Research on the Development of Urban Tourism Resources Based on Text Mining—Taking Chengdu as an Example

Yulu FAN[a], Jinchun KOU [b,*]

School of Economics, Wuhan University of Technology, P.R. China

[a]18384225683@163.com, [b]kouwhut@163.com

*Corresponding author

Keywords: tourism resources, Chengdu, content analysis, text mining

Abstract. In order to promote the development of Chengdu tourism resources, this paper based on text mining, using content analysis method to conduct qualitative research on the network evaluation text of Chengdu tourism. The ROST CM6 software is used to perform high-frequency word analysis, semantic network analysis and sentiment analysis on the text, and the visitors' perception of Chengdu' s tourist destination is obtained. Combined with the characteristics of tourism resources in Chengdu, the paper puts forward some suggestions for improving the development of tourism resources to promote the full utilization of tourism resources in Chengdu.

1. Introduction

As the best tourist city in China, Chengdu is the birthplace of ancient civilizations. It has a history of more than 3,000 years. The long history and culture and the city atmosphere attract many tourists from all corners of the country. According to statistics, in 2018, Chengdu' s total tourism revenue reached 303.3 billion yuan, and the number of tourist receptions reached 217.01 million. The tourism industry has gradually become an important engine for promoting Chengdu' s economic development. Under the situation of the booming tourism industry in Chengdu, studying the image of Chengdu tourism destination and analyzing the development status of tourism resources play an important role in enhancing the image of urban tourism. Under the gradual development and popularization of the Internet, the impact of the network on tourism activities is increasing, and tourism has become a network trend [5]. The characteristics of freedom, openness and sharing of online travel can make it more realistically reflect the tourists' perception of the characteristics of the tourist destination, so that relevant information can be extracted from the travel notes and the development of tourism resources can be studied. The study of tourism destination image is one of the hot spots of tourism research at home and abroad in recent years [2-3]. William used the network information to compare the image of Seoul tourism with traditional print media [1]. Zhang Hong conducted a tourism development study on the characteristic neighborhoods of Beijing Nanluogu xiang based on the network text [4]. He Xing conducted research on the tourism development of Mount Wutai based on tourist satisfaction [5]. Zhang Gaojun and Li Junyi used Mount Hua as a case study to study the image of Mount Hua tourism destination using web logs [6]. In the past, the research on the use of network and content analysis methods focused on the tourism image of popular scenic spots, and did not analyze the overall tourism image of the city. The research case of this paper is to analyze the overall tourism image of Chengdu. By sorting out the network reviews of different platforms, the ROST CM6 software is used to analyze the tourism evaluation texts, and the characteristics of the whole city tourism resources are obtained, and relevant suggestions are put forward to promote the full utilization of urban tourism resources.

2. Data Sources and Research Methods

2.1 Data source

Online travel websites such as Ctrip and Qunar.com are the main medium for tourists to obtain information and share travel experiences. This study uses the travel guide in the online travel network as the source of text data, and uses the "Chengdu Tourism" as the key word to search the "Travel Guide" section of the website, and deletes the comments and travel notes on the web page that repeatedly appear and deviate significantly from the travel theme. Copy all the content into the word document for merging, delete the picture after a total of more than 15 million words, then pre-processed the text to generate analysis text, then process the word segmentation, merge, and generate the final document.

2.2 Research methods

This study mainly uses content analysis method to study the text data such as visitor network reviews and travel notes. The content analysis method is a research method for objective, systematic and quantitative description of the content of the dissemination [7]. Among them, ROST content mining is the most commonly used content analysis tool (referred to as ROST CM6), which filters out a large number of words unrelated to the research content, extracts high-frequency feature words related to the research, and performs semantic network analysis on the sample travels after the word segmentation. Its visual graphics can also complete a variety of research objectives such as sentiment analysis, comparative analysis, status analysis, and trend analysis. In this paper, the previously processed text is imported into ROST CM6 software, which analyzes the high-frequency words, semantic network and emotions in turn, and finally obtains the relevant characteristics of Chengdu tourism destination.

