
Information Technology Support for Rural Homestay Product Innovation: Text Analysis Based on User Comments

Li Na

Department of Art Design, Nanjing University of Finance and Economics, Nanjing, China

Email address:

minilana@163.com

To cite this article:

Li Na. Information Technology Support for Rural Homestay Product Innovation: Text Analysis Based on User Comments. *Urban and Regional Planning*. Vol. 4, No. 4, 2019, pp. 162-167. doi: 10.11648/j.urp.20190404.15

Received: November 18, 2019; **Accepted:** December 9, 2019; **Published:** December 12, 2019

Abstract: Rural tourism has an increasingly significant impact on the design of homestay. Big data analysis of tourists' feedback information should improve the environment art design and promote the innovation of homestay, which provides technical support for design decisions. The research obtains user comment information on the tourism service network platform by the help of web page crawl software. Using ROST CM6 software, text analysis was conducted with the big data of user comment. The results show that the semantic network of tourist comments presents a significant hierarchical relationship. The connection between "homestay", "host", "room" and other nodes is significantly higher than other nodes, so they are important nodes in the network. "Homestay" and its connections can be concluded with three theme categories that constitute the overall dimension of tourists' perception of homestay, among which tourists pay more attention to location. "Room" and its connections present rich perceptual colors, showing the delicacy and comprehensiveness of the tourists' physical sensory experience. "Host" and its connections represent the reception service of homelands. "Luggage", "meal" and "attitude" are the key indicators to evaluate the service quality of the host. From the text analysis, the tourist image perception of homestay facilities is composed of two dimensions: material space and reception service. Tourists' cognition of material space is divided into overall dimension and special dimension. The overall dimension focuses on rational cognition, while the specific dimension focuses on perceptual cognition. The study points out: the homestays presents transition from art design to product design; Constructing the frame of image structure of homestay is the basis of implementing design management; Big data analysis technology can be used to construct image frames and produce design decisions.

Keywords: Multidimensional Image, Rural Homestay Product, Text Analysis, Innovative Design, Decision Support

1. Introduction

Homestay generally refers to a small accommodation facility that shares the spare room of one's own house with visitors and allow them to experience the local natural and cultural characteristics. Homestay is known as B&B (bed and breakfast) in UK, as minshuku in Japanese. The host of homestay can get income by providing overnight bed and breakfast for travelers, meanwhile experiencing local culture is only an additional gain. With the popularity of rural tourism, homestay has become a choice different from standardized hotels for tourists, who enjoy the experience of indigenous culture. For the tourists' experience, more and more hosts pay attention to embody local culture and aesthetic experience in homestays [1], which indicates that the

conceptual design is changing from thinking of homestay as a tourist service facility to as a tourism product. With the changing, the theories and methods of product design should be introduced into the design of homestay.

In term of tourism product, a homestay facility design not only needs to match tourists' needs and satisfactions, but also needs to create extraordinary experience rather than everyday life. In the circumstance of rapid changes of the tourism market, homestay design should be closely related to the production and consumption, which forms a complete cycle [2] and gets efficient iteration of homestay products. It is difficult for homestay designers to grasp the changing taste of consumer by integrating market information only based on their personal experience, and there is a strong demand for

design decisions assist on market research of homestay [3]. Computer aided design is mostly used in drawing, but seldom used in design decisions. Big data analysis plays an increasingly significant role in supporting design decisions due to explosive growth of knowledge and innovation. However, the technical threshold for the application of big data analysis is relatively high. Recently, the development of artificial intelligence technology provides the possibility for cross the threshold by using software on big data analysis.

Text analysis software can assist homestay designer in big data analysis on consumer feedback information [4]. Text analysis is a method to extract useful information from human natural language through computer data-processing [5]. There are three steps of text analysis. First, human natural language is transformed into computer data, then mathematical operation is carried out, and finally, results are interpreted and conclusions are deduced.

