A Research on Tourist Satisfaction of Mountain Resorts Based on Network-Text Hybrid Analysis: A Case Study Of Oct East Shenzhen

Ke Zhang
Shenzhen Tourism College of Jinan University, China
Mu Zhang
JNU-UF International Joint Laboratory on IT& Tourism
Shenzhen Tourism College of Jinan University, China

Abstract

This paper studied the satisfaction of tourists in mountainous scenic spots from the perspective of smart tourism, focusing on the changes brought by the application of information technology. The author collected the network review of the OCT East Shenzhen as a data source, applied the method of comparative analysis of reward and punishment, combined with quantitative and qualitative network text analysis, identified the important factors affecting the satisfaction of tourists in mountainous scenic spots, and analyzed 11 scenic information services attributes, which are categorized into three types; the order of improvement of service attributes, which maximizes visitor satisfaction at the lowest cost, is discussed. The research results show that: (1) The "basic factors" and "performance factors" that provide the level of information services for mountain-type scenic spots that meet the expectations of tourists should be ranked first in the scenic information service improvement plan. (2) The optimization of the "performance factor" and "excitement factor" for providing the most satisfactory service for tourists is an optional improvement based on the budget and development strategy of the scenic spot. (3) According to the improved priority level, mountain-type scenic spots should concentrate on limited resources to enhance various attributes, thereby improving overall tourist satisfaction.

Keywords: Tourist satisfaction; content analysis; comparative analysis of rewards and punishments; Network review; Overseas Chinese Town East

Introduction

In recent years, tourist destinations and many scenic spots have tried to explore the innovation and application of tourism information technology. The informationization and intelligence of the scenic spots have changed the

function and structure of tourism, and have also brought about new changes in tourists' behavior and needs (Gretzel U,2011). Tourists in the information age rely on the Internet, self-service and travel agency booking software. They focus not only on the traditional "clothing, food, travel and entertainment", but also on the convenience and practicality of access to information and intelligent services. This provides a new idea of development and a certain challenge on how best to meet the needs of tourists. In fact, since the emergence of the concept of smart tourism in 2010, it has brought some help to the upgrading of the scenic spot. Building a smart tourist attraction that will satisfy tourists is still a long way to go (Wu B ,2015).

Tourist satisfaction has always been an important indicator to measure the difference between the service level of scenic spots and the expectation of customer demand. It is also one of the research hotspots in academic circles at home and abroad; however, the research content is mostly based on the perspective of the traditional tourism market, using a single quantitative or qualitative method to analyze the factors or model construction for the overall service quality or tourist satisfaction factors of the scenic spot (Huang D, Chen F,2015), there is less concern about the satisfaction of tourists after the promotion of scenic information services from the perspective of emerging smart tourism. In addition, in the face of increasingly competitive domestic and international tourism market, how can the operations managers of the scenic area, with a limited budget and resources, implement the construction of scenic information services and optimize decision-making order, and effectively integrate the construction of scenic spots into the smart tourism structure from the perspective of service and technological innovation? In the smart tourism structure, maximizing the satisfaction of tourists has become one of the important topics to improve the core competitiveness of the scenic spot.

Literature review

There are many research results on the satisfaction of tourists in scenic spots in the world. The main areas focus on the relationship between tourists' satisfaction and the factors affecting tourists' satisfaction based on regional and cultural differences. Yu and Goulden (1989) based on the Brayley model's satisfaction with international tourists, including analysis of tourist attractions, prices, services, facilities and destination images, found that due to regional and cultural differences, tourists' evaluation of historical attractions, staff and service quality was significantly correlated with the background of their own regions (Yu L, Goulden M,2006). Truong and King (2009) pointed out that in tourism marketing, not only the geographical difference will lead to differences in tourist satisfaction, but also having different languages and

cultural backgrounds will have the same effect. Some scholars have also conducted empirical research on tourist satisfaction based on specific scenic spots or tourist destinations, such as research on evaluations of sunny beach resorts (Assaker G, Hallak R, 2012), ski resorts (Bonnefoy-Claudet L, Ghantous N,2013), theme parks (Pearce PL, Wu MY,2016), pilgrimage sites (Krešić D, Mikulić J, Miličević K,2013) and tourist shopping areas (Murphy L, Moscardo G, Benckendorff P, et al,2011),etc.

On the theoretical basis and research methods of tourist satisfaction in scenic spots, Johnson & Gustafsson (2006) conducts a three-factor structural analysis of service attributes such as visitor satisfaction and dissatisfaction in service quality. Simith and Olson (2001) used factor analysis to point out the critical role that scenic services play in tourist satisfaction. Master and Prideaux (2000) also pointed out the link between the quality of service and the success of international destinations when studying the satisfaction of Taiwanese visitors to Queensland, Australia.

In recent years, Chinese scholars have made great achievements in theoretical innovation and empirical analysis on tourist satisfaction research in scenic spots. The theoretical innovation level mainly focuses on the evaluation system construction, measurement methods and theoretical cross-border innovation; while the empirical research focuses on the specific factors of the specific scenic spot satisfaction analysis or indicator (system) creation. Factor analysis and IPA analysis are commonly used research methods for domestic researchers to construct tourist satisfaction evaluation systems in scenic spots. The former is mostly used to select common factors that have a significant impact on tourist satisfaction through regression analysis, while the latter is often used to improve the multi-quadrant optimization based on the case of the evaluation system (Yang W, Hu X,2015; Yu Y,(2014); Dong J, Deng W,2014). With the rise of innovation and cross-border thinking in academic research, structural revisions and innovations based on the above two methods have become popular topics for many scholars. For example, Xia Xiaoyun et al. (2012) used the Fuzzy-IPA method to explore the factors affecting the tourist satisfaction of mountain-type scenic spots; Zhang Huali et al.(2014) introduced the logical model of econometrics, the "Ferner logic model" into the study of tourist satisfaction in scenic spots. Some scholars also used the Philip Kotler evaluation model to compare the pre-tourism expectations of the Haizhou Guandi Temple and the actual perception after the tour to calculate the satisfaction index (Dong J, Yan Y,2014); similarly, the psychological contract (Qin X,2014) and gray correlation analysis (Liu Y, Li W, Liu H, Li X, Zeng Z,2014) are also involved in the study of tourist satisfaction in scenic spots.

