

# Research on the Promotion of Word of Mouth in Tourist Scenic Spots Based on Web Text Mining——the Case Study of Wanlu Valley in Guangdong Province

ZHANG Ying

Heyuan Polytechnic, Faculty of Business Administration, Heyuan Polytechnic, Heyuan, Guangdong, China

**Abstract.** Under the background of Internet economy and sharing economy, tourist scenic spots should pay more attention to tourists' network public opinion and do a good job in cultivating network word of mouth. Taking Wanlu Valley ecotourism area in Guangdong province as an example, the paper collects Baidu index and uses ROST Content Mining software to excavate the post-consumer evaluation text of five tourist websites, such as Tongcheng, Ctrip, Grasshopper's Honeycomb, Meituan, Qunar, etc. By mining the high-frequency characteristic words of the tourist evaluation text, constructing the social semantic network matrix map, and then synthetically analyzing the tourist network attention index and the tourists' evaluation perception information, the result demonstrate that the characteristics of scenic spots, service attitude and tourist facilities are the focuses of tourist evaluation: the number of high-frequency words is large and the degree of praise is high. Therefore, the scenic spots should pay attention to the integration development of "tourism +" industry, improve service quality, enrich tourism experience projects, promote the industrial transformation and update and innovation development of eco-tourism destination.

## 1 Introduction

China Internet Network Information Center released the latest data, showing that the number of Chinese netizens has reached 772 million, the number of mobile Internet users reached 753 million and the Internet penetration rate reached 55.8% with the booming of sharing economy. With the improvement of Internet technology and the subdivision of Internet domains, the proportion of online shopping, network entertainment, tourism e-commerce, etc. is increasing day by day. "Internet +'s" overall service platform for tourism consumption, such as Tongcheng, Ctrip, Grasshopper's Honeycomb, Meituan, Qunar, etc. all focus on the core needs of most consumers and are committed to meeting the traveling consumer behavior such as tourist tickets, hotels, air tickets, car rental, etc. The iteration of "Internet +" to tourism consumption has greatly promoted the dissemination of tourism destination image and made the representation of tourism image become increasingly complex.[1]Any vested interests based on asymmetric information will be eliminated by the internet[2], which is a great challenge to the marketing and management of the word of mouth of tourism enterprises.

The rise of collaborative consumption business in the Internet era has led to the rapid development of online word-of-mouth models such as online tourism consumption, sharing, experience evaluation, etc., and has aroused the widespread concern from the consumers, enterprises and researchers. Tourist attractions are

service reception-oriented enterprises, with the prominent characteristics such as labor intensive, the great service individual variations, etc. and its service quality directly determines the reputation of enterprises and product and market share. Therefore, the scenic spots should pay attention to the study of network word-of-mouth, abandon the traditional bad management practice, take the core value and best interests of tourist experience as the starting point, enrich the service items of tourist attractions, improve the service quality and set up network word-of-mouth and industry integrity, so as to achieve the transformation innovation and development of tourist attractions.

## 2 Research on tourist areas

Guangdong Wanlu Valley leisure resort area is located in Dongxing Village, Xinghilong Town, Dongyuan County. Its geographical location is longitude 11433 "E" and latitude 2342 "30" north, Facing the beautiful Wanlu Lake, it is 15 kilometers away from Guishan Scenic spot and 33 kilometers from the urban area. The scenic spot is located at the nodes of the Ring Wanlu Lake Tour and Pan Wanlu Lake Tour and the location advantages are prominent. Meanwhile, the gorge being deep and secluded, the cliffs being on both sides of the bank, the flying waterfall in the air being long, the resources advantage is obvious. Since its opening on December 27th of 2007, Wanlu Valley has focused on nature sightseeing and has built and operated holiday resorts

brands in line with the concept of environmental protection and catering to the Hakka style of Baizi Wai and Yujia Water Village with high standards and high starting point planning. Having obtained a place to survive and become a tourism brand in Wanlu Lake area, it is also a bright-spot project of the transformation and reconstruction of ecological tourist attractions.

