

HOTEL REVIEW MINING FOR TARGETING STRATEGY: FOCUSING ON CHINESE FREE INDEPENDENT TRAVELER

¹KIM, DONGWON, ²YU, SONG JIN

¹Ph. D, Korea Maritime and Ocean University, Department of Shipping Management, Busan, Korea

²Professor, Korea Maritime and Ocean University, Department of Shipping Management, Busan, Korea

E-mail: ¹sucher1472@hanmail.net, ²coppers@kmou.ac.kr

ABSTRACT

This study aims to construct an appropriate targeting strategy on the hotel guests who belong to a specific cultural boundary by comparing according to text analysis of 12,540 reviews written by Chinese FIT (Free Independent Traveler) and 34,330 by people from others covering 20 hotels in Seoul from Jan. 2013 to June. 2016. This study consists of three steps. In step I, in order to practice topic modeling, we examined using LDA (Latent Dirichlet Allocation). We extracted the keywords for the experiment of the next step. In step II, we investigated which words the keywords extracted from the topic model are linked with by using text-link analysis to make up for the limitations of LDA in terms of interpretation of data. Based on two steps above, in step III, we found the hidden concepts related to keywords and constructed the targeting strategy on Chinese FIT for hotel decision makers.

Keywords: *Chinese FIT, LDA (Latent Dirichlet Allocation), Targeting Strategy, Text-link Analysis, Text Mining.*

1. INTRODUCTION

In today's challenging business world, there is no doubt that the key to success depends on the ability to better understand and cope with customers' requirements. Recently, as online customer reviews have become an important factor in the purchase decision, analysis and research have been being conducted for effective online reputation management [1, 2]. As the use of customer's review data, what is called VOC (Voice of the Customer) data, in CRM (Customer Relationship Management) is regarded as a survival problem for companies, the application of VOC marketing utilizing big data is competitively being carried out. In particular, considering the empirical and the intangible nature of services in the field of tourism, online customer reviews and online e-WOM (Words of Mouth) by actual customer experience have a significant impact on service purchase and evaluation [3, 4]. In other words, when travelers use tourism services and make purchase decisions, they recognize that the reviews of user experiences are more reliable than the information provided by promotion materials or websites. As a result, service providers in the tourism sector are trying to analyze online customer reviews and effectively respond to them. Especially, since the online review contents are

difficult to be controlled or managed directly by the service provider, if it cannot be coped effectively, the negative ripple effect can be getting larger. Therefore, tourism service providers actively respond to service issues such as customer complaints, evaluations and suggestions of online customers, and utilize online communities (e.g. SNS, blogs, web sites, etc.) as channels for communication strategies.

Although the research on the influence and the ripple effect of the online VOC has been carried out in various contexts, it mainly applies the methodology such as the survey and the experimental study [3, 5]. In recent years, research on the analysis of customer reviews using Data mining has been attempted. In particular, some studies have been conducted on text mining analysis in the tourism sector, however, the researches that attempt to target to the customer who belongs to a specific cultural boundary such as Chinese culture are conspicuously lacking, considering the importance of the study. In order to meet the necessity of the research, this study aims to conduct a customer analysis study on hotel services, which is a representative tourist area targeting the Chinese FIT (Free Independent Traveler). In particular, as an international business and tourism city, Seoul, it is necessary to examine

how Chinese customers' perceptions of hotels, and which strategy can meet their needs by applying data mining analysis.

In this study, we used automated software tools designed to analyze large amounts of text and demonstrated that the hotel industry can apply the information from VOC data to develop actionable insights and improve performance. Given this, it's crucial that hotels utilize VOC information, and it can be used to better understand your customers and improve the performance of organizations. Based on VOC information, hotel decision makers can build strategies for their customers in a variety of ways. However, it can be difficult to check them all out and evaluate them. For this reason, this study undertakes a text mining approach that can extract meaningful information from a large amount of text [6]. We proposed two methodologies, LDA (Latent Dirichlet Allocation) and Text-link analysis, among the text mining methods. The LDA provides a probabilistic model of topics at the document level. In the context of text modeling, subject probability provides an explicit representation of the document. Text-link analysis is one of the text mining techniques that identify text-link pattern results or relationships between concepts.

The essence of this research is focusing on establishing a practical targeting strategy optimized for hotel management by extracting and analyzing the hidden information from the extensive volume of the customer's latest hotel data on SNS reflecting actual experiences of customers who have ever stayed at the hotel. In particular, it aims to maximize the business performance of hotel by building a hotel marketing strategy targeting customers in cultural boundaries of China, which is continuously growing.