The case study of tourist satisfaction in scenic spots is also one of the key areas

of concern for Chinese scholars. In the study of tourist satisfaction for a specific type of scenic spots, many research results have been made in the two scenic spots of red tourist scenic spots and forest parks in recent years. The former's research focuses on the discussion of tourist satisfaction and evaluation improvement programs in traditional scenic spots (Shitai JI, 2012; Wang Z, 2013; Zhao W, Sun J, Song H, Xiong Q,2014), while the latter's research dimensions range from eco-tourism perspectives to world heritage sites (Yang R, Liu Y,2015). Other studies on the satisfaction of special scenic spots mainly include: smart scenic spots (Ge J, Gu X, Long Y, 2012), low carbon scenic spots (Gao H, Ren P,2012), mountain-type scenic spots (Yang Y, Li L, Li Q,2015) and cultural tourist attractions (Lü L, Li X,2015), which are most of the main scenic spots in China.

Although the study of tourist satisfaction at scenic spots at home and abroad has been quite large, both are less deep in the discussion about the satisfaction of tourists in the information age and in the context of smart tourism, but they focus more on the construction of theory and models. Tourist's collection of scenic service evaluation data is still use traditional questionnaires and expert interviews as the main method. At the research method level, quantitative analysis methods such as statistics and econometrics are used, or qualitative analysis methods such as psychology's perception, expectation and loyalty, and motivation and intention of consumer behavior are used. Combined with quantitative and qualitative methods, it is rare to conduct more comprehensive and complementary research methods for tourist satisfaction in scenic spots. Therefore, based on the research methods of mainstream tourist satisfaction, a sample analysis of tourist satisfaction of scenic information services from both quantitative and qualitative perspectives will be a useful attempt in this research area.

Therefore, along the above-mentioned literature review and comparative summary ideas, this research will be based on the mainstream research method of the three-factor theory: the reward and punishment comparative analysis method (hereinafter referred to as PRCA), and using the mixed content analysis method combining quantitative method and qualitative method. The content analysis method mines and analyzes the scenic spot information service network comments that can reflect the tourists' true perception in the information age. The focus of attention includes (1) screening the information service attribute factors that have a significant impact on the satisfaction of tourists in mountain-type scenic spots; (2) determining what kind of factors belong to it and the extent of the impact through the PRCA dummy variable model; (3) Finalizing the corresponding improvements, and establishing limited order and related optimization schemes. The above research ideas have theoretical and practical significance

in both research methods and research perspectives to supplement the satisfaction of tourists in mountainous and scenic areas at home and abroad.

Research design

Theoretical basis and research methods

(1) Three-factor theory of customer satisfaction

The study of the three-factor theory began with Kano et al. Martzler and Sauerwein subsequently refined the definition of the three-factor theory. According to the linear correlation and asymmetric relationship between the service attribute performance level and the customer satisfaction, service attributes are classified into "basic factors", "performance factors" and "exciting factors". The attributes in the "fundamental factors" are considered as the basic service functions that need to be provided first. Although a higher level of performance does not lead to the same level of customer satisfaction, service failures or trends will lead to strong dissatisfaction. It is the "performance factor" that the service attribute performance level is positively linear with the overall satisfaction. Among the three factors that can have a decisive influence on customer satisfaction is the "exciting factor". Its high performance will lead to higher customer satisfaction; however, since the average customer has no high expectations, even if there is low performance, it will not cause dissatisfaction.

(2) Network text hybrid analysis

Network text is essentially a complex and large amount of data. A single quantitative or qualitative analysis cannot analyze it thoroughly. For example, in high-frequency word analysis, each keyword generated by quantitative analysis is represented by default as a separate service attribute, but in reality, a keyword of a service attribute can have multiple expressions. For example, the word "online" in the comment may mean "online information service" or "network and communication service". Therefore, after quantitative factor analysis, a qualitative factor interpretation process is required to confirm the internal consistency of the factor corresponding to the same service attribute.

Factor analysis extracts the generalized service attributes, which can show the relationship between service attribute performance and visitor satisfaction from the data; however, it is difficult to understand the specific service content encountered in the corresponding actual journey, and the qualitative content analysis method is used to explain and summarize each service item in the generalized service attribute, which has a guiding role for the subsequent actual scenic area improvement measures.

(3) Comparative analysis of rewards and punishments

PRCA is a common research method in the three-factor theory. It has been more mature and widely used in customer satisfaction research. Compared with other research methods, it has high reliability and corrects the defects that the perspective is not comprehensive enough.

The traditional PRCA data collection comes from questionnaires and expert interviews, which may have subjective understanding and error of the researcher. The tourists evaluate the service attributes in an environment that is set in advance and has a conclusion-oriented environment. While this study takes visitors' online reviews as the research object, and the research samples are authentic and reliable. The purpose of this study is to explore the true evaluation of the tourists' aspects of the scenic spot service.

In summary, unstructured real-time visitor data is difficult to capture by traditional questionnaires or existing scale-based research methods. Therefore, it is necessary to use network text mining technology to analyze the tourist area information service attributes that tourists care about. On this basis, the evaluation of tourists' satisfaction and the optimization of service attributes are more practical. Also, different factors have different influence on the scenic spot, and cannot be promoted by the maximal release method. Instead, based on the reasons of time and cost, the order of the attributes that need to be improved is found. Among them, PRCA is the main research method to help the scenic spot to judge the order of service attributes improvement.