3. Results Analysis

3.1 High Frequency Word Analysis

In this paper, ROST CM6 software is used to segment and count the word frequency of the revised text, in order to extract high-frequency feature words, and obtain preliminary results. Meanwhile, virtual words such as "place", "various", "time" and "some" are included in the filter vocabulary, and irrelevant words are filtered. Finally, the first 48 high-frequency feature words are sorted out, as shown in Table 1. The highest word frequency is "Chengdu", which reflects the tourists' perception of the name of the tourist place is very strong; the frequency of "history" is also very high, indicating that Chengdu's rich historical and cultural heritage is deeply rooted in the hearts of the people; The frequency of attractions such as "Kuan Alley and Zhai Alley" , "Panda Base" , "Du Fu Cottage" , Tianfu "Square "and "Chunxi Road" is generally high, which is consistent with Chengdu' s external tourism propaganda. Highlight the main features of Chengdu tourism.The words "spicy" and "gourmet" have a very high frequency. Visitors have experienced a lot of special snacks in Chengdu, and hot pot has become a business card in Chengdu. In addition, among these high-frequency vocabulary, the number of nouns is the most, 38, mainly the place names and scenic spots and their various food nouns, indicating that Chengdu' s main positioning is leisure and entertainment, attractions and food as the main tourist attraction; followed by adjectives It is mainly the overall feelings and evaluations of tourists for Chengdu, including "characteristics", "good food", "spicy", "leisure" and "lively". It shows that most tourists think that life in Chengdu is comfortable and worth browsing; The minimum number of verbs is 2, with "service "and "experience" , mainly the way of tourists' activities, which indicates that the participation of Chengdu tourism activities is very insufficient and there is no experiential amusement project.

Table 1 The top 48 high-frequency vocabulary of Chengdu travel network review text

Number	Vocabulary	Word Frequency	Part of speech	Serial number	Vocabulary	Word Frequency	Part of speech
1	Chengdu	827	Noun	25	Night	80	Noun
2	History	470	Noun	26	Service	78	Verb
3	Hot Pot	264	Noun	27	Sense of design	77	Adjective
4	Taste	221	Noun	28	Beef	73	Noun
5	Kuan Alley and Zhai Alley	210	Noun	29	Restaurant	72	Noun
6	Variety of dishes	200	Noun	30	Bar	70	Noun
7	Delicious food	170	Noun	31	Ancient town	70	Noun
8	Characteristic	168	Adjective	32	Impression	70	Noun
9	Temple of Marquis	165	Noun	33	String	68	Noun
10	Chunxi Road	150	Noun	34	Daci temple	64	Noun
11	Panda	148	Noun	35	Metro	57	Noun
12	Yummy	132	Adjective	36	Snack	53	Noun
13	Du Fu Cottage	131	Noun	37	Natural	47	Adjective
14	Spicy	130	Adjective	38	Fresh	46	Adjective
15	The people's Park	128	Noun	39	Cheap	46	Adjective
16	Jinli Street	122	Noun	40	Zhimin Road	43	Noun
17	Tianfu Square	107	Noun	41	lively	42	Adjective
18	Museum	106	Noun	42	Famous	40	Adjective
19	Seafood	103	Noun	43	Ruins	38	Noun
20	Leisure time	92	Adjective	44	Dujiang Dam	37	Noun
21	Barbecue	86	Noun	45	Traffic	37	Noun
22	Architecture	85	Noun	46	Jiu Yan Bridge	37	Noun
23	Hotel	81	Noun	47	Bean jelly	36	Noun
24	Memory	80	Noun	48	Experience	33	Verb

3.2 Semantic Network Analysis

In this paper, using the social network and semantic analysis functions in ROST CM6 software, input the modified network comment text in the text box to be processed, and start the NET DRAW graphics tool to generate the Chengdu tourism semantic network analysis graph, as shown in Figure 1. To reflect the relationship between the meaning of the phrase and the overall structural relationship of the text. It can be seen from the network graph that "Chengdu", "hot pot" and "history" are three important nodes with high frequency

in the sample travel, and the closer the words are to the central nodes, the closer the relationship is to the nodes. Several more obvious conclusions can be drawn from the semantic analysis graph:

(1)The semantic network radiated by "Chengdu", in addition to Chengdu's most distinctive tourist attractions, also includes many cognition about Chengdu's tourism image, such as "natural", "leisure" and "features". Reflects the impression of tourists on Chengdu tourism.

(2) The semantic network radiated by the word "history" brings together many historical and cultural attractions such as "Temple of Marquis"and"Kuan Alley and Zhai Alley" , which reflects Chengdu's profound cultural heritage and rich cultural landscape, and is a powerful driving force to attract tourists.