This study consists of three steps. In step I, in order to practice topic modeling, we examined using the LDA. We extracted the keywords for the experiment of the next step. In step II, we investigated which words the keywords extracted from the topic model are linked with by using text-link analysis to make up for the fault of LDA in terms of interpretation of data. In step III, based on the results according to two-step text analysis, we established a targeting strategy for Chinese FITs related to the hotel industry in Seoul, Korea.

2. LITERATURE REVIEWS

As the importance of online customers reviews grows in management activities and operations, previous studies on the influence of

them have been conducted. In the service management field, there were also studies on online reviews, and they have been proved to be crucial factors in the process of the brand image formation and purchase decision [1-3, 7-9]. However, through the previous research, even though the significance of the online review has been verified, the research on the contents implied by the online review is relatively lacking. Recently, in terms of the development of technology related to data analysis, text data analysis researches on VOC contents are proceeding.

Kim and Oh proposed a market segmentation method by applying a text classification algorithm on the online VOC [10]. He et al. (2012) was applied text analysis to online reviews on Twitter to elicit analysis and strategies on competitors [11]. Bjorkelund et al. (2012) utilized sentiment analysis of online hotel reviews in order to categorize and evaluate the level of service by region [12]. In the precedent researches, customer reviews have been analyzed, consumer and market trend researches have been conducted, the implications for the management strategy have been deducted by deriving meaningful results. As can be seen from these preceding studies, we can confirm that the analysis of VOC can provide useful strategic guidelines for service management and tourism.

Topic modeling is a field of text mining that proposes extracting potentially meaningful topics from multiple documents based on a procedural probability distribution model [13]. In other words, it is an algorithm that stochastically computes a set of constituent words in a document and provides a set of specific topic keywords [14]. In particular, topic modeling is a method of effectively analyzing data in a large amount of text, and simultaneously constructing and presenting topics through a collection of frequently appearing words [15]. Topic modeling is more effective in terms of analyzing and extracting trends than frequency or keyword analysis methodology [16]. Jin (2013) demonstrated that topic modeling is an effective way to track changes of online trends on SNS [17]. Cha et al. (2015) applied topic modeling to suggest implications for smartphone marketing strategy [18]. Bae (2014) proposed an issue tracking system of large scale of Twitter data by using topic modeling [14]. We accept the topic modeling method, since it is suitable for extracting trends or meaningful results from large-scale data.

In the previous studies, most researches were mainly conducted to solve problems with only LDA model for topic modeling. However, these

researches have somewhat limitations when it comes to reflecting accurately the complicated and diverse customers' opinions and analyzing precisely the data to grasp their hidden intentions from the reviews.

In order to compensate the defect of LDA, therefore, in this paper, we established 5 topic domains from the reviews of hotel users by utilizing the LDA model and extracted key vocabularies related to each topic domain. In addition, by adding a text-link analysis model, we identified the overall context of the customers' opinions and overcame the disadvantage of LDA, as comparing and analyzing between the key words extracted from the topic model and the words linked with them. Based on these methodologies, we developed and suggested the targeting strategy for hotel management.

3. METHODOLOGIES

In order to analyze hotel reviews, the research was conducted through the following process (Figure 1).

To begin with, the VOC data, types of online reviews, were collected, preprocessed, and then data analysis was performed. As we mentioned above, research composed of the three-step proceeds. Keywords extracted from step I were applied to experiment on step II. Based on this process, we built the targeting strategy for Chinese FIT. Figure 1 demonstrates the three-step research flow.

To be specific, in the first step, keywords are extracted by five topics using the LDA for topic modeling. The keywords from the reviews of Chinese FITs and other foreigners are separated by subject. In the step II, Text-link analysis technique was used to extract the vocabularies, which are linked and related to the keywords extracted in the first step and to grasp the context between the keywords collected and the vocabularies. Finally, in the third step, a targeting strategy with focus on Chinese FIT was constructed.

In this study, the subjects for analysis are VOC data from "Expedia and Tripadvisor," web site (www.expedia.com and www.tripadvisor.com) related to hotels. On these sites, 12,540 pieces of guests' reviews written by Chinese FITs and 34,330 by people from other countries covering 20 hotels - The five hotels of twenty are 5-star hotel and the others are business hotel from 4 stars to 3- in Seoul, Korea, from January, 2013 to June, 2016, had been collected including titles and status information (Figure 2).