Model selection

Tourist's overall satisfaction dummy variable regression model

As a commonly used three-factor structure research method, PRCA supplements the traditional three-factor theory to calculate the defects of each attribute frequency separately, which provides a more comprehensive research perspective for the study of tourist satisfaction. In essence, PRCA encodes customer-perceived performance evaluations into two sets of dummy variables, X_j+ stands for high performance, X_j- stands for low performance, and these two variables are used as independent variables, while a two-category regression analysis was performed using customer satisfaction as a dependent variable. By judging the regression coefficient corresponding to the high and low performance levels, the reward coefficient $\beta+$ and the penalty coefficient $\beta-$, the attribution of each service attribute in the three factors is determined: $|\beta-|>|\beta+|$ is marked as "basic factor".$|\beta+|>|\beta-|$ is classified as "exciting factor"; when two indicators are similar, it is recorded as "performance factor".

$$\text{Overall Satisfaction (OS)} = a_0 + \sum \left(\beta_+ X_{i+} + \beta_- X_{i-} \right)$$

Methodology

Step 1: Excavate the network text, extract the high-frequency feature words related to the scenic information service and its frequency, structure the unstructured comments, and verify the basic dimensions of the service attributes through dimensionality reduction, condensing and exploratory factor analysis and verification. This study is based on the online review of the OCT East Shenzhen as the research object. Firstly, it extracts the service attributes that have significant influence on the scenic information service by quantitative analysis methods such as online commentary on high-frequency feature word statistical analysis and principal component analysis.

Step 2: Based on the original review, the above structured processing data is subjected to secondary verification and factor interpretation through qualitative content analysis. Then, based on the scores of the selected service attributes in its original comments, a high (low) performance level frequency satisfaction distribution table was prepared, and the qualitative content analysis and summarization of the factors were carried out by commenting on the specific service scenarios and details involved in each attribute. If the details of the scenic spot information service after the qualitative content analysis can be applied to the structure and content of the scenic spot information service attribute extracted by the factor analysis, the reliability of the first step factor analysis is verified again.

Step 3: According to the results of factor analysis, the binary regression analysis is carried out and classified into factors. Also, according to the classification results, the overall satisfaction degree is maximized, and the order of optimization and improvement of service attributes is determined. PRCA is performed on the service attributes after quantitative and qualitative analysis. On the one hand, by judging the relationship between the regression coefficients of the two dummy variables, it is judged which satisfaction factors are involved in the information service attributes of the scenic spots. Also, the ratio of the high (low) performance score of a service attribute in the original comment to the total frequency is also compared from another perspective to verify whether the two results match; On the other hand, according to the size of the reward coefficient and the penalty coefficient, the priority level of the scenic spot information service attribute improvement is determined, and the

limited resources are concentrated to improve each attribute in an orderly manner, thereby improving the overall tourist satisfaction.

Study area

This paper selects Shenzhen Overseas Chinese Town as a case study. OCT East Shenzhen Tourist Resort is one of the first "national tourist resorts" issued by the National Tourism Administration. During the 7th National Day holiday in 2018, it has received more than 160,000 tourists, becoming a "Shenzhen special tourist business card". Since the concept of "smart tourism" was put forward in 2010, the OCT East Shenzhen has carried out many upgrades from online portals, official Weibo, WeChat, scenic APP, to in-house information and intelligent services. They have also created a full-service visitor experience from ticket purchases, transportation, itinerary and travel guides. Exploring the changes in tourist satisfaction and the influencing factors after the upgrade of the information service of the scenic spot can not only measure the effectiveness of existing information services, but also concentrate resources on further building a national-level tourist resort to provide reference for smart services, smart marketing and smart management.

Sample collection and processing

Dianping.com as the world's first independent third-party consumer review website, ranking first in the domestic daily life essential tools. Compared with the OTA websites such as Ctrip.com and Fliggy.com, the online commentary of the public comment network is rich and true. It also directly displays the user's evaluation and satisfaction degree of tourism consumption in the scenic spot, and has the characteristics of initiative, individuality and high reference value. Therefore, this article uses the user comments of the OCT East Shenzhen on the Dianping.com as the data source.

Taking "OCT East Shenzhen" as the search key, the relevant travel commentary search was carried out in the data source of the user comments on the public comment online, and the "LocoySpider" software was used to collect 1283 related comments on "OCT East Shenzhen" published between March 31, 2017 and March 31, 2018. The data was then cleaned and preprocessed for comments:

(1) Excluding comments from any of the above features in which the same

visitor has repeated publications, with commercial nature, obvious complicity or malicious bad reviews, and less than two factors affecting visitor satisfaction;

(2) Unifying the different ways of expressions (synonyms, synonyms) of the same thing to ensure the consistency of the words. For example: WIFI (keywords), wireless network (synonym), Internet (synonym), networking (synonym), network speed (synonym), traffic (synonym); APP (keywords), mobile client (synonym), mobile phone software (synonym), application (synonym). Replace these homographs and synonyms with unified keywords to prevent them from being ignored by ROSTCM6 software because of the low frequency of partial and synonymous words;

(3) Delete the uploaded photos, videos, etc. in the comments. The final output contains 543 valid web comments and is stored as a .txt text file recognized by ROSTCM6 software for next quantitative and qualitative content analysis.

Method of analysis

(1) Quantitative content analysis: ROSTCM6 software was selected as the text analysis program for high-frequency word extraction in this paper. Through the segmentation and word frequency statistical analysis, the information matrix of scenic information service factors affecting tourists' satisfaction was obtained. Then SPSS Statistic 22 software was used for factor analysis to screen and eliminate the test items, and to extract the scenic spot information service attributes with high reliability and validity.

(2) Qualitative content analysis: Use the original web comment to verify whether the scenic spot information service that tourists are most concerned about can be better distributed in the factors obtained from quantitative content analysis. Through the qualitative interpretation method, the specific service content corresponding to the scenic spot information service attribute is enriched, which paves the way for comparative analysis of frequency satisfaction and interpretation of research results.