Network evaluation studies of homestay facility are mainly concentrated in the field of tourism. O. L. Fan (2011) constructed the evaluation model of homestay facility based on the perspective of customer perception [6]. S. Li et al. (2015) studied the perception of tourism experience of mainland residents in Taiwan [7]. Y. ZH. Liu et al. (2016) used stratified sampling and systematic sampling to extract the network evaluation of 53 family hotels in hongcun village, and studied the experience process and results of tourists in family hotels from the aspects of emotional tendency, experience factors and tourism perception [8]. M. Fang et al. (2016) studied the network evaluation of homestay facilities in Lin'an based on the image of tourist destinations [9]. J. Tong et al. (2018) conducted an empirical study on the attributes and characteristics of experience perception of residents of Dalian homestay facility by combining qualitative and quantitative methods [10]. Q. Sun et al. (2018) analyzed the service quality of rural homestay facilities in Nanyuanling of Wuyi mountain by matching the comments with the service quality evaluation index of homestay facilities [11]. Y. Tang et al. (2019) studied the network evaluation of service quality of Zhenmeilige inn in Lugu lake [12]. T. T. Shi et al. (2019) studied the experience factors of residents in metropolitan homestay [13]. The above studies mainly used the tourist review information of ctrip, qunar, mafengwo, airbnb tourism service platforms, and conducted text analysis with the help of ROST CM6 or grounded theory. The research uses the software ROST CM6 provided by the ROST virtual learning team of Wuhan university, which integrates language, statistics and computer learning technologies to provide visual result output and meet the technical requirements of conventional text analysis. Through text analysis, the perception image of homestay for tourists is obtained, and the data is interpreted from the perspective of product design positioning of homestay.

2. Methods

2.1. Case

Xidi and Hongcun are ancient villages located at south

china in Anhui province. They were included on the world heritage list in 2000, representing traditional huizhou culture, one of the three dominant regional cultures. Xidi and Hongcun have high quality of tourism resource, which gain great attraction and satisfaction from tourists. As one of the earliest tourism exploit village in China, tourism has become its leading industry instead of agriculture. Surrounding villages, such as Lucun, Bishan and Nanping, also benefited from tourism development in Xidi and Hongcun villages.

There is a large quantity of homestay in Xidi and Hongcun villages, with strong features and long operation. The cases of homestay in Xidi and Hongcun villages are suitable subjects for study because there are many kinds of homestays from common local houses to tourism products.

2.2. Research Conception

Comments of the tourists on the homestay accommodation was collected through the tourism website, and text analysis was conducted with the Rost CM6 software. After the programmed operation of Rost CM6, the feature words of tourists' comments were obtained, and the relationship between feature words was used to build semantic network. Semantic network structure presents text content, intuitive and concise, easy to analyze user feedback information. By interpreting the perception and attitude of users on semantic network, it helps designers to make their decisions of homestay products.

2.3. Data Collection

Ctrip was founded in 1999 and listed on NASDAQ stock exchange in the United States four years later. Ctrip has comprehensive tourism business, large and stable number of users, and high information reliability. Therefore, it is selected to obtain the website of research data. User comments were obtained by setting classification of homestay inns, entering the destination and check-in and check-off information (staying for one-night from July 25 to July 26, 2019, when wasn't weekend nor holiday time). Due to the research on tourists' perceptions and attitudes on design of homestay products, the middle and high price homestay products are more representative. The user comments of 37 homestay inns was obtained through the setting one room with the price above 300 yuan. Then, webpage crawl software "octopus" was used to obtain the user comments of these homestays on July 25, 2019. Data merge was carried out on 37 captured files on excel software, duplicate values were deleted, and 13,262 valid comments were finally obtained.

2.4. Data Analysis

Rost CM6 was used for text analysis. Convert the comments on excel into text format and enter them into ROST CM6 for data processing. First, the text data was sorted out, English words and blank lines were removed, synonyms were merged (such as "inn", "hotel" and "homestay" were uniformly changed to "homestay"), and

word segmentation and word frequency analysis were performed. Word frequency analysis was used to assess the importance of a word to a text and to identify the features of a visitor's comment. By default, the software outputs the top 300 words (table 1).

Table 1. Statistical table of word frequency analysis.

ranking	vocabulary	frequency	ranking	vocabulary	frequency
1	host	8076	16	breakfast	1111
2	homestay	6595	17	delicious	1086
3	room	5299	18	decoration	998
4	enthusiasm	3516	19	quiet	972
5	environment	3038	20	satisfaction	950
6	clean	2894	21	facility	933
7	service	2796	22	night	893
8	scenic area	1938	23	please	878
9	location	1937	24	south lake	816
10	convenience	1332	25	gateway	811
11	next time	1258	26	considerate	743
12	check in	1227	27	taste	739
13	feature	1214	28	experience	737
14	comfortable	1204	29	tidy	719
15	hygienic	1183	30	friend	615