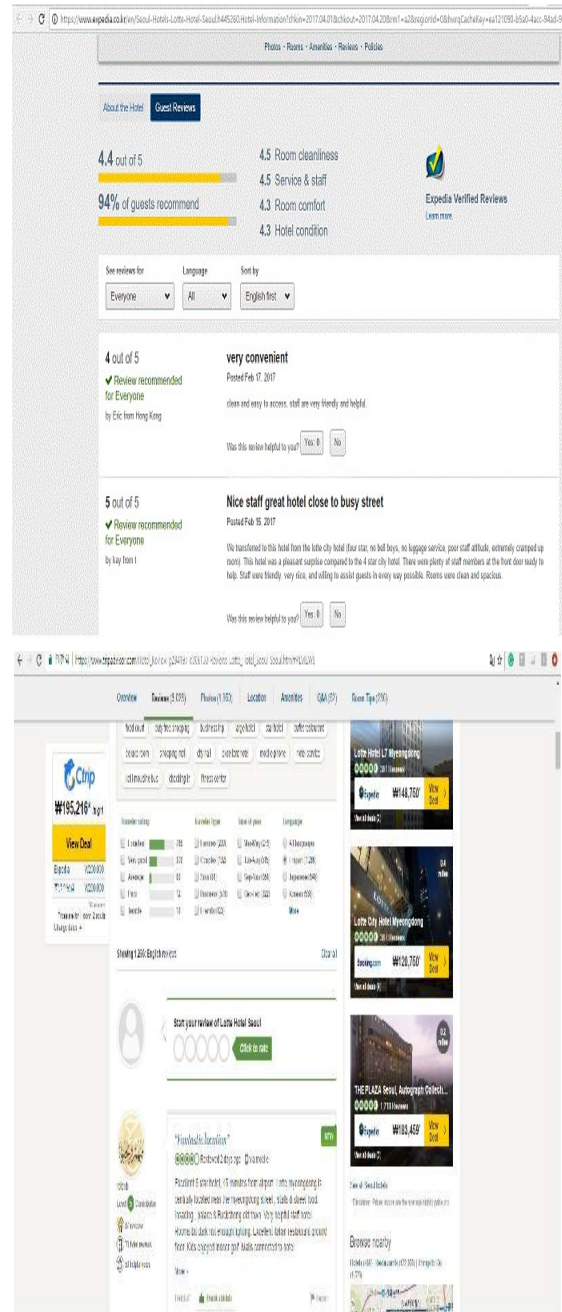


Figure 2: Sources Of VOC Data

3.1 LDA (Latent Dirichlet Allocation)

We utilized LDA for topic modeling. The topic modeling algorithm automatically encapsulates large archives of documents by discovering hidden topics or themes found in a set of documents [19]. The background of this method is that all documents share the same set of topics, but each document displays a mixed topic with a

different probability. In short, certain words are likely to be included in a particular topic. The LDA takes up a Bayesian estimation framework for text to infer the topics (word distribution) and decomposes each document into multiple topics.

The LDA deals with multinomial weights $P(z|d)$ as latent random variable for the topic. The LDA model is extended by sampling these weights from the Dirichlet distribution, the conjugate prior to the multinomial distribution. With this extension, the model allocates the probabilities to the data outside the training data and uses fewer parameters reducing over-fitting [20]. The LDA assumes the generative process for a corpus as follows:

1. Choose $N \sim \text{Poisson}(\xi)$.
2. Choose $\theta \sim \text{Dir}(\alpha)$.
3. For each of the N words w_n : (a) Choose a topic $z_k \sim \text{Multinomial}(\theta)$. (b) Choose a word w_n from $p(w_n|z_k, \beta)$, a multinomial probability conditioned on the topic z_k .

By using some generative variables to control objects of interest (documents, words, and topics), the LDA can overcome the limitations of local observation and the linear increase in the number of parameters [21]. The variable α controls the documents, β controls the words and θ controls the topics. Given the Dirichlet parameters α and β , with a topic mixture θ , a set of K topics $\{z_k\}$ and a set of N words $\{w_n\}$, we have a marginal distribution of $\{z_k\}$ of the document [21]:

$$P(d|\alpha, \beta) = \int p(\theta|\alpha) \left(\prod_{n=1}^N p(z_k|\theta) p(w_n|z_k, \beta) \right) d\theta \quad (1)$$

Each item in the collection is modeled as a finite mixture of an underlying set of topics. In turn, each topic is modeled as an infinite mixture of the underlying set of subject probabilities. In LDA, this three-level hierarchical probability model has the strength to surmount the limitations of local training set indexing and parameter over-fitting.

Our topic modeling utilized the software Topic Modeling Toolbox 0.4.0 for LDA and guidelines on parameter values by Gruen and Hornick [22] and Gibbs sampling is applied for the main computation. We decided to determine the five topics for simplicity and increased interpretability. The topics are "Location, Amenities, Price, Facilities and Experience [23]." Based on five topics, we extracted and analyzed keywords from Chinese reviews and others for the experiment on step II.