(3) Comparative analysis of rewards and punishments (PRCA): According to the results of quantitative and qualitative content analysis, a two-category regression analysis based on the structure of reward and punishment comparison analysis was carried out, and the identified satisfaction factors of

the various scenic spot information service attributes were attributed to the performance level of each service attribute, and then each attribute specifies an optimization scheme and a priority improvement order.

Empirical research

Text analysis

High frequency word extraction

ROSTCM6 was selected as the text analysis software for the high-frequency word extraction of this paper, and the 543 valid comments collected above were analyzed for batch content. By extracting the relevant high frequency words of the scenic spot information service and classifying them, the attributes that can represent the scenic spot information service are obtained.

(1) Custom word segmentation vocabulary and word frequency statistics filter vocabulary

First, typo check and format correction are performed for each comment to improve the accuracy of ROSTCM6 high frequency word extraction. Add names such as "Shenzhen", "OCT", "OCT East" and other scenic spots such as "Tea Valley" and "Yunhai Valley" and the names of "wooden roller coaster", "Tianzen" and "roaring mountain torrent" to ROSTCM6 custom word segmentation vocabulary and word frequency statistics filter vocabulary to ensure the validity of software analysis. At the same time, the high-frequency non-substantial vocabulary appearing in the comments, such as: service, scenic spot, play, and insufficient, is added to the meaningless filter vocabulary to prevent the vocabulary from interfering with the word frequency analysis and statistical process.

(2) Participle and word frequency analysis statistics

Using the word segmentation and word frequency analysis functions in ROSTCM6, the operation was repeated 10 times, during which the custom vocabulary was continually updated to exclude high-frequency words unrelated to this study. In the end, 73 high-frequency words that were mentioned by tourists and whose meanings were not repeated were generated (See Table 4.1).

Table 4.1 Scenic area information service high frequency vocabulary

Serial number	Key words	Frequency	Serial number	Key words	Frequency	Serial number	Key words	Frequency	Serial number	Key words	Frequency
1	Project	368	21	Personnel	87	41	Marketing	30	61	Open	21
2	Overhaul	333	22	Cableway	80	42	Position	30	62	Ticketing	21
3	Group buy	324	23	Management	77	43	Time	29	63	Burst	19
4	Cost-effective	324	24	WIFI	76	44	Souvenir	29	64	Problem	19
5	Surroundings	283	25	Family	72	45	Facility	29	65	WeChat	18
6	Small train	264	26	Consumption	70	46	Evaluation	28	66	Promotion	18
7	Effectiveness	227	27	Convenience	62	47	Commentary	28	67	Check-in	17
8	Old man	206	28	Couple	57	48	Year card	26	68	Child	17
9	Cable car	205	29	Dissatisfied	53	49	Identification	26	69	Elevator	17
10	Phone	198	30	Buy	53	50	Rest	26	70	Program	17
11	Customer service	183	31	Maneuver	51	51	Safety	25	71	Waste	16
12	Handle	155	32	Holiday	48	52	Improve	25	72	Remote	16
13	Performance	148	33	WeChat	46	53	Feel	23	73	Signal	15
14	Attitude	138	34	Entertainment	43	54	Parking	23			
15	Food	136	35	Low season	39	55	Praise	29			
16	Battery car	132	36	Queue	39	56	Management	29			
17	Map	112	37	Stroke	38	57	Features	22			
18	Vacation	96	38	Package	36	58	Fast	22			
19	Website	95	39	The weather	33	59	APP	22			
20	Value for money	88	40	Decoration	31	60	Check-in	22			

34

Factor Analysis

Using the search and search function of Microsoft Excel software, the number of occurrences of 73 high-frequency words in each visitor's comment was calculated separately, thus forming a frequency matrix of the number of occurrences of 73 high-frequency words in 543 comments.

(1) Correlation test of original variables

Before the factor analysis, the above-mentioned 73 variables should be subjected to KMO test (Kaiser-Meyer-Olkin) and Bartlett Test of Sphericity to determine whether the correlation coefficient between the variables and the correlation coefficient matrix is suitable for factor analysis. The frequency matrix was tested using SPSS22 with a KMO value of 0.831 and a Sig value of 0.000 in the Bartlett test, which is less than 0.05 and at a significant level (see Table 4 2). The above data indicates that the rectangle is a significant correlation matrix (Correlation Matrix) for factor analysis.

Table 4.2 KMO and Barlett test results

KMO sample measure		**0.831**
Bartlett's sphericity test	Approximate card square Approx. Chi-square	1081.985
	Degree of freedom df	231
	Significant probability Sig.	0.000

(2) Principal component analysis

In this study, principal factors analysis was used to extract common factors in quantitative analysis. High frequency words with a commonality (Communalities < 0.5) or a factor load value less than 0.35 will be removed from the correlation matrix to maintain the reliability and validity of the data. In the end, 42 keywords of 73 high-frequency words were summarized into 16 factors with eigenvalue greater than 1, and the cumulative variance reached 62.34% (> 60%), indicating that extracting 16 factors can better explain the service attributes they represent (See Table 4.3).

When the factor internal consistency test was performed, the coefficient of the factor 6, factor 11 and factor 13 (Cronbach's alpha test) was found to be less than 0.7 (see Appendix 1). In order to maintain the internal consistency of the factors, the three factors were deleted. However, in order to ensure the integrity of the 16 common factors, the study still named these three factors: "Emergency Management Service", "Tourism Product Purchase Service" and "Security Management Service" (see Table 4.4). The remaining 13 service

attribute common factors are: "Scenic Information Service", "Site Traffic Service", "Visitor Consultation Service", "Online Ticket Service", "Office Automation Service", "Online Information Service", "Electronic Tickets Services", "Network and Communication Services", "Visitor Complaint Services", "Smart Navigation Services", "Visitor Management Services", "Travel Route Planning Services" and "Online Payment Services" (see Table 4.4).