After the text was processed one sentence in one-line, semantic network analysis is carried out. A semantic network was a network of nodes and their interrelationships that is

used to structure and present textual content. In a semantic network, information is expressed as a set of nodes, which are connected to each other through a set of marked directed lines to represent the relationship between nodes [14]. The generation of semantic network in Rost CM6 includes five steps: extraction of high-frequency words, filtering of nonsense words, extraction of line features, construction of network, and construction of co-occurrence matrix (table 2). Co-occurrence matrix is to establish the relationship between vocabulary and context through the similar occurrence frequency of high-frequency words, so as to grasp the meaning expressed by words more accurately [15]. Finally, the drawing software Netdraw was used to draw the network graph and carry out degree centrality analysis (figure 1). Degree centrality reflects the centrality of nodes in the network by the degree of nodes, which is indicates by the number of direct connections between nodes. The larger the degree of nodes, the higher the degree centrality of nodes, and the more important the nodes are in the network. Degree centrality includes two indexes of degree centrality and degree centrality. In a directed network, degree centrality represents the node as the "receiver" of the relationship, and degree centrality represents the node as the "sender" of the relationship.

Table 2. Co-occurrence glossary statistics.

Co-occurrence vocabulary	frequency	Co-occurrence vocabulary	frequency	
enthusiasm	host	enthusiasm	considerate	282
room	clean	complete	facility	279
scenic area	gateway	room	feature	263
homestay	location	room	host	250
clean	tidy	delicious	host	227
enthusiasm	service	personally	host	217
clean	hygienic	gateway	host	212
service	host	homestay	service	211
considerate	service	grace	environment	209
homestay	host	homestay	clean	209
scenic area	homestay	considerable	host	206
homestay	environment	room	hygienic	205
attitude	service	homestay	feature	204
room	tidy	style	decorate	198
homestay	gateway	room	facility	193

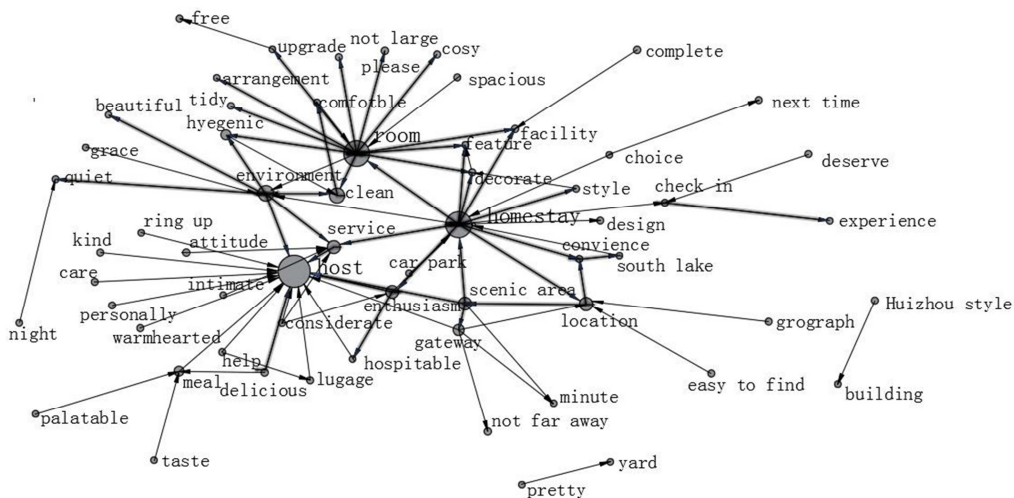


Figure 1. Semantic network degree centrality analysis diagram.

3. Results

Degree centrality shows that there is a significant hierarchical relationship in the semantic network. The connections between nodes of "homestay", "host", "room" and other nodes are significantly higher than that of others, which is 22, 21 and 21 respectively. Therefore, they are important nodes in the network. Node of "Host" has the highest degree of centrality, which means its network influence is the biggest. The out-degree centrality of nodes of "homestay" and "room" are obviously higher than that of in-degree centrality, which indicates that their network control is large. Visitors' comments on homestay revolve around the three dimensions of "homestay", "host" and "room". "host" belongs to the service environment of homestay, while "homestay" and "room" belong to the material environment of homestay. The reception service of the host ("host") directly influences the attitude of tourists' comments, while tourists' perception of the environment of the homestay includes two levels: concrete ("room") and overall ("inn"). "Room" is the specific dimension of tourists' perception of homestay. The sleeping behavior mainly occurs in the room, which leads to the body perception of being fully aware of details. "homestay" is the overall dimension of tourists' perception, and the external connection and internal sharing services (parking, catering, leisure, etc.) of the homestay generate a comprehensive evaluation of the overall dimension of the homestay. As the study focuses on the material environment of homestay products, it focuses on the analysis of the relationship between "homestay", "room" and their connecting.