3.2 Text-link Analysis

In the next step, text-link analysis was applied to VOC. The text-link analysis identifies

text-link pattern results or relationships between the concepts [24]. That is, text-link analysis is a pattern-matching technology that enables to define pattern rules and compare these to actual extracted concepts and relationships found in the text. For example, extracting ideas about an organization may not be interesting enough. Using this technique, we could also learn about the links between this organization and other organizations or the people within an organization. Text-link analysis makes researchers be able to extract opinions on products or, for some languages, the relationships between genes.

As applied text-link analysis to this study, frequently pairs of words in VOC are identified. We chose keywords from the outputs of LDA in this analysis. It means that the software had found all combinations including keyword and other words related to the keyword. Although LDA technique plays a role to extract and classify keyword by each topic, it is impossible to identify that a keyword has a relation and meaning with other words. Text-link analysis technique was used to make up for the fault of LDA in terms of interpretation of data. We can not only use text-link analysis technique to find the words related and linked to the keywords, but also infer the contextual semantic relationships between the keywords and other words. In this study, in addition to the inference of the relation between words, the contextual semantic relationship was more clearly analyzed by actually verifying the sentences including the keyword. For this study, we applied SPSS Modeler 15.0 as an experimental tool.

4. RESULTS OF EXPERIMENTS

4.1 Step I: Topic Modeling for Extracting Keywords

The LDA topic modeling outputs posterior probability that captures the likelihood that a specific word belongs to a certain topic. Using the probability, we presented the more than thirteen most likely words in each topic and how likely it was that each word belongs to that topic as shown in table 1. Through LDA, topic modeling, five topics were derived such as "Location, Amenities, Price, Facilities, and Experience."

After selecting the five topic domains, we classified subject into Chinese FITs and other countries' FITs and analyzed key words collected in each subject. To begin with, we collected the keywords from each topic. Compared to aspect between Chinese and others, in general, even though common keywords appear, some keywords

show a different aspect. We divided the extracted keywords into common words and separate words.

Table 1: Keywords For Each Topic

Topic	Subject	Extracted Words	
		Common Words	Separate Words
Location	Chinese	metro, away, shuttle, walk, airport, walking, restaurants, street, close, location, shopping, shops	taxi, fare
	Others		comfortable, city
Amenities	Chinese	tv, ice, room, refrigerator, machine, maker, iron, dryer	Chinese, tea, discrimination, microwave, kitchen
	Others		English, floor, bar, night, buffet, coffee, smoking
Price	Chinese	rate, paid, pay, worth, price, cost, extra, charge, day, fee, cash, cheap, parking	tip, over
	Others		night, stay, deposit, credit, card
Facilities	Chinese	room, internet, business, store	free, access, speed, wireless, lobby, high, computer, available,
	Others		pool, area, tub, fitness, spa, kids, noise, swimming, indoor, nice
Experience	Chinese	service, good, street, enjoy, first, spacious, food, beautiful, perfect, outstanding	revisit, welcome
	Others		old, free

Compared two sections by subject, in the first topic, "Location," Chinese and others mentioned means of transportation like metro and shuttle from a specific place such as airport or shops for shopping to hotel. Especially, as the taxi and fare were mentioned by the Chinese, we can infer that these topics are sensitive to them. We analyzed "taxi" using text-link analysis. In the second topic, "Amenities," we focus on two words "tea and discrimination." As dealing with these words on the next step, we analyzed them by using text-link analysis. In the topic "Price," "tip" was

selected, in the topic "Facilities," "speed" was, and finally in the topic "Experience," "service," and "revisit" were chosen as keywords for text-link analysis in step II. We selected and analyzed "service" in spite of common words from the last topic "Experience", since it is one of the most crucial elements of hotel business.

In addition, we were able to compare and analyze the universal thinking of foreign tourists and the characteristic thought of Chinese FITs through LDA results. However, we focused on the Chinese FITs in this research. Thus, the comparison between a specific group and other groups will be taken into account as the next subject in order to overcome the limitations of this research.

4.2 Step II: Text-link Analysis for Targeting Strategies

The results of the text-link analysis demonstrate which words the keyword extracted from LDA are linked with as follows (Table 2).

Linked words the keyword "taxi" in the topic "Location" are composed of "fare, expensive, rip-off, and kind, etc." However, they are not associated with hotel management. So we tried to find an important keyword from "Common words." The keyword "shuttle" was selected and linked with "convenient, shopping, department, airport and free." Looking at the links, we can identify that Chinese and other guests are satisfied with the free shuttle bus from the hotel to a department store in Myung-dong or Dongdaemun for shopping.