Table 4.3 Factor Cronbach internal consistency coefficient

Factor	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
a-coefficient	0.820	0.738	0.837	0.865	0.842	0.613	0.795	0.811	0.806	0.788	0.512	0.844	0.467	0.732	0.782	0.734

Table 4.4 Scenic spot information service factor identification concerned by tourists

Serial number	Factor name	High frequency word	Communalities	Variance contribution rate (%)	Serial number	Factor name	High frequency word	Communalities	Variance contribution rate (%)
1	Scenic information service	Overhaul	0.677	8.352	8	Electronic ticket service	Check-in	0.603	3.424
		Time	0.601				Year card	0.642	
		The weather	0.707		9	Network and communication services	Signal	0.632	3.378
		Holiday	0.578				WIFI	0.623	
2	Scenic traffic service	Battery car	0.589	6.552	10	Tourist complaint service	Management	0.725	3.189
		Small train	0.584				Dissatisfied	0.584	
		cableway	0.614		11	Tourism product purchase	Souvenir	0.621	3.013
3	Tourist advisory service	Customer service	0.623	4.764			Consumption	0.653	
		Staff member	0.641		12	Intelligent navigation service	Map	0.657	2.926
		Phone	0.553				Commentary	0.502	
4	Online ticket service	Group buy	0.690	4.254	13	Security management service	Safety	0.621	2.895
		Package	0.534				Old man	0.589	
		Cost-effective	0.655				Child	0.557	
5	Office automation service	Effectiveness	0.728	4.008	14	Tourist management service	Parking	0.678	2.781
		Convenience	0.638				Queue	0.645	
6	Emergency management service	Burst	0.553	3.864	15	Travel line rate planning	Stroke	0.662	2.644
		Facility	0.619				Identification	0.633	
		Many people	0.584				Promotion	0.707	
7	Online information service	Website	0.654	3.706	16	Online payment service	Alipay	0.724	2.588
		WeChat	0.554				WeChat	0.655	
		APP	0.566				Review	0.614	
Cumulative variance contribution rate (%)				35.5					62.34

Qualitative analysis

In order to verify whether the scenic spot information services of tourists in the online comment can be better distributed among the 13 factors obtained from the above analysis, the author uses the search function in the Word software to search the keywords covered by each factor in online reviews. A comparison between the search results and the scenic information service factor is tabulated (Table 4.5). The 13 factors from the factor analysis shown in the table can be found in the original tourist commentary, corresponding to the demand representation of the scenic spot information service.

Table 4.5 Qualitative explanation of the details of the scenic spot information service

Scenic information service attribute	Scenic information service content
Scenic information service	·Facility maintenance time notice
	·Project announcement operation time is consistent with actual operation
	·Weather forecast and corresponding scenic area resource regulation
	·Planning and promotional notices for important festivals
Scenic traffic service	·Directions and directions for the route of the battery car in the park
	·Train ride instructions and safety precautions across the scenic area
	·Propaganda of sights with visual and auditory experience inside the ropeway
Tourist advisory service	·Service quality of customer service staff in the visitor center
	·The level of help and attitude of the staff in the park
	·Use the call center for consultation before and after the tour
Online ticket service	·Special group purchases provided by distributors such as Meituan.com, Dianping.com.
	·Regular family or couple packages
	·Discounted prices experienced through online ticketing
Office automation service	·Automated, informative and intelligent office management software
	·Speed up the exchange of electronic tickets with a ticket machine
	·Employee efficiency and quality in the office process (low error rate)
Online information service	·The timely and new information of the portal website
	·Online customer service receives timely and efficient responses after consultation
	·Social interaction of WeChat public account
	·Applicability of APP updates and internal use of the scenic area
Electronic ticket service	·Detailed explanation and timely reminder of the ticket collection after purchasing the electronic ticket
	·Convenience when collecting tickets
	·VIP customer service for the annual card
Network and communication services	·Cell phone call signal sensing in the park
	·Cover the entire park's WIFI signals and services
	·Tourism digital footprint tracking built using WIFI connections
Tourist complaint service	·Old equipment and frequent maintenance
	·Project and management settings are unreasonable
	·Less complaint methods and avenues and few responses
Intelligent navigation service	·Campuses are large, It is difficult to play all project in the parks
	·There are a large number of campuses, but there is no more attractive commentary service or equipment configuration
Tourist management service	·Connection between parking lot and scenic spot
	·The tourist flow in the scenic spot is large, and the queue is serious.
Tourist route planning service	·Intelligent personal travel customization
	·Clear and clear park play and public service logo
	·Complete park attraction introduction and recommendation service
Online payment service	·Electronic payment options for consumption in the scenic area
	·Promotions with third-party payment or consumer review sites

Satisfaction structure analysis

Comparative Analysis of Awards and Punishments (PRCA)

(1) Constructing a virtual variable regression model

After the above quantitative and qualitative analysis, the factor scores of 13 scenic information services are used as independent variables, and the overall satisfaction of tourists is used as the dependent variable. Using SPSS 22 based on PRCA-based dummy variables for two-category regression analysis to determine the factor attributes of scenic information services that plays a decisive role in the satisfaction of tourists, and based on the three-element theory of customer satisfaction, the regression coefficient determines the influence of the service attributes on the degree of satisfaction of tourists.

Since it is difficult to directly measure the degree of tourist satisfaction from the text, only the overall satisfaction of the dependent variable is coded as two categorical variables, satisfaction is expressed as (OS=1), and unsatisfactory is (OS = 0). For example, when a customer gives a positive evaluation of the entire journey in a review, has an intention to revisit the scenic spot, or is willing to recommend the scenic spot to other people, the overall satisfaction of the review is OS=1. Conversely, negative comments are coded as OS=0, which means that visitors are not satisfied with the entire journey.

The 13 factors in this study have two dummy variables, one representing a higher level of performance (Xj+) and the other representing a lower level of performance (Xj-). When a visitor describes a very bad scenic information service performance level, the two dummy variables will be encoded in the form of Xj+= 0 and Xj-= 1; conversely, when the service performance level is high, the two variables will be encoded in the form of Xj+=1 and Xj-=0. In addition, if a certain scenic information service attribute is not mentioned in the comment, then both variables have a value of 0.

