This is the Pre-Published Version.
Nao Li, Vincent Tung & Rob Law (2017) A fuzzy comprehensive evaluation algorithm for analyzing electronic word-of-mouth, Asia Pacific Journal of Tourism Research, 22:6, 592-603.
This is an Accepted Manuscript of an article published by Taylor & Francis in Asia Pacific Journal of Tourism Research on 31 Mar 2017 (published online), available at: http://www.tandfonline.com/10.1080/10941665.2017.1308395.

Title

A fuzzy comprehensive evaluation algorithm for analyzing electronic word-of-mouth

Abstract

This study evaluates tourism experiences shared through electronic word-of-mouth (eWOM) across four Chinese attractions. The objective is to develop a framework for evaluating eWOM by constructing an indicator system and implementing an analytic hierarchy process with the use of a fuzzy comprehensive evaluation algorithm. This framework is achieved by mapping more than 6,000 websites related to Chinese tourism attractions and filtering over 200,000 useful reviews to measure service performance. Results indicate that ecological-biological attractions failed to make tourists feel "very satisfied" in various aspects, such as overall evaluation, infrastructure, traffic, natural environment and social environment. Overall, the study contributes by presenting a framework that can be adopted by tourism researchers and industry practitioners to understand tourist preferences and evaluate service performance to improve service quality.

Keywords: electronic word-of-mouth, big data, analytic hierarchy process, fuzzy comprehensive evaluation algorithm

1. Introduction

Facilitating satisfactory and memorable tourism experiences are critical to destination competitiveness as these experiences are often reflected through tourists' word-of-mouth (WOM) (Baloglu et al., 2004; Tung & Ritchie, 2011). WOM reflects the difference between product performance and user perspectives in the life cycle of a product (i.e., from introduction to growth, maturity, and decline) (Mahajan & Muller, 1984). It also plays a significant role in understanding customer segmentation, overcoming product limitations, and achieving market development.

Given the advances in information technology and big data, the scope of tourism research on WOM has greatly expanded. Researchers and industry practitioners have realized the importance of electronic word-of-mouth (eWOM) in investigating tourist experiences and preferences. For example, previous studies have analyzed eWOM based on user-generated content (UGC) for information dissemination, purchase involvement, and consumer experiences (Hennig-Thurau et al., 2004; Hennig-Thurau& Walsh, 2003; Arenas-Márquez et al., 2014; Shin et al., 2014). Despite the tremendous innovation in scientific research from the use of big data, certain challenges must be addressed in applying big data to tourism research (Bryant et al., 2008). These challenges include data fragmentation, complexity of technology, data accuracy, right to use, business and technology alignment, and requirement of data specialists (Davenport, 2013). Furthermore, existing studies on big data and eWOM are largely concentrated in a hospitality context that measures factors based on the satisfaction of hotel guests, including location, room, price/value, and food and beverage.

In light of this research gap, the present study aims to contribute to the literature on big data in a tourism setting by evaluating tourism experiences shared through eWOM. The objective is to develop a framework for evaluating eWOM by constructing an indicator system for data acquisition and implementing an analytic hierarchy process (AHP) with the use of a fuzzy comprehensive evaluation (FCE) algorithm. This framework is achieved by mapping more than 6,000 websites related to Chinese tourism attractions and filtering over 200,000 useful reviews to measure the gaps between attraction service performance and the "very satisfied" sentiment of tourists. In doing so, this study contributes toward adopting a novel framework by combining a big data platform and sentimental analysis to interpret the eWOM of tourist experiences.

2. Literature review

2.1 eWOM in tourism

eWOM is the diffusion of information across communities that use network information technology (Lee & Hu, 2005). This process communicates positive, negative, and neutral comments on the features and services of target products by past consumers to present and potential consumers (Hennig-Thurau et al., 2004). Early adopters used information obtained through eWOM to determine product properties, advantages, and disadvantages, as well as to predict product trends (Moore & McKenna, 1999).