At present, Guangdong Wanlu Valley leisure resort area has become a popular ecological leisure-time tourist attraction in Hong Kong, Macao and Pearl River Delta. Since the opening of Wanlu Valley, it has experienced three stages: development period, leap-forward development period and steady development period. Using Baidu index to analyze the trend value of the network attention of Wanlu Valley, it can be seen that the attention degree of the network has steadily increased and the average trend value is 115 from 2011 to 2018. (see Table 1).

**Table 1.** "Guangdong Green Valley Leisure Tourist Area" Network Attention Analysis

| Names of Index      | Observation Time Domain  | Trend Value |
|---------------------|--------------------------|-------------|
| Overall Trend       | 2011                     | 21          |
|                     | 2012                     | 90          |
|                     | 2013                     | 120         |
|                     | 2014                     | 133         |
|                     | 2015                     | 118         |
|                     | 2016                     | 131         |
|                     | 2017                     | 115         |
|                     | January to March of 2018 | 59          |
| Average Trend Value | 115                      |             |

Data sources: Baidu Index January 1st, 2011 to March 13th, 2018.

### 3 Research methods and data sources

Ecotourism can be classified as experiential tourism and its tourism aesthetics can be divided into three stages: pre-tourism expectation, in-tour experience and post-tourism recall.[3]No matter at which stage, the tourists' participation and experience provided by the scenic area and tourists' sentiment of happiness from the activities have a significant impact on the formation of good word-of-mouth effect. Though without the prominent location advantages, the maximization of resource characteristics, experience aesthetics and service efficiency of the ecotourism destination, especially the focus on the collection of network propaganda and network evaluation, are of critical vitality for the sustainable management and development of a scenic area. Among the unstructured data mining methods such as tourism public opinion and network public opinion, the content mining system software- ROST Content Mining developed by Professor Shen Yang is very appropriate with favorable effects.

By using ROST Content Mining software and taking the period from January 1st 2015 to March 10th 2018 as

the time node, excavating the visitor experience on Tongcheng, Ctrip, Mafengwo, Meituan and Qunar, the paper has made a thorough analysis of the evaluation texts of Guangdong Wanlu Valley Tourist Destination and a total of 333 valid samples were received (see Table 2). Based on the effective sample, the text is comprehensively analyzed, such as word segmentation, high word frequency statistics, semantic network and social network. From the perspective of qualitative and quantitative analysis, this paper studies the effect of tourist experience, perception of word-of-mouth, and the current situation and trend of network public opinion in scenic area on the basis of qualitative and quantitative analysis so as to offer the specific suggestion on its brand building, project promotion, service quality optimization, etc.

**Table 2.** "Guangdong Green Valley Leisure Tourist Area"Network Evaluation Text Collection Statistics

| Network Names | Sample Numbers |
|---------------|----------------|
| Tongcheng     | 242            |
| Ctrip         | 61             |
| Mafengwo      | 15             |
| Meituan       | 8              |
| Qunar         | 6              |
| Total         | 333            |

Statistical time: January 1st 2015 to March 10th 2018

## 4 Research procedures

### 4.1 High-frequency Word Comparison Analysis

This paper summarizes 333 evaluation texts by excavating five tourism websites and puts them together as "Tourist Experience Evaluation of Wanlu Valley Leisure Resort Area in Guangdong .txt", as well as imports them into ROST Content Mining software for word frequency analysis. The filter table is used to filter the words which are not related to the research subject and have no obvious representative meaning. After filtering, 70 effective high-frequency words such as "drift", "scenic spot", "very good" and "air" are formed(see Table 3).