First, the 26 dummy variables of the 13 factors are measured by the regression analysis model of the overall tourist satisfaction below (see the following formula). OS is the overall satisfaction of visitors; Pos=1 is a satisfactory predictive value that visitors feel after experiencing the scenic information service, and a0 is a constant. β + and β - are the overall satisfaction of the j-th factor in the virtual variable (Xj+) with higher service performance level and the lower level virtual variable (Xj-) respectively.

$$Odd(OS(os = 1)) = \frac{P_{os \, = \, 1}}{1 - P_{os \, = \, 1}} = \exp\left[a_0 + \sum_{j=13} (\beta_{\,+} X_{j\,+} + \beta_{\,-} X_{j\,-}) \right]$$

After the first operation, excluding the two service attributes that did not reach the significant level of satisfaction with the tourists: tourist complaint service and travel route planning service, and finally 11 factors were added to the second round of regression analysis (see Table 4.6). The degree of matching of the model to the data was verified by null hypothesis, where Likelihood ratio chi square (16) = 551.17, p<0.001, which proved that the model matched well. In addition, the correct classification rate of the model is 94.55%, which is far above the maximum probability criterion of 55.26%, indicating that the 11 scenic information service attributes have an impact on tourist satisfaction. The possibility of experiencing a scenic information service at a specific high (low) performance level shows the likelihood that a visitor will be satisfied (not satisfied).

(2) Judgment of factor types of service attributes

Based on the three-factor theory of customer satisfaction, the overall satisfaction of the tourist experience scenic spot information service is based on the comparison between the "reward coefficient" β + and the "penalty factor" β -. The Wald test was used in this study to detect equal magnitudes between β + and β - in order to identify whether a particular scenic information service is a positive or negative contribution to the satisfaction of visitors. After two-category regression analysis, only five of the service attributes have significant effects on β + and β -, including "Site Information Service", "Visitor Consultation Service", "Online Ticket Service", "Office Automation Service" and "Visitor Management Services."

The null hypothesis H0: β -+ β +=0 applies to the above five factors and represents the reward effect and penalty effect of the specific scenic spot information service attribute on the overall satisfaction of the tourists. At the same time, this also means that the five service attributes are "performance factors", and its performance level is linearly related to the overall satisfaction of visitors. "Site Traffic Service", "Online Information Service" and "Electronic Ticket Service" are only significant in terms of penalty effects and are therefore "basic factors". Similarly, the remaining three factors: "network and communication services", "smart navigation services" and "online payment services" are only significantly related to the reward effect, so they are classified as "exciting factors".

Table 4.6 Two-category regression analysis and results of comparative analysis of rewards and punishments

Scenic information Service attribute	High performance level dummy variable					Low performance level dummy variable					Significant H_0: $\beta_+ + \beta_- = 0$	Types of factor
	Reward Coefficient (β_+)	Standard error (S.E)	Wald test	Overall satisfaction (%)	Satisfied predictive value (%)	Punishment coefficient (β_-)	Standard error (S.E)	Wald test	Overall satisfaction(%)	Not satisfied predictive value (%)		
Scenic information service	3.057	0.838	3.65	21.267	0.89	-2.089	0.644	-3.25	0.124	0.90	3.76	Performance factors
Scenic traffic service						-1.829	0.687	-2.66	0.161	0.89	0.764	Basic factor
Tourist advisory service	1.784	0.538	3.32	5.952	0.69	-2.087	0.739	-2.82	0.124	0.90	0.764	Performance factors
Online ticket service	3.07	0.478	6.42	21.539	0.81	-2.461	0.985	-2.5	0.085	0.93	0.594	Performance factors
Office automation service	4.065	1.453	2.8	58.25	0.97	-3.567	0.947	-3.77	0.028	0.97	0.767	Performance factors
Online information service						-2.098	0.749	-2.8	0.123	0.90		Basic factor
Electronic ticket service						-1.897	0.592	-3.2	0.15	0.89		Basic factor
Network and communication services	1.028	0.617	1.67	2.795	0.60							Exciting factor
Intelligent navigation service	1.941	0.656	2.96	6.963	0.78							Exciting factor
Tourist management service	2.724	1.063	2.56	15.241	0.87	-2.447	0.588	-4.16	0.087	0.92	0.824	Performance factors
Online payment service	3.442	1.28	2.69	40.003	0.95							Exciting factor

Analysis and discussion

Three-factor based tourist satisfaction optimization plan

First, OCT East Shenzhen should give priority to improving the "basic factors" and "performance factors" of service attributes that have a significant negative impact on visitor satisfaction due to service failure. The management objectives of the scenic spot should be positioned as these "basic factors" and "performance factors" to play a normal performance level, thus avoiding the unsatisfactory evaluation of tourists in the process of experiencing the scenic spot information service. If there are more resources available at the disposal level, then the scenic area operational objectives can be considered to upgrade from avoiding dissatisfaction evaluation to maximize the satisfaction of tourists. For example, improve the service attribute allocation in the "exciting factors" of OCT East, and further optimize the performance level of "performance factors". Based on the above-mentioned scenic spot information service attribute improvement scheme, it is possible to efficiently allocate and control limited scenic spot resources.

(1) Tourist satisfaction optimization plan based on "basic factors"

According to the above research results, the "basic factors" in the information service attributes of the OCT East are: "Site Traffic Service", "Online Information Service" and "Electronic Ticket Service". Combining the nature of the "fundamental factors" in the three-factor theory, on the one hand, the scenic spot should provide the "basic factor" service performance level that is comparable to what tourists expect, so as to avoid the strong dissatisfaction caused by the inconsistent perception and expectation of tourists; on the other hand, the service attributes in this category do not need to exceed the expectations of tourists, because the superior performance of the service in the "fundamental factors" does not bring about the same degree of improvement in visitor satisfaction. Scenic managers can adopt some low-cost, simple and effective short-term strategies to apply to this factor service attribute optimization. For example, (1) improving the official website information, WeChat public number and APP and other online information publishing platform content, timely response, timely response to consultation, friendly interaction design and visitor search preference tracking; (2) introducing more informative and intelligent products to create more practical instructions for transportation services and signs in the park; (3) improving the efficiency of existing information equipment such as ticket pick-up machines and electronic access control, and attaching importance to the training of employees in information aids.