Recent studies on eWOM have focused on the relationship between electronic information and customer preferences. For example, consumers can browse Internet sources for target products and recommendations to aid their decision making (Yilmaz &Asli, 2013), especially when they do not have previous experience in purchasing the target product or service. The motives behind the information posted, read, and diffused by consumers have been investigated (Hennig-Thurau& Walsh 2003; Park & Kim, 2009).

eWOM, as a cost-effective means for tourism marketing, has two key points, namely, production of information and generation of revenue (Litvin et al., 2008). The former refers to an effective and efficient framework for collecting positive, negative, and neutral reviews disseminated on the Internet when conducting eWOM analysis. The latter refers to the importance of applying eWOM analysis to the industry for revenue generation. Previous studies typically employ questionnaires or interview surveys to investigate the mechanism and spread of eWOM generation. Bronner and Hoog (2010) conducted interviews to study the relationship between vacationers and eWOM. Sotiriadis and Van Zyl (2013) administered a questionnaire survey to investigate the effects of eWOM on the purchase decisions of customers. Other researchers have also manually collected review data from online channels (Gretzel &Yoo, 2008; Pan & Crotts, 2012; Ayeh et al., 2013). Ye et al. (2011) collected hotel rating data from Ctrip and found that traveler reviews can affect business performance. Pan et al. (2007) collected 40 blogs from travelblog.org, travelpod.com, and travelpost.com to evaluate service quality in particular destinations.

2.2 Big data and sentiment analysis in tourism research

Rigorous academic work requires high-quality data. Advances in information technology (i.e., big data technology) have enabled researchers to collect high-quality data from investigation samples that are frequently scattered. Although the concept of big data does not have a generally agreed-upon definition, it is typically defined to contain three features: volume, velocity, and variety (Chen et al. 2014). According to the International Data Corporation (Gantz & Reinsel, 2011), "big data technologies describe a new generation of technologies and architectures, which are designed to economically extract value from very large volumes of a wide variety of data by enabling high-velocity capture, discovery, and/or analysis." Big data technology involves the collection and exploration of data to generate decisions and predictions.

Big data has introduced opportunities to tourism research by annually producing large amounts of data (Xiang & Gretzel, 2010). The approach provides a comprehensive data spread by collecting integrated information from various organizations. Thus, customer preferences can be identified from individual travelers, enabling efficient and adequate decisionmaking by tourism marketers. For example, extant tourism research has applied data resources and research goals in three aspects.

First, researchers have used GPS tracking data to explore tourist behaviors. Edwards and Griffin (2013) used GPS tracking devices to record tourist movement data in Sydney and Melbourne to help destination managers improve visitor experiences by improving the tourist capacity to find systems. McKercher et al. (2012) investigated the difference between first-time and repeat visitor behavior patterns (i.e., travel distance and intermittent activity) based on GPS data in Hong Kong.

Second, user-generated data (UGD) have been used to analyze guest experiences and visitation rates to destinations. Xiang et al. (2015) collected 60,648 customer reviews across 10,537 hotels in Expedia to explore the relationship between hotel guest experiences and satisfaction. Wood et al. (2013) proposed a big data approach to estimate the visitation rates and tourist origins of 836 recreational sites on the basis of the locations and visitor profiles of 197 million geotagged photographs uploaded on Flickr from 2005 to 2012. Crotts et al. (2009) developed a stance-shift analysis to evaluate hotel guest satisfaction based on Internet blogs.

Third, sentiment analysis has been used to identify the attitude of a speaker or a writer from UGD (Pang& Lee, 2008). Sentiment analysis can resemble an extension analysis of questionnaire surveys. For example, after tourists visit a scenic spot, they usually conduct reviews that include subjective observations. Scores in a traditional questionnaire survey are provided by a tourist, while scores for sentiment analysis are provided by reviews and travel blogs. For example, Ye et al. (2009) studied the performance of three sentiment classification techniques: Naïve Bayes, SVM, and the character-based N-gram model were used on travel blogs and reviews. Waldhör and Rind (2008) developed a semi-automatic tool based on linguistic parsing methodology and terminology extraction, called etBlogAnalysis, to feature valuable information in travel blogs. García et al. (2012) used the WordNet lexicon database to calculate the sentiment score of review keywords, which are important for evaluating customer decisions with regard to accommodation or food and beverage consumption.

The present study maps more than 6,000 websites related to Chinese tourism attractions and filters more than 200,000 useful reviews to measure the gaps between attraction service performance and the "very satisfied" sentiment of tourists. This study is an attempt to use big data in terms of data volume and sentiment analysis.

3.Research framework

A research framework is developed (Figure 1) to apply big data in evaluating tourism experiences shared through eWOM.