**Table 3.** "Guangdong Green Valley Leisure Tourist Area"Visitors evaluate major high frequency feature words

| Tourist evaluation feature words | frequency | Tourist evaluation feature words | frequency | Tourist evaluation feature words | frequency |
|----------------------------------|-----------|----------------------------------|-----------|----------------------------------|-----------|
| Drifting                         | 77        | Hour                             | 16        | Nice environment                 | 9         |
| Scenic area                      | 40        | Tongcheng                        | 16        | All the way                      | 9         |
| Well                             | 40        | Project                          | 16        | Drive                            | 9         |
| Air                              | 39        | Original ecology                 | 15        | Uphill                           | 9         |
| Surroundings                     | 38        | A bit                            | 15        | Very suitable                    | 9         |
| Inside                           | 35        | Feel                             | 15        | Beautiful                        | 9         |

|               |    |            |    |               |   |
|---------------|----|------------|----|---------------|---|
|               |    |            |    | scenery       |   |
| Green Valley  | 33 | Casual     | 15 | Overall       | 9 |
| Worth it      | 31 | View       | 15 | Not much      | 9 |
| Fun           | 30 | Green Lake | 14 | Refreshing    | 9 |
| Tickets       | 30 | Service    | 13 | Along the way | 9 |
| Attractions   | 29 | Kilometers | 13 | Pretty        | 9 |
| Child         | 24 | High speed | 12 | Go            | 9 |
| Fresh air     | 23 | Very happy | 12 | Okay          | 8 |
| Suitable for  | 22 | Natural    | 11 | Creek         | 8 |
| landscape     | 22 | Comparing  | 11 | Like          | 8 |
| Happy         | 22 | Summer     | 11 | Super         | 8 |
| play          | 21 | Next time  | 11 | It's nothing  | 8 |
| Acceptable    | 20 | Ideas      | 10 | A lot of fun  | 8 |
| Mountain Road | 19 | Swing      | 10 | Far           | 8 |
| Facility      | 16 | Children   | 10 | Go again      | 8 |

#### 4.2 The Analysis of Classification of High Frequency Words and Social Semantics

In the age of the Internet, before booking for hot spring tourism consumption, tourists always use its official websites, online travel platforms, travelogues and other channels to obtain a thorough knowledge of the facilities, catering services, price levels, transportation guides, and other important information of the scenic spots, then form the important psychological expectation, and finally determine the purchase behavior. In order to better analyze the types of tourist experience evaluation in Wanlu Valley leisure and vacation area in Guangdong Province, 70 high frequency feature words are evaluated on network. It is classified and summarized from six kinds of perceptual objects, such as tourism attraction, tourism environment atmosphere, tourism scenic spot culture, tourism experience, tourism propaganda and recommendation, tourism resource evaluation, etc.(see Table 4).Through comparative analysis of categories and items, we can get a clearer view of the characteristics of tourism resources, cultural characteristics of scenic spots, tourism service level and other key public praise evaluation situation.

**Table 4.** "Guangdong Wanlu Valley Leisure Tourist Area" Tourist Perception Information Analysis

| Perceived Object           | High-frequency word Quantity | Perceived Object  | High-frequency word Quantity |
|----------------------------|------------------------------|-------------------|------------------------------|
| Ecological Environment     | 9                            | Status of Charges | 9                            |
| Tourist Facilities         | 12                           | Service Attitude  | 12                           |
| Scenic Spot Characteristic | 20                           | Cultural Customs  | 8                            |

In order to further study the situation of tourists' perceived objects, by analyzing the high frequency characteristic words of tourists' evaluation, the paper constructs the social semantic network matrix map of

network text in Wanlu valley leisure resort area of Guangdong province. It can be seen from the picture that high frequency words such as "drift", "scenic spot", "very good" and "air" are important nodes. According to the distribution rules of the social semantic network matrix map, the relationship between the high-frequency words and the core high-frequency words depends on the distance they are adjacent: closer the distance is, more intimate the relationship is. The result demonstrate that as the representative of ecotourism destination in northern Guangdong, its scenic characteristics, service attitude and tourist facilities are the focal points of tourist experience.