(2) Tourist satisfaction optimization plan based on "performance factors"

As the five "performance factors" of the information service of OCT East, the service attributes such as "Site Information Service", "Visitor Consultation Service", "Online Ticket Service", "Office Automation Service" and "Visitor Management Service" provide long-term and stable contribution to the overall satisfaction level of OCT East tourists. Like the "fundamental factors", "performance factors" may also lead to more unsatisfactory evaluations; therefore, it is also necessary to maintain a performance level close to the expectations of tourists. However, further optimization of this factor should be based on the positioning of the scenic development and the actual budget to operate.

According to the frequency distribution in the qualitative content analysis, the favorable rate of the first three of the above five attributes is much higher than the latter two, which proves that there is a certain degree of scenic spot information service defect in the "performance factor". This paper attempts to provide the following solutions to solve this problem: (1) under the guidance of the general direction of building a smart scenic spot, properly introduce queues or virtual tour experience scenarios to buffer the inevitable long queues caused by large passenger traffic, and use the APP or the interactive function on the WeChat public platform to compensate the visitors to a certain extent (souvenirs, e-coupons, extra rewards and extra service upgrades); (2) comprehensive use of the sensor in the park, electronic access control, GIS, GPS and other information equipment for crowd forecasting and regulation to ensure the best tourist experience; (3) some physical scenic defects can be compensated by the excellent service quality of the staff, and a complete training and supervision system for front-line service personnel in the scenic area is established. The efficient help and friendly service attitude of the staff in the park during the tourism process can offset the loss of the scenic image brought by some bad reviews.

(3) Tourist satisfaction optimization plan based on "exciting factors"

Compared with the two factors discussed above, "exciting factors" ("network and communication services", "smart navigation services" and "online payment services") play an important role in improving and changing the long-term information service capabilities of the scenic spots. However, based on short-term resource and budget constraints, it is only appropriate to consider further enhancements to these service attributes when the first two factors match the expected level of visitors. Under the guidance of this idea, realize (1) the WIFI coverage of the entire OCT East scenic spot; (2) the precise intelligent navigation system inside the scenic spot, and (3) the intelligent electronic payment system, such as the universal electronic payment system, for enhancing the tourist experience and recording the intelligent footprint of the

tourism consumption, which needs to be gradually improved based on long-term strategic planning.

Optimization of scenic spot information service attribute's sequence based on regression coefficient

All in all, the "fundamental factors" and "performance factors" that provide the level of information services to meet the expectations of visitors should be ranked first in the scenic information service improvement plan. The "performance factor" and "exciting factor" optimizations to provide the most satisfying services for visitors are optional improvements based on budget and scenic development strategies.

Based on the degree of influence of each service attribute on the evaluation of tourist dissatisfaction, this paper gives the following short-term scenic spot information service capability optimization order according to the size of the two-class regression analysis coefficient and the priority level of service improvement from high to low: (1) office automation service (β -=-3.567), (2) online payment service (β -=-2.461), (3) online ticket service (β -=-2.447), (4) online information service (β -=-2.098), (5) scenic area information service (β -=-2.089), (6) tourist consultation services (β -=-2.087), (7) electronic ticket services (β -=-1.897), and (8) scenic traffic services (β -=-1.829).

Similarly, since each service attribute affects the degree of satisfaction of visitors, according to the size of the regression analysis coefficient, the priority level of the "performance factor" and "exciting factor" service attributes is raised from high to low, given the following long-term scenic spot information service capacity upgrade sequence: (1) office automation service (β +=4.065), (2) online payment service (β +=3.442), (3) online ticket service (β +=3.07), (4) scenic spot information service (β +=3.057), (5) tourist management services (β +=2.724), (6) smart navigation services (β +=1.941), (7) tourist advisory services (β +=1.784), (8) network and communication services (β +=1.028).

Conclusion

The mixed content analysis method throughout this article adds a large amount of web text mining content to the original PRCA, which is different from the current scale research perspective based on questionnaires. On the one hand, it reflects the tourists' satisfaction with the scenic spot information service. On the other hand, compared with the traditional single quantitative or qualitative web text mining studies that focus on image perception and sentiment analysis, the quantitative and qualitative content analysis method also provides more accurate and objective analysis ideas for the web text analysis used in this paper. The quantitative information method is used to

collect the information of the scenic spot information service, which greatly improves the efficiency of data collection and analysis. The qualitative interpretation and analysis on the basis of quantity makes the whole service attribute structured and concrete, which provides useful theoretical guidance and reference for the management group of the scenic spot.

The empirical study conducted in the case of Shenzhen Overseas Chinese Town as an example can be applied to other tourist attractions in specific scenic spots and specific service groups. It can also be other types of scenic spots where smart scenic spots are strategic development goals, such as cultural heritage scenic spots, national forest parks, rural tourist attractions, etc. It is even possible to screen out the order of service attribute optimization that affects the satisfaction of a city's entire tourism market based on the analysis of tourist satisfaction of web text analysis. In addition, the mixed content analysis method may not only be applied to the satisfaction research, but also applied to the research of tourists' purchasing decision, revisiting decision, and promotion motivation research.

Due to the limited ability of the author, there are some inevitable research flaws in this paper. For example, in the process of factor analysis of unstructured network comments, information may be missed or visitors may not fully describe their overall satisfaction with the scenic information service in the comments. In addition, this study only intercepts the online reviews of OCT in the relatively short period of time, and does not cover all the online reviews of OCT East. Therefore, the author also expects to continue to collect more comprehensive online comments and analyze them in the future when personal abilities are improved, in order to update more accurate and comprehensive research results.