---- Please insert Figure 1 here ----

The framework is based on a big data platform called DINFO-OEC Text Mining Platform. The platform is a Hadoop-based patented product for Chinese text acquisition, processing, and sentimental analysis (http://www.dinfo.cn/html/product/product_OEC.html). The use of the DINFO-OEC text mining platform in industries (i.e., Industrial and Commercial Bank of China) has been demonstrated to be effective and efficient in handling unstructured data. In essence, DINFO-OEC uses an O (ontology) - E (element) - C (concept) model to specify a text. O represents layered business categories or objectives composed of E and C. E represents domain relevant expressions, such as research objects and their attributes. C represents domain irrelevant expressions, such as time, place, emotion, and attitude. The main idea of the OEC model is separating the expressions of business and nature language and thus enabling researchers to focus on domain knowledge rather than the diversity and complexity of nature language expressions. When the OEC model is used to process the tremendous amount of data acquired, non-structured texts are modeled with structured data and semantics and can be further analyzed according to research aims.

The study began with an iterative data acquisition and indicator construction. The iteration assured that the data were available for evaluating service performance according to the constructed indicator system and that the constructed indicator system was effective for evaluating the data. After constructing an optimized indicator system, the acquired data were processed from non-structured UGC text to structured data to prepare for sentimental analysis and evaluation. The weights of indicators were identified by an AHP. Finally, an FCE algorithm was used to evaluate the service performance of four categories of attraction, and the resulting scores were used to rank the eWOM of the attractions.

4. Data acquisition and analysis

4.1 Data acquisition and cleaning

This study was designed to adopt the manual and machine learning of collected UGC content from the Internet. At the initial stage, more than 6,000 websites related to Chinese tourism were searched, including tourism news, marketing, organizations, reviews, and blogs. A filtration process was conducted to exclude websites with non-UGC content and those with outdated and false links to further examine the different dimensions of data resources. A list of 100 websites was obtained. The websites were classified based on a pattern of gaining UGC content to obtain data resources of landscape reviews. A total of 14websites (e.g., blog.sina.com.cn, ctrip.com, and qunar.com) were consequently accessed during the data acquisition process. Data were collected from June 2013 to September 2014 to cover 20 Chinese destination cities, 228 tourist attractions, and more than 4,000 scenic spots. After data cleaning, which included the removal of duplicates and very short and meaningless reviews, a total of 257,635 reviews were considered valid.

4.2 Construction of the indicator system

Three steps were conducted to construct the indicator system based on big data for evaluating the service quality of different attractions. First, first-class indicators covering six tourism elements, namely ,food, accommodation, transportation, travel, shopping, and entertainment (Smith, 1988),were initialized based on the classic ServQual model (Parasuraman, 1988). Each element covers five dimensions, namely, tangibles, reliability, responsiveness, assurance, and empathy, and 22 statements following the keyword list. Then, the indicator system was tested on the collected data to determine its effectiveness. The result indicated that the indicator system could not adapt completely to the UGC of the Internet, because the UGC contains unstructured expressions of what tourists think and does not follow the indicator system. Moreover, the data that could be used for evaluation in terms of the initialized indicator system were limited and difficult to acquire. Therefore, the third step was conducted to adjust and optimize the initialized indicator system through a manual selection and computer experiment process. A sample acquisition website (www.dianping.com) was used to collect reviews. Through word segmentation, the keyword frequency was recorded and high-frequency keywords were manually integrated with those in the keyword list defined in the initialized indicator system. The integrated keywords were classified into first- and second-class indicators. A test was subsequently conducted to examine the coverage of data acquisition based on the new indicator system. If keywords were missing, these words would be added to the classification to update the new indicator system. Thus, the new indicator system could be used to collect the majority of tourist reviews.

Figure 2 shows that the final indicator system has seven first-class indicators (i.e., social environment, natural environment, traffic, infrastructure, suitability, single evaluation, and overall evaluation) and 31 second-class indicators. The indicator system is different from our previous assumptions and from that of the questionnaire survey because tourists are more concerned about their emotional experiences and less on established services (i.e., security and attitude of staff).

---- Please insert Figure 2 here ----

4.3 Data processing and sentiment analysis

The DINFO-OEC platform processed all the reviews using the OEC model. Given that the indicators have been constructed, E in the OEC model could be identified. Therefore, all the reviews were structured as Attractions: Indicators +Sentiments (e.g., Summer Palace: traffic +dissatisfied). This study used sentiment analysis to capture certain indicators expressed in a review. The following score scale was first established for the sentiment analysis: "5" for very satisfied, "4" for satisfied, "3" for neutral, "2" for dissatisfied, and "1" for very dissatisfied. Sentiment analysis was conducted with the DINFO-OEC platform. The platform uses word segmentation to split the sentences and annotates the property of each Chinese character. The sentiment can be identified in several levels based on the lexicons provided by the platform(e.g., "very" + "good" for the score of 5). The experiments show that the precision of DINFO-OEC sentiment analysis is more than 85%, whereas recall is more than 90%.