Table 2: The Example Of Comparison Of Text-link Analysis Results

Topic	Keyword	Linked words
Location	taxi	fare, expensive, hotel, rip-off, wrong, way, kind, driver
	shuttle	convenient, shopping, department, airport, free
Amenities	tea	enjoy, free, afternoon, cocktails, good service, breakfast, impressive
	discrimination	against, Chinese, people, items, bathing, serious, afraid, slippers, take, away, feeling, bad uncomfortable
Price	tip	no, pleasant, service, included, charge, room, impressed, not, accept

Facilities	speed	internet, service, room, information, business
Experience	service	excellent, good, chicken, beer, room, warm, reception, free, shuttle, friendly, thoughtful, disappointing
	revisit	good, service, friendly, response, location, business, convenient

In the topic "Amenities," the Chinese response to "tea" demonstrates guest satisfaction and positive attitude towards the hotel. Drinking tea is one of most important eating habits of Chinese and they enjoyed and got impressed with tea service. On the other hand, comments in reference to the keyword "discrimination" dealt with a serious problem. Chinese who wrote this review felt discriminated because of the hotel's response on asking a small item for bathing and made an example about other Chinese guests that the hotel did not change wet slippers. The guest's words clearly indicate guest's disappointment and negative attitude towards the hotels as a Chinese.

In the topic "Price," we can recognize that Chinese and foreigners want to stay at a cheaper hotel from the result of LDA. Furthermore, we selected the keyword "tip." Chinese hotel review imparts that even though they had gotten a good service, hotel staffs did not require any tip and they were really impressed with it.

In the topic "Facilities," Chinese tourists are significantly satisfied with the speed of wireless internet in the room.

Finally, in the topic "Experience," two words, "service and revisit," were selected. In the middle of the various services, "fried chicken with beer delivery service" is impressing to Chinese FIT. In the review of "revisit," they posted the reasons why they would revisit the hotel. The words linked with "revisit" are related to good and friendly response to the guest, convenient business circumstance and accessibility to hotel.

Through the result of two analysis methods, we could find and face on the unexpected opinion of Chinese guests. However, these aspects give managers lessons to build a strategy and tactic for a specific group and this meaningful information is also able to have a significant impact on hotel's performance.

4.3 Step III: Building Targeting Strategies

We intended to construct and present targeting strategies to improve the hotel business

performance. In general, in order to build the targeting strategy, the managers should analyze the business circumstance and 3C (Customer, Company, and Competitor), maintain and expand the experience of target customer online, and verify the customer's needs by using the ICT (information and communication technology) analysis techniques such as data mining, and web mining. As companies practice the targeting strategy considering the business circumstance and customer experience utilizing ICT, they can establish the differentiation strategy for target guests and, based on this, secure business competitiveness. This strategy is able to be applied to any industry field in common.

In this study, based on the results of the analysis, we constructed the targeting strategy for Chinese FIT as follows (Figure. 3)

The recent Korean hotel business circumstance has rapidly changed due to the increase of Chinese tourists. This phenomenon requires managers in hotel industry to establish a Chinese custom-built strategy. They collect Chinese CX (customer experience) in the form of reviews online, which are unstructured and raw data. We utilized two text mining techniques, LDA and text-link analysis, to excavate CX after refining these type of data. This text mining process is not a simple classification or grouping process, but it is an indispensable procedure to extract the objective and hidden attributes of customer needs. That is, through text analysis, we could discover the experiences of target customers, which are that hotel managers are unpredictable and unexpected. These hidden data can be exploited as the key materials to build a targeting strategy for Chinese FIT. The results of analysis are delivered to hotel staffs and they promote improvement in service at the moment of truth. The execution of staffs on the frontline considering the hidden attributes of customer needs makes the customer evoke the satisfaction of service and induces to revisit the hotel. The analysis of CX and the execution of hotel staffs improve hotel business performance. In this study, we focused on construction of strategy utilizing the analysis of Chinese VOC. The analysis results can let the differentiation strategy be established for target Chinese FIT and be secure business competitiveness in the field of hotel industry.

5. CONCLUSIONS

Although tourism-related service is one of the representative experience goods, which is not able to be appreciated the quality without

experiencing it in person, it cannot be used to verify or do the real things in the trial due to constraints on time and space. In addition, if the wrong choice is made, since the material and psychological damage are large, the customer should try to make an optimal decision by collecting sufficient information before using it. Therefore, customers should refer to not only the information from the service provider, but also the reviews from the active users. Until now, it has been difficult to analyze the vast amount of online reviews due to technical constraints. However, recently, big data analysis techniques have been developed to analyze unstructured text data and performed the researches to find their meanings. In this study, we analyzed the user experiences of 3years and 6months of the hotel service in Seoul by using topic modeling method, LDA and text-link analysis.

The results of this study are summarized as follows. In the first step, the five domains of topics were extracted, utilizing the LDA, topic modeling technique. The topics were classified into "Location, Amenities, Price, Facilities and Experience." We classified the keywords by each topic and divided common words and separate words in Chinese reviews and other reviews. In the step II, we selected a few words used by the Chinese as the keywords and carried out the text-link analysis in order to verify the relationship between the keywords and the words linked to them. Finally, in the step III, based on the analysis processes and results, we established a targeting strategy for Chinese FIT.