3.1. Cluster Analysis of "Homestay"

The connection node of "homestay" can be classified into three theme categories, the first one is location, the second one is form, and the third one is reception. The first kind of location-related nodes include nouns: "scenic spot", "doorway", "location", "parking lot", "south lake" (the core scenic spot of Hongcun village); Adjective: "convenient"; Time quantifier: "minute". According to the co-occurrence matrix, the semantics expressed by location-related nodes are accessibility, including the entrance to the scenic spot to the homestay, the homestay to the scenic spot (south lake), the homestay to the parking lot, and the homestay to the service facilities. The tourists' perception of accessibility had two measurement dimensions: time and target, visitors emotion perception for convenient, search words in the text of the context, found that more tourists to 2, 5, 10, 15, 20 minutes to walk time perception scale, 10 minutes feel convenient, 10-15 minutes attitude appear differentiation, generally feel inconvenient to 20 minutes. Some tourists put forward the problem of road identification, which indicates that psychology influences the perception of distance. If the path is easy to identify, the perception of distance is close.

Formal-related nodes include nouns: "environment", "feature", "room", "facility", "style", "design" (noun or verb),

"sanitation" (noun or adjective), and adjectives: "clean", "warm". According to the co-occurrence matrix, there are three levels of word frequency in homestays. The highest frequency is "environment", followed by "feature" and "room", followed by "facility", "style", "design" and "sanitation". It indicates that tourists perceive the homestay, and "environment" expresses tourists' perception of the homestay atmosphere. "Feature" and "room" are the expression of the image and function of the homestay. "Facility", "style", "design" and "hygiene" are more detailed perceptions. Formal-related nodes produce three semantics: experience, characteristics and reception conditions. "Environment" relates to 9 nodes. In addition to "homestay", "room" and "host", there are four adjectives, "graceful", "elegant", "clean" and "quiet", and two nouns, "service" and "sanitation", indicating that "environment" reflects the overall experience of the homestay facility and expresses higher aesthetic pursuit. "Characteristics", "style" and "design" reflect the uniqueness of the homestay, while "room", "facilities" and "sanitation" reflect the reception conditions of the homestay. Tourists' emotional perception is "clean" and "warm".

Reception related nodes include: "host", "service", "passion", "breakfast", "the host" is significantly higher than other common word frequency, host of the homestay facility has the greatest effect in the reception service, the host and the employee's behavior ("services"), attitude ("enthusiasm"), product ("breakfast") affect visitor perception of hospitality services of a homestay facility.

In conclusion, "homestay" consists of three theme categories that constitute the overall dimension of tourists' perception of homestay, among which tourists pay more attention to location. "Homestay" is an integrated concept, including material space and service content, and visitors can experience the homestay through the overall atmosphere perception.

3.2. Cluster Analysis of "Room"

In the degree centrality analysis diagram, "room" not only shares connection node with "homestay" and "host", but also has separate connection. All nodes of the form category of "room" and "homestay" are connected, which indicates that no matter at the specific or overall level, the form has always attracted tourists' attention. Compared with the co-occurrence matrix, the co-occurrence frequency of "room" and "clean" is far higher than that of other rooms, up to 1106, indicating that "clean" is the focus of tourists' attention to guest rooms. The second is "room" and "tidy", which amount to 334, which further indicates that tourists mainly comment on guest rooms through perceptual cognition.

The nouns of "room" cluster nodes include: "feature", "host", "homestay", "sanitation" (noun or verb), "facility", "environment", "style", "design" (noun or verb); Adjectives, including: "clean", "clean", "comfort", "sweet", "comfort", "big", "small", "passion", "beautiful", "complete". In the cluster nodes of "room", there are more adjectives describing

sensibility than "homestay", which indicates that perceptual cognition dominates tourists' perception and attitude. The guest room is a private space for tourists, who stay in it for a long time and use the guest room facilities at close quarters, so that the physical senses are more stimulated. "Clean" is the most important sensory experience, followed by "clean", again for "comfortable", "warm" and "comfortable" and "wide", "small", "passion", "beautiful", "complete". "Clean" and "neat" express the perception of the function and order of the room; "Comfortable", "warm" and "comfortable" express the perception of agreeableness and comfort; Spacious, not much and space form; Passion is the perception of service; The beauty is coexisting with the courtyard, showing the visitors' perception of the external landscape of the room. Completeness is the perception of facilities. In addition to nodes shared with "homestay" and "host" clusters, separate connection nodes can better reflect tourists' unique perception dimensions of rooms, such as "clean", "comfortable", "please", "spacious", "not big", "complete" and "beautiful".