4.4Evaluation of attractions

4.4.1 Four categories of attractions

Attractions have various categories. This study adopts the four categories identified by USAID (Stange & Brown, 2013): geophysical–landscape–aesthetic, ecological–biological, cultural–historical, and recreational categories. The geophysical–landscape–aesthetic category includes mountains, gorges, big rocks, rock formations, caves, rivers, water bodies, scenic views, and overlooking forests. The ecological–biological category includes the environments of organism behavior, reproduction, predation, and migration. The cultural–historical category includes manifestations of human evolution. The recreational category includes attractions that aim to entertain and educate people. All 228 attractions in our list were classified into four categories for further analysis.

4.4.2 Analytic hierarchy process

AHP (Saaty, 1990) is a combination of qualitative and quantitative approaches used to model complex systems and provide assistance in the decision-making process. The influential factors of complex problems can be decomposed into several top-down layers by AHP. The factors in each layer belong to the top-layer factors and control the lower-layer factors. AHP implementation includes establishing a judgment matrix through pairwise comparison, determining the weight of each influential factor, and conducting a consistency check. The advantage of AHP is the use of less quantitative data to make decisions.

This study adopted AHP to calculate the weight of the first- and second-class indicators based on the number of reviews. Many reviews note that indicators contribute more to eWOM evaluation and should be labeled with higher weights. The numbers of reviews of indicators widely varied. The review numbers of “natural environment” and “suitability” were 74,589 and 6,206 respectively. Thus, the measured scale constructed by neighboring comparative methods exceeded the value of 9. The measured scale was generally limited to the value of 9 to describe the significant ratio between two indicators. A measure scale that exceeds 9 does not make sense. Therefore, the original calculated measurement scales were adjusted and limited from 1 to 9. The adjusted measurement scale, such as the first-class indicators, is shown in Table 1.

Two criteria were set in weight calculation to fit the standard AHP calculation process:
1. The weight of all indicators was determined by the number of reviews.
2. Measurement scales were adjusted to fit the principle of AHP and were limited from 1 to 9.

All weights calculated in AHP were determined by a random consistency check to avoid the contradiction of indicator significance (e.g., A>B, B>C, and C>A). If C.R. <0.1, then the constructed judgment matrix has acceptable consistency and can be considered acceptable. As shown in Table 1, the consistency check (see Table 1) of first-class indicators is C.R. = 0.0072.

---- Please insert Table1 here ----

Table 2 presents the final weight of the indicators determined by AHP. Given that weight was calculated based on the number of eWOM reviews, the results appear consistent with the data distribution of the indicators. Weights represent the dynamics of the data being collected. Data collection and analysis were performed daily. The weights were dynamically changed as data grew in terms of different years, seasons, weeks, and days.

---- Please insert Table2 here ----

4.4.3 Fuzzy Comprehensive Evaluation

FCE (Zadeh 1976) combines fuzzy theory and mathematical models from low to high levelsto comprehensively evaluate the operational effects of measures. FCE is widely used to quantify fuzzy and uncertain problems. It can be taken into account when additional factors should be considered in decision making. Tangible or intangible factors are included, and these factors are hierarchical. FCE was applied to combine the first-class indicators and comprehensively evaluate the eWOM of tourist attractions.

$C_i = \{C_{i1}, C_{i2} ... C_{ik}\}$ is one of the second-class indicator sets for a certain first-class indicator set $C = \{C_1, C_2 ... C_i\}$. The weight vectors obtained from AHP are $w = \{w_1, w_2 ... w_i\}$ and $w_i = \{w_{i1}, w_{i2} \cdots w_{ik}\}$ for the first- and second-class indicator sets respectively. The detailed algorithm of FCE is divided into six steps:

Step 1: The ranking set is determined according to the sentiment analysis order $v = \{v_1, v_2 ... v_m\}$.