We built the targeting strategy to secure the competitiveness of hotel management through differentiation strategy for target. For the strategy, unstructured data based on online reviews are collected and objective and hidden attributes to Chinese FIT's needs from them are extracted by using methodologies of ICT such as LDA and text-link analysis. Hotel managers utilize the attributes as key materials, which are applied to offline services. They promote customer satisfaction through offline service improvements and induce Chinese FITs to revisit the hotel. This has a direct impact on improving the hotel's business performance.

Through this study, it is proved that topic modeling is a useful methodology to understand customer's perception of service. Therefore, the academic contribution of this study can be summarized as follows. First, online VOC of hotel service reflects current trends, and effective political and strategic implications can be derived from VOC analysis. In other words, the validity of

text mining methodology has been verified in the study of customer's service quality perception, and it is necessary to carry out studies to use various text analysis methodologies. Second, future studies related to service quality in the hotel sector will be able to improve and supplement existing service quality parameters and measurement items in consideration of the topics derived from this study. Third, we used LDA and text-link analysis to overcome some limitations of topic modeling. Although we classify the words in the reviews into topics, we can extract the words related to the keywords by using the text-link analysis to overcome the constraints of LDA, which has limitations in verifying the relationship between words.

In addition to academic contributions, practical and policy implications can be presented. First of all, the hotel service provider can periodically analyze online reviews to identify overall customer trends and issues associated with customer service. In particular, the topic modeling method can be used to analyze a large amount of text data relatively easily and make comparisons with competitors easier. In addition, through the text-link analysis used to supplement the constraints of the LDA technique, we can identify more detailed customer needs and improve unpredictable service elements. Furthermore, the targeting strategy can be used in marketing and promotional strategies. For example, in addition to universal services, we will be able to take advantage of customized service offerings, promotions and marketing tailored to specific target customers. Finally, we will be able to change the online customer management centered on short-term response to complaints into the sustained customer relationship management by identifying the determinants of service quality via text analysis, which is not recognized by the service provider as well as the customer.

We have tried to secure rigidity in the methodology and the research process, but there are a few limitations and it is necessary to complement them in future studies. First of all, although topic modeling has the advantage of extracting issues or topics effectively based on keywords, there are disadvantages, which are that it is impossible to verify the causality and score the degree of positive / negative intensity due to the limitations of the algorithm. In future studies, causality analysis combined with various methodologies such as sentiment analysis and network analysis will lead to more specific and useful implications. In addition, as we mentioned above, we were able to compare

and analyze the universal thinking of foreign tourists and the characteristic thought of Chinese FITs through LDA results. However, we focused on the Chinese FITs in this research. Thus, the comparison between a specific group and other groups will be taken into account as the next subject in order to overcome the limitations of this research. There is another problem with the generalization of research results because of analyzing 20 hotels only in Seoul. We will continuously expand the area and the target for experiment in the future to secure the better strategy for business implication.

REFERENCES:

