cleanliness and hygiene management, potentially borrowing hotels' standardized management models to establish rating standards for bedding supplies and facility cleanliness.

References

[1] Blal I, Singal M, Templin J. Airbnb' s effect on hotel sales and hotel performance[J]. International journal of hospitality management, 2017, 67: 174-184. [2] Xie K L, Kwok L. The effects of Airbnb' s price positioning on hotel performance[J]. International journal of hospitality management, 2017, 67: 174-184. [3] Zhao Xuefeng, Tang Qing, Li Yue. Online hotel reputation dimension mining based on customer reviews and corpus[J]. Library and Information Service, 2012, 56(12): 124-129. [4] Cheng M M, Jin X. What do Airbnb users care about? An analysis of online review comments[J]. International journal of hospitality management, 2019, 76: 58-70. [5] Lu Changbao, Lin Sijie. Research on tourists' motivation for choosing online short-term rental accommodation[J]. Economic Management, 2018, 40(12): 153-167. [6] Dogru T, Hanks L, Moody M, et al. The effects of Airbnb on hotel performance: evidence from cities beyond the United States[J]. Tourism management, 2020, 79: 104090. [7] Wu Weifang, Gao Baojun, Yang Haixia, et al. The impact of review text on hotel satisfaction: a sentiment analysis-based approach[J]. Data Analysis and Knowledge Discovery, 2017, 1(3): 62-71. [8] Duan W, Yu Y, Cao Q, et al. Exploring the impact of social media on hotel service performance: a sentiment analysis approach[J]. Cornell hospitality quarterly, 2016, 57(3): 282-296. [9] Li Hui, Wang Liting. Research on microblog hotspot topic discovery based on term heat[J]. Information Science, 2018, 36(4): 45-50. [10] Xia Huosong, Li Baoguo, Yang Pei. Online news comment theme extraction based on improved K-means clustering[J]. Journal of the China Society for Scientific and Technical Information, 2016, 35(1): 55-65. [11] Ishikawa S, Arakawa Y, Tagashira S, et al. Hot topic detection in local areas using twitter and wikipedia[C]//ARCS 2012. Prague: IEEE, 2012: 1-5. [12] Tang Xiaobo, Qiu Xin. Research on high-quality comment mining model for themes[J]. New Technology of Library and Information Service, 2015(Z1): 104-112. [13] Tu Haili, Tang Xiaobo, Xie Li. Research on user demand mining model based on online reviews[J]. Journal of the China Society for Scientific and Technical Information, 2015, 34(10): 1088-1097. [14] Wang Hongwei, Gao Song, Lu Pin. Research on online news hotspot identification based on LDA and SNA[J]. Journal of the China Society for Scientific and Technical Information, 2016, 35(10): 1022-1037. [15] Ye Chuan, Ma Jing. Research on theme discovery algorithm for multimedia microblog comment information[J]. New Technology of Library and Information Service, 2015(11): 51-59. [16] Huang Wei, Zhao Jiangyuan, Yan Lu. Construction and empirical research on network hotspot event topic

drift index[J/OL]. Data Analysis and Knowledge Discovery: 1-15. [2020-07-18]. http://kns.cnki.net/kcms/detail/10.1478.G2.20200713.1744.012.html. [17] Feng Kun, Yang Qiang, Chang Xinyi, et al. Fresh e-commerce customer satisfaction evaluation based on online reviews and stochastic dominance criteria[J/OL]. Chinese Journal of Management Science: 1-12. [2020-07-18]. https://doi.org/10.16381/j.cnki.issn1003-207x.2019.2108. [18] Ye Chunlei, Xing Yanli. Research on personalized retrieval recommendation model for graduate students based on LDA and social network centrality[J]. Library and Information Service, 2015, 59(13): 104-112. [19] Tussyadiah I P, Zach F. Identifying salient attributes of peer-to-peer accommodation experience[J]. Journal of travel & tourism marketing, 2017, 34(5): 636-652. [20] Zhao Changyu, Wu Yaping, Wang Jimin. Twitter text theme mining and sentiment analysis under the Belt and Road Initiative[J]. Library and Information Service, 2019, 63(19): 119-129. [21] Li Yu, Liu Hong, Sun Jianjun. Atlas analysis of research themes in Chinese library science doctoral dissertations[J]. Library Journal, 2018, 37(6): 22-30. [22] Mikolov T, Sutskever I, Chen K, et al. Distributed representations of words and phrases and their compositionality[C]//Advances in neural information processing systems. Lake Tahoe: NeurIPS, 2013: 3111-3119. [23] Zhou Jian, Liu Yanbao, Liu Jiajia. Analysis of knowledge structure and research frontiers in sentiment analysis[J]. Journal of the China Society for Scientific and Technical Information, 2020, 39(1): 111-124. [24] Liu Ziqiang, Xu Haiyun, Yue Lixin, et al. Research on core technology theme identification method based on Chunk-LDAvis[J]. Library and Information Service, 2019, 63(9): 73-84. [25] Zhang Tairui, Chen Yu. Research on health information user transfer behavior based on LDA model factor extraction[J]. Library and Information Service, 2019, 63(21): 66-77. [26] Zhang Meng, Zhang Guangyu, Ye Zuoliang. Research on the impact of online information on hotel online booking—based on Ctrip hotel online booking data analysis[J]. Tourism Tribune, 2011, 26(7): 79-84. [27] Yan R, Zhang K Z K, Yu Y. Switching from hotels to peer-to-peer accommodation: an empirical study[J]. Information Technology & People, 2019, 32(6): 1657-1678. [28] Hou Z, Cui F, Meng Y, et al. Opinion mining from online travel reviews: A comparative analysis of Chinese major OTAs using semantic association analysis[J]. Tourism management, 2019, 74: 276-289. [29] Zervas G, Proserpio D, Byers J W. The rise of the sharing economy: estimating the impact of Airbnb on the hotel industry[J]. Journal of marketing research, 2017, 54(5): 687-705.

Author Contributions:
Pan Meiyu: Data analysis and paper writing;
Chi Maomao: Topic selection, research framework design, paper revision;
Wang Weijun: Research idea formulation, paper revision.

Note: Figure translations are in progress. See original paper for figures.