Step 2: The evaluation matrix $R = (r_{ij})_{m \times n}$ is calculated based on the ranking set.

$$R_i = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ r_{k1} & r_{k2} & \cdots & r_{km} \end{bmatrix}.$$

Step 3: The i th subordinated degree matrix for each second indicator $B_i^{'}$ is

$$B_i^{'} = W_i \circ R_i = (w_{i1}, w_{i2} \cdots w_{im}) \circ \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ r_{k1} & r_{k2} & \cdots & r_{km} \end{bmatrix} = (b_1, b_2, \cdots b_m).$$

Step 4: $B_i^{'}$ is normalized as

$$B_i = \left(\frac{b_1}{\sum b_m}, \frac{b_2}{\sum b_m}, \cdots, \frac{b_m}{\sum b_m} \right).$$

Step 5: The first-class subordinated degree matrix is constructed.

$$B_{(i \times m)} = (B_1, B_2 \cdots B_i)^T.$$

Step 6: If a score matrix is $G = (G_1, G_2 \cdots G_m)^T$, then the final evaluation score of a specific first-class indicator i for a certain attraction is computed as

$$S_i = B_{i \times m} \times G = (B_1, B_2 \ldots B_m) \times (G_1, G_2 \cdots G_m)^T.$$

Each first-class indicator was considered a single indicator system and processed in each attraction as a onetime FCE. The evaluation result of each indicator of each attraction was then obtained.

An indicator set and its weight are represented in this study by c and w respectively. The ranking set was determined by sentiment analysis as described in Section 3.3. Therefore, $v_i = \{verysatisfied, satisfied, common, dissatisfied, verydissatisfied\}$. The eWOM evaluation score can be easily obtained if the score matrix is $G = (100,80,60,40,20)^T$, which corresponds to scores of 5, 4, 3, 2, and 1 in the sentiment analysis.

5 Results

5.1 Distribution of data quantity

The distribution of data quantity in this study is shown in Table 3. Single evaluation and natural environment account for 35% and 22% of all reviews respectively. The review of visitor flow rate accounts for 45% in the single evaluation of first-class indicators. The value is reasonable for the large Chinese population. An accurate proportion was obtained in this study based on the big data platform. Moreover, charge rationality accounts for 28% in the single evaluation. This result denotes that the cost of tickets to tourist attractions is an important concern of tourists in their travels under current Chinese living standards. Second-class indicators for the natural environment are more evenly distributed. For example, landscape accounts for 25%, scenic environment for 16%, air quality for 13%, water quality for 22%, and greening for 24%. The suitability indicator occupies the least proportion in the distribution. This result may be attributed to the fact that most Chinese tourists, especially the young (2%) and student (5%) groups, are less concerned about personalized tours.

---- Please insert Table3 here ----

5.2 Tourist sentiment

The results of the sentiment analysis showed the satisfaction of tourists when they travel to various attractions. The proportions of “very satisfied” and “satisfied” are 36% and 27% respectively. More than half of the tourists hold a positive WOM on their travel experiences. Approximately 14% of the tourists were “very dissatisfied” with the tour, and only 1% was “dissatisfied.”

5.3 Evaluation of four categories

FCE was conducted to determine the score of each first-class indicator for all attractions. The eWOM score of each first-class indicator was obtained (see Table 4) by integrating the scores of the indicators in the four categories. For example, the score of “social environment” indicator in the cultural-historical attraction category is 87.399. Reviews related to “suitability” indicator were classified as “satisfied” (score 80) in the sentiment analysis.

---- Please insert Table4 here ----

The gap was calculated since the tourist "very satisfied" sentiment score was 100, and then the gap score was normalized. The final results are presented in Figure 3. “Overall evaluation” has the largest gap between attraction service performance and the "very satisfied" sentiment of tourists.

---- Please insert Figure 3 here ----

This study further examined the gap between attraction service performance and the "very satisfied" sentiment of tourists in second-class indicators (Fig. 4). As shown in Figure 4, tourists expressed their dissatisfaction in all types of attractions, especially in the ecological-biological attraction category, with a gap of more than 60%. The “cost performance” indicator is another gap generator. Ecological-biological attractions have a “traffic” gap in the first-class indicator. However, the reasons for the gap in the second-class indicators should still be further explored. “Parking” performed well but “traffic outside the land area” did not. This result indicates that ecological–biological attractions must promote outside traffic service quality to improve tourist satisfaction.

The Chinese tourist satisfaction survey (The Chinese Tourism Academy, 2014) showed that national tourist satisfaction declined because of poor performance, inadequate basic infrastructure, and poor health conditions, all of which coincided with the results of the present study.