- [1] S.J. Doh, and J.S. Hwang, "How Consumers Evaluate eWOM (Electronic Word of Mouth) Messages", *CyberPsychology & Behavior*, Vol. 12, No. 2, 2009, pp. 193-197.
- [2] Y.W. Fan, and Y.F. Miao, "Effect of Electronic Word-of-mouth on Consumer Purchase Intention: The Perspective of Gender Differences", *International Journal of Electronic Business Management*, Vol. 10, No. 3, 2012, pp. 175.
- [3] A.G. Mauri, and R. Minazzi, "Web Reviews Influence on Expectations and Purchasing Intentions of Hotel Potential Customers", *International Journal of Hospitality Management*, Vol. 34, 2013, pp. 99-107.
- [4] R. Law, D. Buhalis, and C. Cobanoglu, "Progress on Information and Communication Technologies in Hospitality and Tourism", *International Journal of Contemporary Hospitality Management*, Vol. 26, No. 5, 2014, pp. 727-750.
- [5] M.D. Sotiriadis, and C. Van Zyl, "Electronic Word-of-mouth and Online Reviews in Tourism Services: The Use of Twitter by Tourists", *Electronic Commerce Research*, Vol. 13, No. 1, 2013, pp. 103-124.
- [6] E. Turban, R. Sharda, and D. Delen, *Decision Support and Business Intelligence Systems* (9th Ed.), Upper Saddle River, NJ: Prentice Hall, 2010.
- [7] T.W. Gruen, T. Osmonbekov, and A.J. Czaplewski, "eWOM: The Impact of Customer-to-customer Online Know-how Exchange on Customer Value and Loyalty", *Journal of Business Research*, Vol. 59, No. 4, 2006, pp. 449-456.
- [8] M.R. Jalilvand, and N. Samiei, "The Impact of Electronic Word of Mouth on a Tourism Destination Choice: Testing the Theory of Planned Behavior (TPB)", *Internet Research: Electronic Networking Applications and Policy*, Vol. 22, No. 5, 2012, pp. 591-612.
- [9] Z. Zhang, Q. Ye, R. Law, and Y. Li, "The Impact of e-word-of-mouth on the Online Popularity of Restaurants: A Comparison of Consumer Reviews and Editor Reviews", *International Journal of Hospitality Management*, Vol. 29, No. 4, 2010, pp. 694-700.
- [10] G.-H. Kim, and S.-Y. Oh, "Methodology for Applying Text Mining Techniques to Analyzing Online Customer Reviews for Market Segmentation", *International Journal of Contents*, Vol. 9, No. 8, 2009, pp. 272-284.
- [11] W. He, S. Zha, and L. Li, "Social Media Competitive Analysis and Text Mining: A Case Study in the Pizza Industry", *International Journal of Information Management*, Vol. 33, No. 3, 2013, pp. 464-472.
- [12] E. Bjorkelund, T.H. Burnett, and K. Norvag, "A Study of Opinion Mining and Visualization of Hotel Reviews", *14th International Conference on Information Integration and Web-based Applications & Services*, 2012.
- [13] J.K. Ahn, K.H. Lee, and H.W. Kim, "Trend Analysis of Intelligence System Studies: Topic Modeling and Networks Analysis", *The Korea Society of Management information*, 2015, pp. 561-570.
- [14] J.-H. Bae, N.-G. Han, and M. Song, "Twitter Issue Tracking System by Topic Modeling Techniques", *Journal of Intelligence and Information Systems*, Vol. 20, No. 2, 2014, pp. 109-122.
- [15] M. Steyvers, and T. Griffiths, "Probabilistic Topic Models", *Handbook of Latent Semantic Analysis*, Vol. 427, No. 7, 2007, pp. 424-440.
- [16] W.-J. Ryu, J.W. Ha, Md. Hijbul Alam, and S.K. Lee, "Extracting Trends from Twitter using a Topic Modeling Technique", *40th Conference of Computing Science and Engineering*, 2013, pp. 191-193.
- [17] S.A. Jin, C.E. Heo, Y.K. Jeong, and M. Song, "Topic-Network based Topic Shift Detection on Twitter", *Journal of the Korean Society for Information Management*, Vol. 30, No. 1, 2013, pp. 285-302.
- [18] Y.J. Cha, J.-H. Lee, J.-E. Choi, and H.W. Kim, "A Topic Modeling Approach to Marketing Strategies for Smartphone Companies", *Knowledge Management Research*, Vol 16, No. 4, 2015, pp. 69-87.

- [19] D. Blei, and J. Lafferty, "*Topic Models, in Text Mining: Classification, Clustering, and Applications*", ed. A.N. Srivastava, & M. Sahami, CRC Press, 2009, pp. 71-94,.
- [20] B.C. Russell, A.A. Efros, J. Sivic, W.T. Freeman, and A. Zisserman, "Using Multiple Segmentations to Discover Objects and their Extent in Image Collections", *Proceedings of CVPR*, 2006.
- [21] T. Kakkonen, N. Myller, and E. Sutinen, "Applying Latent Dirichlet Allocation to Automatic Essay Grading", *FinTAL 2006, LNAI 4139*, 2006, pp. 110– 120.
- [22] B. Gruen, and K. Hornik, "Topicmodels: An R Package for Fitting Topic Models", *Journal of Statistical Software*, Vol 40, No. 13, 2011, pp. 1-30.
- [23] K. Berezina, A. Bilgihan, C. Cobanoglu, and F. Okumus, "Understanding Satisfied and Dissatisfied Hotel Customers: Text Mining of Online Hotel Reviews", *Journal of Hospitality Marketing & Management*, Vol. 25. No. 1, 2015, pp. 1-24.
- [24] IBM. "IBM SPSS Modeler Text Analytics 15 User's Guide", 2012.
- [25] scholarship.sha.cornell.edu.