Acknowledgement

References

Assaker G, Hallak R. (2012).European travelers' return likelihood and satisfaction with Mediterranean sun-and-sand destinations: A Chi-square Automatic Identification Detector-based segmentation approach . *Journal of Vacation Marketing*, 18(2):105-120.

Bonnefoy-Claudet L, Ghantous N.(2013). Emotions' impact on tourists' satisfaction with ski resorts: the mediating role of perceived value. *Journal of Travel & Tourism Marketing*, 30(6):624-637.

Dong J, Deng W.(2014). Empirical study on tourist satisfaction in cultural scenic spots based on principal component analysis: taking Shanxi Yuncheng as an example. *Science & Technology and Industry*, (03):40-44.

Dong J, Yan Y.(2014). Research on the satisfaction evaluation of tourists in Jiezhou Guandi Temple by Philip Kotler Evaluation Model. *Journal of Jiangsu Business Theory*, (4):36-38.

Gao H, Ren P.(2012). Research on the improvement strategy of tourist satisfaction in low carbonized scenic spots. *Journal of Statistics and Decision*, (23):60-63.

Ge J, Gu X, Long Y.(2012). Analysis of the construction of smart scenic spots based on stakeholder theory, *Journal of Productivity Research*, (5):183-184+225.

Gretzel U.(2011). Intelligent systems in tourism: A social science perspective. *Annals of Tourism Research,* 38(3):757-779.

Huang D, Chen F.(2015). Review of domestic and foreign tourism satisfaction research. *Journal of Chongqing Technology and Business University(Social Science Edition)*, 32(1):49-55.

Johnson M, Gustafsson A.(2006). Improving customer satisfaction, loyalty and profit: An integrated measurement and management system. John Wiley & Sons.

Krešić D, Mikulić J, Miličević K.(2013). The factor structure of tourist satisfaction at pilgrimage destinations: The case of Medjugorje . *International Journal of Tourism Research*, 15(5):484-494.

Liu Y, Li W, Liu H, Li X, Zeng Z.(2014). Evaluation of tourist satisfaction in free scenic spots based on grey correlation analysis. *Journal of Travel Overview* (4):176-178.

Lü L, Li X.(2015). Research on tourist satisfaction survey in cultural tourism scenic spots in Henan Province. *Journal of Anyang Teachers College*, (5):59-65.

Master H, Prideaux B.(2000). Culture and vacation satisfaction: a study of Taiwanese tourists in South East Queensland. *Tourism Management*, 21(5):445-449.

Murphy L, Moscardo G, Benckendorff P, et al.(2011). Evaluating tourist satisfaction with the retail experience in a typical tourist shopping village. *Journal of Retailing and Consumer Services*, 18(4):302-310.

Pearce PL, Wu MY.(2016). Tourists' evaluation of a romantic themed attraction expressive and instrumental issues. *Journal of Travel Research*, 55(2):220-232.

Qin X.(2014). Research on tourist satisfaction survey in rural tourism scenic spots in Shanxi Province based on psychological contract. *Economist*, (4):218-219.

Shitai JI.(2012). Research on tourist satisfaction and its influencing factors in red scenic spots. *Journal of Northwest Normal University (Natural Science)*, (5):115-120.

Smith RK, Olson LS.(2001). Tourist shopping activities and development of travel sophistication. *Visions in Leisure and Business*, 20(1):23-33.

Truong TH, King B.(2009). An evaluation of satisfaction levels among Chinese tourists in Vietnam . *International Journal of Tourism Research*, 11(6):521-535.

Wang Z.(2013). Research on red tourist tourist satisfaction system from the

perspective of public service . *Journal of Frontiers*, 21:97-99.

Wu B.(2015). Cracking chaos, smart tourism ushered in rational return. China Publishing Media Business Daily, 2015-08-14(13).

Xia Q, Wang Z.(2012). Research on tourist satisfaction of mountain-type scenic spots based on fuzzy-IPA: taking Huangshan scenic spot as an example. *Journal of Anhui Normal University (Natural Science)*, (5):471-476.

Yang R, Liu Y.(2015). Research on tourist satisfaction of eco-tourism scenic spots based on IPA Model: Taking Taiping National Forest Park in Xi'an as an example. *Travel Research*, (2):69-74.

Yang W, Hu X.(2015). Research on tourist satisfaction of Wild Elephant Valley scenic spot in Xishuangbanna based on IPA analysis . *Journal of Southwest Forestry University*, (3):73-77.

Yang Y, Li L, Li Q.(2015). Discussion on the development policy of mountain tourism destinations based on satisfaction survey: Taking Damingshan in Nanning as an example. *Journal of Market Forum*, (2):83-84.

Yu L, Goulden M.(2006). A comparative analysis of international tourists' satisfaction in Mongolia. *Tourism Management*, 27(6):1331-1342.

Yu Y.(2014). Evaluation of tourist satisfaction in Huaguo Mountain scenic spot based on IPA analysis. *Journal of Changsha University*, (5):108-111.

Zhang H, Li Q, Liu H.(2014). An empirical study on the satisfaction of inbound tourists in Xi'an based on the "Ferner Logical Model" . Journal of Shaanxi Preschool Teachers College, (04):93-97.

Zhao W, Sun J, Song H, Xiong Q.(2014). Research on the measurement and improvement strategy of psychological satisfaction of tourists in red tourism service. *Contemporary Education Theory and Practice*, (6):119-121.

About the Authors

Ke Zhang，master student at Shenzhen Tourism College, Jinan University. She specializes on smart tourism and intangible cultural heritage tourism.

E-mail: Constancezhangke@Gmail.com

Mu Zhang, Ph.D., Professor, Associate Dean, Shenzhen Tourism College of Jinan University, Director of JNU-UF International Joint Laboratory on IT& Tourism. He is the corresponding author of this article and he engaged in researches on geography, geographical information system, and teaching of tourism management, etc.

E-mail: zhangmu@163.com