---- Please insert Figure 4 here ----

6. Discussion and implications

Triggered by the research gap among big data and tourism, tourism experience, and eWOM, this study developed a framework for evaluating eWOM. The framework is based on a big data platform that supports data acquisition, data processing, indicator system construction, and sentimental analysis. Data were acquired across platforms, that is, data were collected not only from one website but from many websiteswith tourist reviews. This study maps more than 6,000 websites related to Chinese tourism attractions and filtered more than 200,000 useful reviews. This method of collecting data avoided potential biases in social media data, which a single and specific source could cause (Ruths& Pfeffer,2014).

The indicator construction is an iteration process of data acquisition and indicator optimization based on the data, which gave rise to a new approach of understanding and using UGC to evaluate services. Currently, no indicator system based on UGC data exists to evaluate the service performance of attractions. When the initial indicator system according to classic ServQualwas defined, our data acquisition showed that there were not enough reviews on the Internet with regard to the indicators. This finding indicated that UGC does not follow the indicators in classic ServQual but contains free expressions of what tourists think or want to share. Classic ServQual does not reflect what UGC mainly focuses on. To construct an effective indicator system for review-based evaluation, an iterative process was conducted to construct the indicator system. This involves adding high-frequency words or focusing on collected reviews and testing new indicators until the indicator system can be used to analyze the majority of tourist reviews. Furthermore, the constructed indicators shed a light on a certain structure among various aspects andattributes related to attraction services; such structure is deemed a promising research area in social media analyticsin hospitality and tourism (Xiang, Du, Ma, & Fan, 2017).

On the basis of the data acquired and the indicator system, the evaluation of four categories of attractions was implemented by an AHP with the use of an FCE algorithm. First, the eWOM scores of the first-class indicators in the four categories were calculated. Overall, only the scores in Social environment and Suitability were over 80 and no obvious score difference in the four categories of attractions on the indicators was obtained, except in Traffic, Infrastructure, and Overall evaluation. Culture–historical attractions received the highest score and ecological–biological the lowest in Traffic. This result indicated that tourist experience coincided with the degree of transportation systems of the two types of attractions. In China, most culture–historical attractions were developed earlier and have well-built transportation systems for easier access, whereas most ecological–biological attractions are far from cities, and tourists feel they are in convenient and difficult to access. Culture–historical attractions also obtained the highest score in Infrastructure, and recreational attractions obtained the lowest score. This result again indicated that culture–historical attractions in China are better built and developed, and that tourists are more satisfied with the infrastructure compared with the other types of attractions. On the other hand, recreational attractions should improve their infrastructure to meet the rapid recreation population. Geophysical–landscape–aesthetic attractions received the highest score in Overall evaluation, and ecological–biological attractions obtained the lowest score. As emerging attractions in China, ecological–biological attractions should improve in the following indicators: overall evaluation, popularity, cost performance, satisfaction, and visiting experience.

Second, the gaps between attraction service performance and the "very satisfied" sentiment of tourists were measured. Significant gaps exist between service performance and the "very satisfied" sentiment of tourists across all four categories of attractions. In terms of the type of attraction, ecological–biological attractions have larger gaps than the other types. Ecological–biological attractions have a significantly large gap between service performance and the "very satisfied" sentiment of tourists in Infrastructure. Similar to recreational attractions, ecological–biological attractions should improve their infrastructure to facilitate tourism development. As regards the Health condition, ecological–biological attractions also have significantly larger gaps than the other attractions. Ecological–biological attraction managers need to address these gaps because the ecological environment is critical for ecological–biological attractions and needs to be sustainable, and yet, tourists are not satisfied. Ecological–biological attractions also have a gap in Traffic, a result that coincides with the lowest score in Traffic mentioned above. For ecological–biological attractions, managers need to be cognizant of the increasing tourist requirement for accessibility, convenience, and ecological sustainability. In terms of a certain indicator, the Social environment indicator has the best performance in all first-class indicators. Educational significance, Local specialty, and Humanistic characteristics have very small gaps. These outcomes indicated that tourists approve the cultural characteristics that most attractions feature. However, other indicators, such as Charge rationality, Basic facilities, and Parking, have gaps between attraction service performance and the "very satisfied" sentiment of tourists. Charge rationality is always a popular topic among social media in China. Tourists are very sensitive to an increase in prices. Service facilities, such as parking, should also be improved by most attractions in China.