Visitors' perception of the "room" cluster is dominant. "clean" and "tidy" determine the attitude of visitors' comments. Visitors pay attention to the room's agreeableness and comfort, and have requirements on the quality of space and facilities. In a word, the "room" cluster presents rich perceptual colors, showing the delicacy and comprehensiveness of visitors' physical sensory experience.

3.3. Cluster Analysis of "Host"

"Host" represents the reception service of homelands. The behavior of the host: "luggage", "meal" and "attitude" are the key indicators to test the service. "Luggage" is the embodiment of the pick-up process, perceived as "help"; "Delicious", "palatable", "taste" of "food" is the direct embodiment of service ability; The "attitude" of the host (or waiter) is the temperature of service, which is perceived by visitors as welcome. The perceptual words of "host" cluster include: "warm", "considerate", "hospitable", "intimate", "cordial", "attentive" and "warm-hearted", indicating the influence of the host's appeal and insight on the perception of tourists.

4. Discussion

Based on the analysis of user feedback big data, the three dimensions of "host", "homestay" and "room" constitute the perception framework of homestay, and the network structure is formed with them as the center. "Homestay" and "room" are the perceptual dimensions of the physical environment, which are coupled with the accommodation behavior of tourists. "Host" is the perceived dimension of service environment.

The overall dimension of tourists' perception of homestay ("homestay") presents a rational color. Tourists take "homestay" ("accommodation") as the coordinate to establish connection with other five elements of tourism: food, travel, tour, entertainment and purchase. The location of homestay

affects the smoothness of tourism activities and thus the choice of tourists. The perception of location is closely related to the positioning needs of people. Increasing the convenience of location, such as the identifiable route to the homestay, the clear route to the parking lot, the striking name of the homestay and the recognizable facade, can improve the perception of the location of the homestay. The interior of the homestay is the space for tourists' accommodation activities. Tourists' "habitual" living behaviors are carried out in the "abnormal" homestay space, and the abnormal form is easy to attract tourists' attention. Tourists' perception of form is mainly reflected in co-occurrence nouns, which are characterized by rational cognition and rich in conceptualization. Therefore, recognizable symbols, classifiable styles and understandable programs appear in the design, which are easily perceived by tourists through subjective processing. However, the popular form of homestay is more easily recognized by tourists.

The specific dimension ("room") of tourists' perception of homestays presents perceptual color. Perceptual words show the characteristics of multi-sensory participation. The perception of vision, hearing, taste and touch makes the guest room directly related to the body experience. Guest rooms are private Spaces with clear boundaries and clear functions. Visitors like to have a clean and tidy environment when they have a rest in the guest room. Materials, colors and layout have a significant impact on the environment. The physical properties of the room, such as feeling of bright light, oppressive space, appropriate shape, better sense of layout, affect the body feeling, so design of room physical properties, such as change the window size, enlarge eye shot to improve the small room of oppressive feeling, divide the space to change the shape of the discomfort, are help to improve tourist experience. Color, vein, texture, light and other easy to stimulate the senses of the design elements, should be actively used in the room design. In addition, the relation of beautiful and courtyard in co-occurrence shows, the scenery of doorway, window causes beautiful feeling easily. Attractive design should be used for where that do not interfere with sleep, such as toilet, balcony, balcony.

Visitors perceive the service environment of homestay through the service ability and social ability of the "host". Host and family members (or employees) provide accommodation services that are not only part of the visitor experience, but also part of the local cultural environment. The host's personality charm and social skills are the key to improve the perceived quality of the homestay facility. The host's local experience, cultural taste and communication skills affect the connection between tourists and local knowledge. As readers and practitioners of local culture, hosts act as a bridge between tourists and places.

5. Conclusion

Transition of homestay facility design into products. Product design is controlled by the integrated process of design, production and consumption. As a tourism experience

product, homestay is also subject to the rule of production and consumption. With the expansion of capital operation of homestay facility industry, the product design features of brand management are increasingly significant. The transition to productization of homestay inevitably requires the application of product design and management methods across specialties, the establishment of homestay product system, the realization of overall planning, continuous homestay ovation and rapid update.