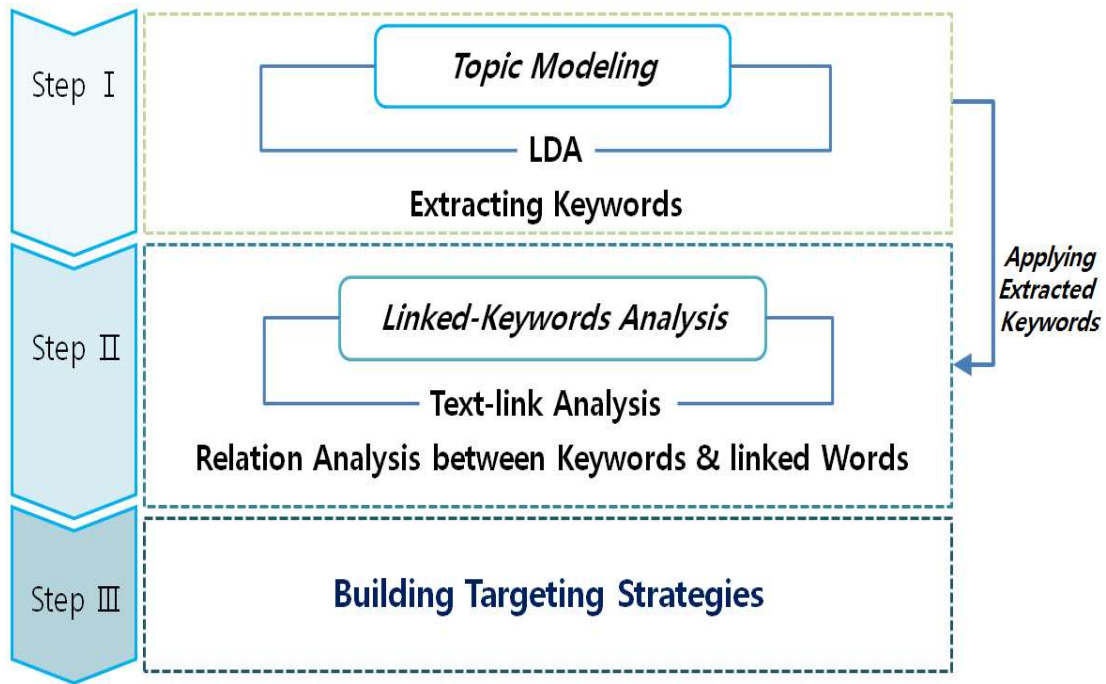


Figure 1: Research Flow

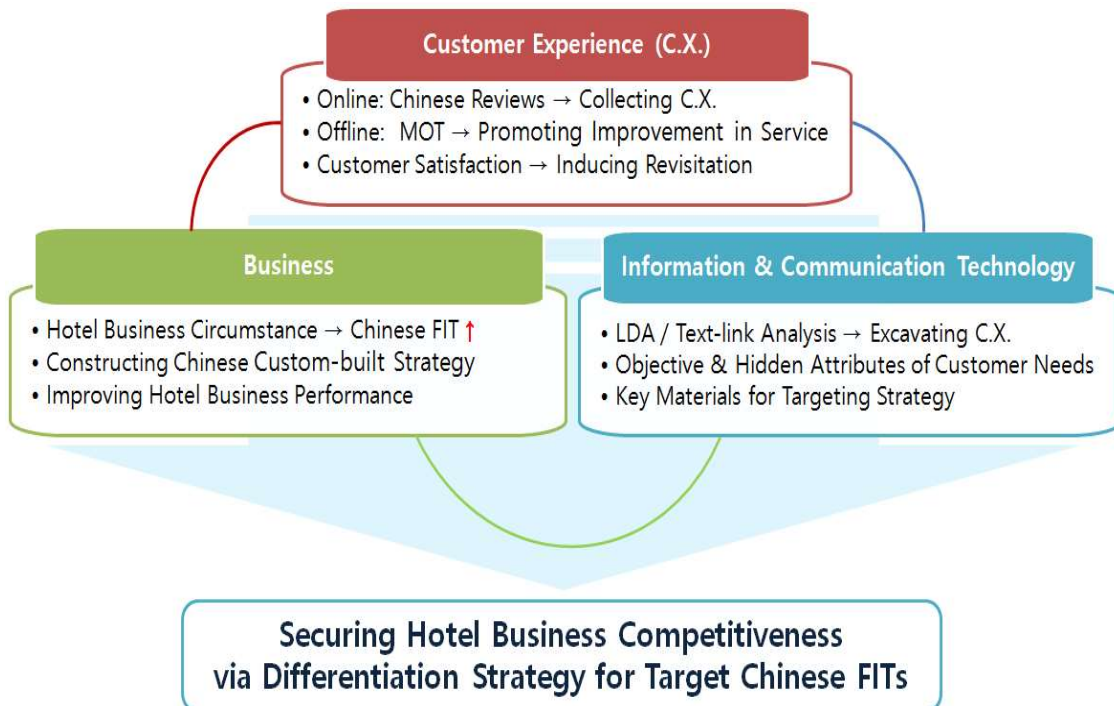


Figure 3: Targeting Strategy For Chinese FIT