The findings of this study not only generated managerial insights for attractions, i.e., that UGC can help them understand tourist experiences, what tourists pay attention to, and how they evaluate their services, but also provided the services that have gaps between service performance and the "very satisfied" sentiment of tourists. These findings can help managers in tourist attractions make managerial strategies, policies, and decisions to improve their services. However, services related to tourism attractions are not only provided inside the attractions but also involve transportation, traffic, and pricing, all of which cannot be regulated by the attractions. Therefore, the findings of this study also gave insights and suggestions for governments on government-related services and policies.

7. Conclusion and limitations

With the increase of UGC on tourism in the Internet, eWOM has become an important source to capture tourism service performance. To investigate the gaps between service performance and the "very satisfied" sentiment of tourists across four Chinese attractions through eWOM, this study proposed a framework for analyzing eWOM by including a two-class indicator system for data acquisition and an FCE algorithm by implementing an AHP process, sentiment analysis, and FCE algorithm. The proposed approach allowed for the evaluation of tourism service performance when an increasing number of tourists express their opinion on experienced services in the Internet.

This study has limitations that should be addressed in future research. First, the indicator system is primarily designed for manual and semi-automatic handling, which may lower the reasonability of the experiment results. An improved indicator production system should be designed to use machine learning technology for automatically producing, filtering, and optimizing indicators. Second, the weight calculated by AHP can be changed with the increase in the number of reviews, which may also increase the uncertainty of the final evaluation results. Additional statistical work should be conducted to analyze eWOM changes because the number of reviews varies.

References

Arenas-Márquez, F. J., Martínez-Torres, M. R., & Toral, S. L. (2014). Electronic word-of-mouth communities from the perspective of social network analysis. *Technology Analysis & Strategic Management, 26(8)*, 927-942.

Ayeh, J. K., Au, N., & Law, R. (2013). Predicting the intention to use consumer-generated media for travel planning. *Tourism Management, 35*, 132-143.

Baloglu, S., Pekcan, A., Chen, S. L., & Santos, J. (2004). The relationship between destination performance, overall satisfaction, and behavioral intention for distinct segments. J*ournal of Quality Assurance in Hospitality & Tourism, 4(3-4)*, 149-165.

Bronner, F., & de Hoog, R. (2010). Vacationers and eWOM: Who posts, and why, where, and what? *Journal of Travel Research, 50(1),* pp. 15–26.

Bryant, R., Katz, R. H., & Lazowska, E. D. (2008). *Big-data computing: creating revolutionary breakthroughs in commerce, science and society*: A white paper prepared for the Computing Community Consortium committee of the Computing Research Association. Retrieved from: http://cra.org/ccc/resources/ccc-led-whitepapers/.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: a survey. *Mobile Networks and Applications, 19(2)*, 171-209.

Crotts, J. C., Mason, P. R., & Davis, B. (2009). Measuring guest satisfaction and competitive position in the hospitality and tourism industry an application of stance-shift analysis to travel blog narratives. *Journal of Travel Research, 48(2),* 139-151.

Davenport, T. H. (2013, June 26). At the big data crossroads: Turning towards a smarter travel experience. Retrieved from http://www.bigdata.amadeus.com/assets/pdf/Amadeus_Big_Data.pdf.

Edwards, D., & Griffin, T. (2013). Understanding tourists' spatial behaviour: GPS tracking as an aid to sustainable destination management. *Journal of Sustainable Tourism, 21(4),* 580-595.

García, A., Gaines, S., & Linaza, M. T. (2012). A lexicon based sentiment analysis retrieval system for tourism domain. *Expert System Application International Journal, 39(10)*, 9166-9180.

Gretzel, U., & Yoo, K. H. (2008). Use and impact of online travel reviews. In P. O'Connor, W. Ho¨pken, & U. Gretzel (Eds.), *Information and Communication Technologies in Tourism2008* (pp. 35–46). New York: Springer

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of Interactive Marketing, 18(1),* 38-52.

Hennig-Thurau, T., & Walsh, G. (2003). Electronic word-of-mouth: Motives for and consequences of reading customer articulations on the Internet. *International Journal of Electronic Commerce, 8(2),* 51-74.

Gantz, J.,& Reinsel, D. (2011, June). *IDC iView: Extracting value from chaos.* Retrieve from http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf.