(3)The semantic network radiated by the word "hot pot" is not only related to all kinds of food, but also contains a large number of taste words about "taste", "yummy", "restaurant" and "spicy", reflecting as a food capital, Chengdu has a good impression for tourists and has a good experience.
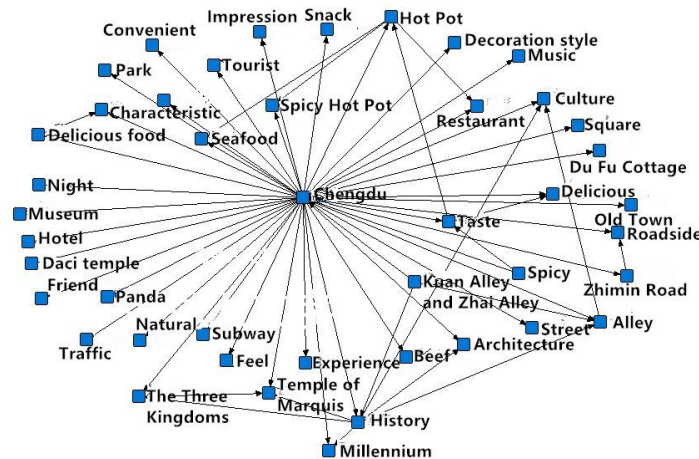

Fig.1 Chengdu tourism semantic network analysis.

3.3 Sentiment Analysis

Emotions have positive and negative points, as well as strength and weakness [8]. The sentiment analysis in the content analysis method can be used to analyze the tourist's emotions. It analyzes the emotional distribution of the tourists by quantitatively evaluating the emotional expressions in the tourist text information. This paper uses the sentiment analysis method to analyze the sentiment network text of Chengdu tourists, as shown in Table 2. Among the various types of emotions, positive emotions accounted for 49.33%, neutral emotions accounted for 36%, and negative emotions accounted for 14.67%.In the specific performance of various emotional intensity, the moderate level of positive emotions is dominant, while the general level of negative emotions is dominant. In general, the emotional elements of tourists' impressions of Chengdu's tourist destinations are mainly positive emotions, and there are also some neutral emotions and less negative emotions. Through research, the negative factors in the emotional analysis of tourists' texts are mainly: crowding and commercialization. In order to promote the effective development of tourism resources, the tourist attractions should appropriately rectify these problems.

Table 2 sentiment analysis table.

	Proportion	Strength	Proportion
Positive emotions：	49.33%	General strength（0—10）：	16.00%
		Moderate strength（10—20）：	20.00%
		High intensity（20 以上）：	13.33%
Neutral mood：	36.00%		
Negative emotions：	14.67%	General strength（0—10）：	12.00%
		Moderate strength（10—20）：	2.67%
		High intensity（20 以上）：	0.00%

4. Tourism Development Status and Research Summary

This study takes Chengdu as a case study, based on the network text, using the content analysis method, carries out high-frequency word analysis, semantic network analysis and sentiment analysis, and studies the tourists' perception of the image of Chengdu tourism and summarizes the tourism resources of Chengdu. The following conclusions are drawn:

(1) Through the analysis of high-frequency words, the names of various major attractions, various foods, and historical atmosphere have become the three most frequent words, which fully reflects the characteristics of Chengdu tourism: scenic spots + attractions history and culture + food. In terms of high-frequency words, nouns account for the majority, including a small number of adjectives, and fewer verbs, reflecting the lack of participation in Chengdu tourism.

(2) Through semantic network analysis, it is concluded that tourists have high perceptions of various scenic spots in Chengdu. Through a variety of scenic spots, a series of impressions about Chengdu, such as "leisure", "easy" and "characteristics" are reflected. The comfort brought by Chengdu life has become the label of Chengdu tourism; and Chengdu, as a historical and cultural city, has historical and cultural attractions such as "Temple of Marquis" and "Kuan Alley and Zhai Alley". rich Chengdu' s historical and cultural heritage also has tourists A strong attraction.

(3) In the tourist destinations of Chengdu, the emotional components are mainly positive emotions and neutral emotions. The proportion of negative emotions is small, and tourists are generally satisfied. However, the negative word-of-mouth brought by negative emotions may have a significant amplification effect. Therefore, in order to promote the effective development of tourism resources, the tourist attractions should effectively manage the negative word-of-mouth.

5. Countermeasures For Tourism Development

Through the above research and analysis, combined with the problems of high degree of commercialization and low participation of tourists in some scenic spots in the process of Chengdu tourism development, this paper puts forward some reasonable suggestions on improving the development of tourism resources in Chengdu.