#### 4.3 An Analysis of Affective Tendency after Tourism Experience

The emotional attitude of online texts mainly focuses on textual data and analyzes emotionally-sensitive nouns and adjectives in tourist evaluations to obtain visitors' perception experience and visual impressions on tourist attractions and destinations. Valley Web Text performs "sentiment analysis." In the analysis table, we can see that the proportion of positive emotions of tourists is 73.57%, neutral emotion is 5.4%, and negative emotion is 21.03%. It can be seen that tourists have a high degree of praise, and the proportion of neutral and negative emotions is relatively small (see Table 5).

**Table5.** "Guangdong Wanlu Valley Leisure Tourist Area" Travel Experience Evaluation Sentiment Analysis

| Emotion category  | Number of words (articles) | Proportion(%) |
|-------------------|----------------------------|---------------|
| Positive emotions | 245                        | 73.57         |
| Neutral emotion   | 18                         | 5.4           |
| Negative emotion  | 70                         | 21.03         |
| Total             | 333                        | 100           |

### 5 Conclusion

As a typical representative of ecological and leisure tourist destinations in Guangdong Province, the Wanlu Valley Leisure and Holiday Tourist Area has attracted the attention of tourists on the network for its scenic features, service attitude, tourist facilities, etc. The scenic spot needs to focus on the tourist experience effect, perception of word of mouth, network public opinion status and trends. Through thematic sentiment analysis of web text data, the results showed that positive emotions accounted for 73.57%, neutral emotions accounted for 5.4% and negative emotions accounted for 21.03% after the tourists travel there. Taking Ctrip as an

example, by analyzing the evaluation of reputation of tourism experience, the following is drawn out: tourists think highly of the place: over 96% of the tourists render the acknowledgement and recommendation and the average score is 4.5 points (5 Points), which all reflects the scenic spot in the hardware and the software comprehensive operation has obtained the visitor's acknowledgement. The satisfaction degree of service behavior in tourist attractions consists of complex elements, including the construction of natural environment and the construction of management environment. [4]. "Internet +" has promoted the reform of the operation mode of tourist attractions and the sharing economy has become a new trend in the development of tourism economy. It has also increasingly influenced the word-of-mouth effect, brand shaping and innovative development of tourist attractions. In the age of experience economy, the scenic spot marketing needs to satisfy consumers' desires and increase customer experience [5]. The high frequency words of negative evaluation of Wanlu Valley reveal that the scenic area has been running almost for a decade. Therefore, some facilities need to be maintained and updated and the experience project of scenic spot needs to be supplemented so as to ensure the quality of participatory and interactive facilities for tourists, for the purpose of improving the network reputation.

## Acknowledgments

The author would like to thank support of 2017 Guangdong education department scientific research platform and project - youth innovation talent category (project no. : 2017GWQNCX061), Philosophy and Social Sciences in Heyuan, "13th Five-Year" planning project (No. : HYSK17P06), Heyuan Polytechnic research project (No. : 2017\_sk03).

## References

1. Ma Qiufang. The Content Analysis of Tourist Destination Media Symbols: A Case Study of Shaanxi Province. *Tourism Science*, 2011, 25(3): 35-44.
2. Hou Denghua. The Legal Status of Network Platform under Sharing Economy. *Forum on Political Law: Journal of China University of Political Science and Law*, 2017(1)157-164.
3. Zheng Huawei. A Study on the Internet Text of the Internalization of Red Tourism Value--Also on the Formation Mechanism of National Well Being. *Tourism Journal*, 2016, 31(5): 111-118.
4. Ryan, Li Meizhen, Wang Lin. *Research methods of Tourism Science: Research on Tourist Satisfaction*. Beijing: Tourism Education Press, 2012.
5. Li Zhifei. *Tourism Purchase Behavior: the Situation of Different Places, Experience Marketing and Post Purchase Effect*. Wuhan: Huazhong University of Science and Technology Press, 2009.