Lee, C. C., & Hu, C. (2005). Analyzing hotel customers' E-complaints from an internet

complaint forum. *Journal of Travel & Tourism Marketing, 17(2-3)*, 167-181.
Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism Management, 29(3)*, 458-468.
Mahajan, V., Muller, E., & Kerin, R. A. (1984). Introduction strategy for new products with positive and negative word-of-mouth. *Management Science, 30(12)*, 1389-1404.
McKercher, B., Shoval, N., Ng, E., & Birenboim, A. (2012). First and repeat visitor behaviour: GPS tracking and GIS analysis in Hong Kong. *Tourism Geographies, 14(1)*, 147-161.
Moore, G. A., & McKenna, R. (1999). *Crossing the chasm: Marketing and selling high-tech productsto mainstream consumers*. New York, NY: HarperCollins Publishers
Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval, 2(1-2)*, 1-135.
Pan, B., MacLaurin, T., & Crotts, J. C. (2007). Travel blogs and the implications for destination marketing. *Journal of Travel Research, 46(1)*, 35-45.
Pan, B., & Crotts, J. C. (2012). Theoretical models of social media, marketing implications, and future research directions. Social Media in Travel, Tourism and Hospitality, 46, 73-83.
Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). Servqual. *Journal of Retailing, 64(1)*, 12-40.
Park, D. H., & Kim, S. (2009). The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews. *Electronic Commerce Research and Applications, 7(4)*, 399-410.
Ruths, D., & Pfeffer, J. (2014). Social media for large studies of behavior. *Science*, 346(6213), 1063-1064.
Saaty, T. L. (1990). An exposition of the AHP in reply to the paper “remarks on the analytic hierarchy process”. *Management Science, 36(3)*, 259-268.
Shin, D., Song, J. H., & Biswas, A. (2014). Electronic word-of-mouth (eWOM) generation in new media platforms: The role of regulatory focus and collective dissonance. *Marketing Letters*, 25(2), 153-165.
Sotiriadis, M. D., & Van Zyl, C. (2013). Electronic word-of-mouth and online reviews in tourism services: the use of twitter by tourists. *Electronic Commerce Research, 13(1)*, 103-124.
Smith, Stephen L. J. (1988). Defining tourism: A supply-side view. *Annals of Tourism Research*, 15, 179-190.
The Chinese Tourism Academy. (2014, January 13). *Chinese tourist satisfaction survey (Quarter 4, 2013 and 2013 Full Year).* Retrieved from http://www.ctaweb.org/html/2014-1/2014-1-13-11-27-56041.html
Tung, V. W. S. &Ritchie, J. R. B. (2011). Exploring the essence of memorable tourism experiences. *Annals of Tourism Research*, 38(4), 1367-1386.
Stange, J., & Brown, D. (2013). *DMO Workbook: Tourism Destination Management Sustainable-achieving Sustainable and Competitive Results*. Retrieved from http://www.usaid.gov/sites/default/files/documents/2151/DMOworkbook_130318.pdf
Waldhör, K., & Rind, A. (2008). etBlogAnalysis—Mining virtual communities using statistical and linguistic methods for quality control in tourism. . In P. O’Connor, W. Ho¨pken,

& U. Gretzel (Eds.), *Information and Communication Technologies in Tourism2008* (pp. 453-462). New York: SpringerWood, S. A., Guerry, A. D., Silver, J. M., & Lacayo, M. (2013). Using social media to quantify nature-based tourism and recreation. *Scientific Reports 3, Article number: 2976.doi:10.1038/srep02976.*

Xiang, Z., Du, Q., Ma, Y., & Fan, W. (2017). A comparative analysis of major online review platforms: Implications for social media analytics in hospitality and tourism. *Tourism Management*, 58, 51-65.

Xiang, Z., & Gretzel, U. (2010). Role of social media in online travel information search. *Tourism Management, 31(2)*, 179-188.

Xiang, Z., Schwartz, Z., Gerdes, J. H., & Uysal, M. (2015). What can big data and text analytics tell us about hotel guest experience and satisfaction? *InternationalJournal of Hospitality Management, 44*, 120-130.

Ye, Q., Law, R., Gu, B., & Chen, W. (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings. *Computers in Human Behavior, 27(2)*, 634-639.

Ye, Q., Zhang, Z., & Law, R. (2009). Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. *Expert Systems with Applications, 36(3)*, 6527-6535.

Yilmaz, S.S. and Asli, D.A.T. (2013), "Internet as an information source and social distance: any relationship?". *Journal of Hospitality and Tourism Technology, 4(2)*, 188-196.

Zadeh, L. A. (1976). A fuzzy-algorithmic approach to the definition of complex or imprecise concepts. *International Journal of Man-machine Studies, 8(3)*, 249-291.