Construct the image structure system of homestay. The comments on the website reflect the perceptual image of the homestay. Users summarize the appearance into a rational image and reflect it through the corresponding relationship between language and form, showing the interaction process between users and products. The image structure formed by the clustering of user-perceived images, such as the branches of the homestay product tree, can bring forth homestay product with the changes of the consumer market. The homestay product design can be diversified and varied like a tree, to improve the efficiency of design innovation.

Big data analysis supports design decisions. In the era of knowledge economy, design homestay innovation is embodied in transforming the market and consumers' cognition into new products, and big data information analysis plays an important supporting role. The text analysis software is used to digitize the user feedback information, and the data analysis is realized with the help of artificial intelligence technology, to serve as the basis for the product design and positioning of the homestay facility and avoid the blindness of subjective experience. Specifically, the effective management of product design homestay innovation of homestay facility can be realized by creating sample database and perceptual word database of homestay facility product.

Acknowledgements

The author received funding from Jiangsu university philosophy and social science research project with the serial number 2016SJB760, and Innovation and entrepreneurship course construction project of Nanjing university of finance and economics, and teaching reform project of Nanjing university of finance and economics with the serial numbers JGY1655 and JGY1539, and Pre-research project of Nanjing university of finance and economics with the serial number ZW01.

References

- [1] L. L. Li, Y. Zhang, "Visual Criticism by Looking at Inside Objects and Outside Landscapes of Country Inns," *Zhuangshi*, vol. 2019, 4, pp. 22–27.
- [2] M. Erlhoff and T. Marshall, *Design Dictionary, Perspective on Design Terminology*. Wuhan: Huazhong University of Science and Technology Press, 2016, pp. 284.
- [3] S. S. Deng, *Research on text analysis methods for big data support decision making*. Shanghai: Zhongxi Book Company, 2017, pp. 172–177.
- [4] D. J. XU and L. ZHANG, "The Application of Text Mining in Social Science Research: Present Situation, Problems and Prospects," *Journal of Science and Society*, vol. 2015, 3, pp. 75-89.
- [5] C. X. Zhai and S. Massung, *Text Data Management and Analysis: A Practical Introduction to Information Retrieval and Text Mining*. Beijing: China Machine Press, 2019, pp. 142-147.
- [6] O. L. Fan, "Evaluation model construction of home stay facility from the perspective of customer perception: based on grounded theory," *Journal of Jiangsu Commercial Forum*, vol. 2011, 10, pp. 37–39.
- [7] S. Li, X. L. ZHOU, Y. DING and Y. C. WU, "Research on Experience Perception of Mainland Residents to TAIWAN Travel—Text Analysis of 98 travel notes from Mafengwo site," *Journal of Tourism Forum*, 2015, 6, pp. 7–20.
- [8] Y. X. LIU and S. B. YIN, "Analysis of the Tourist Experience of Family Hotel in Hongcun Village Based on Network Reviews," *Journal of Anhui University of Technology (Social Sciences)*, 2016, 6, pp. 44-46.
- [9] M. FANG, Y. S. XU and Y. ZHANG, "On the Development Strategies of the Linan Home Stay Based on Tourist Destination Image," *Journal of Central South University of Forestry & Technology (Social Sciences)*, 2016, 2, pp. 50-53.
- [10] J. Tong and J. S. Liu, "Study on Perception of Dalian B&B Experience Based on Web Text Analysis," *Journal of Yangling Vocation and Technology College*, vol. 2018, 4, pp. 1–6.
- [11] Q. ZHANG, H. ZH. SUN, Z. R. LAN and Q. Y. YU, "An Analysis of the Tourist Service Quality of Nanyuanling Guesthouses in Wuyi Mountain Based on Netw," *Journal of Xichang University (Social Science Edition)*, 2018, 2, pp. 73-76.
- [12] Y. Tang and X. B. Feng, "Research on Online Evaluation of Home Stay Service Quality Based on Content Analysis—Taking the Zhenmei Lige Inn in Lugu Lake as an Example," *Journal of Leshan Normal University*, 2019, 9, pp. 51-56.
- [13] T. T. Shi and Y. Q. Han, "Elements of Guest Experience in Metropolitan B&B-Analysis of Online Comments of Airbnb," *Journal of Pioneering with Science & Technology Monthly*, 2019, 6, pp. 39-42.
- [14] M. Young, *Content Analysis of Communication: Research and Applications*. Chongqing: Chongqing University Press, 2014, pp. 262.
- [15] S. P. Wang and L. Ling, *Study on structural characteristics of industrial symbiotic network*. Shanghai: Tongji University Press, 2012, pp. 75.