(1) Controlling the degree of commercialization and creating excellent tourist products: In order to expand the scale of some popular tourist attractions, modern commerce, entertainment and catering are integrated into them on the basis of the original scenic features. Although it seems more diversified, it ignores the preservation and excavation of the original flavor of the scenic spots, resulting in heavy commercial flavor and people. The literary atmosphere is insufficient. Such as "Kuan Alley and Zhai Alley", "Jinli Street "and other ancient streets and ancient towns have a homogenization trend, and are filled with the same snacks and shops. For the long-term development of tourism resources, each ancient town scenic spot should establish a long-term supervision mechanism to review the businesses entering the scenic spot, gradually phase out the

low-end business, promote industrial transformation and upgrade, avoid the same shop and create special products. In terms of catering, it is necessary to appropriately reduce the proportion of snack shops, to make reasonable pricing, and to increase snacks with Chengdu characteristics such as "Dragon reading hands", "Zhong Dumpling" and "Sichuan noodles". Other tourist souvenir shops such as postcards and bookmarks should reflect local characteristics and add features such as "Panda"and "Sichuan Opera", so that visitors can truly feel the cultural atmosphere and heritage of the scenic spot and enhance the recognition of the commercial symbols of the scenic spots.

(2) Enrich tourism activities and increase participation of tourists: Chengdu has abundant tourism resources, but at present, the main mode of tourism is to visit scenic spots and appreciate them. The types of activities are relatively single. Tourists stay for a very short period of time. It is difficult to understand the historical and cultural connotations of scenic spots profoundly. With the rapid development of tourism today, tourists are no longer just satisfied with the traditional way of appreciation and guided tour. The design and development of Chengdu tourism products should be from a single and one-way "static sightseeing" to a two-way "Experience the tour and the "deep tour" transformation. While maintaining the advantages of rich scenic spots and delicious food in Chengdu, we can also rely on the advantages of rich history and culture to plan and participate in highly experiential tourism activities, so as to realize the upgrading of historical and cultural tourism to comfortable experience tourism and the transformation of traditional tourism to fashion tourism.

(3) Improve the business card of Chengdu Tourism City: Chengdu is the birthplace of ancient Shu civilization, one of the ten ancient capitals of China. Well-known tourist attractions such as Kuan Alley and Zhai Alley,Temple of Marquis and Du Fu Cottage have given important cultural connotations and enhanced the attractiveness of Chengdu' s tourist landscape.In addition, Chengdu is situated in the panda protection base, and the image of the panda has become a business card for Chengdu tourism. However, it may be monotonous to confine itself to intra-city tourism. We can combine the surrounding tourist destinations such as "Mount Emei", "Mount Qingcheng" and "Scenic Spot of Daocheng Yading" to develop a number of special tourist routes and gradually improve Chengdu' s tourist business cards by improving the degree of convenience and creating diversification and diversification of tourism.

References

[1]Hunter W C, The social construction of tourism online destination image: A comparative semiotic analysis of the visual representation of Seoul, J. Tourism Management, 2016, 54(2): 221-229.

[2] Asuncion BeerliJ, osefa D Martin, Factors Influencing Destination Image, J. Annals of Tourism Research, 2004, 31(3) : 657 -681.

[3]Chew E Y T, Jahan S A, Destination image as a mediator between perceived risks and revisit intentions: A case of post-di-saster Japan, Tourism Management, 2014, 40(2): 382-393.

[4]Hong Zhang, Tingting Shi , Mengxue Yu , Research on Tourism Development of Characteristic Streets Based on Network Text Analysis——Taking Beijing Nanluoguxiang as an Example, J. Journal of Anhui Agricultural University (Social Science Edition), 2019,28(01): 49-54.

[5]Xing He, Research on Tourism Development of Mount Wutai Based on Tourist Satisfaction, J. Commercial Economy, 2019(09):63-64.

[6] Gaojun Zhang, Junwei Li, Research on Tourism Image Perception of Mount Hua Scenic Spot——Based on Text Analysis of Visitors' Web Log , J. Tourism Science, 2011, 25(4) : 87 - 94.

[7]Zhanlei Yao , Xin Xu , Analysis of Co-occurrence Phenomenon in Scenic Spots——Taking the First National 5A Tourist Attractions in East China as an Example, J. Tourism Science, 2011, 25(2): 40-47.

[8]Yongming Wang , Meixia Wang, Rui Li, Dianting Wu , Study on the image perception of The Old Town of Phoenix tourism based on network text content analysis, J. Geography and Geo-Information Science,2015,31(01):64-67.