Journal of Theoretical and Applied Information Technology
30th November 2018. Vol.96. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

CTQ FOR SERVICE QUALITY MANAGEMENT USING WEB-BASED VOC: WITH FOCUS ON HOTEL BUSINESS

[1]DONGWON KIM
[1]Professor, Korea Maritime & Ocean Univesity, Liberal Eduction, Busan, S. Korea
E-mail: [1]sucher1472@kmou.com

ABSTRACT

This study aims to derive CTQ(Critical-to-Quality) for service quality management strategy optimized to the hotel service. In order to derive CTQ, text mining analysis of 46,870 reviews written by foreign visitors covering 20 hotels in Seoul, Korea, is conducted from Jan. 2014 to Jun. 2017. In this study, KJ(Kawakita Jiro) method applied by text mining technique, LDA(Latent Dirichlet Allocation) for topic modeling, is examined and raw VOC data were translated to CTQ characteristics. In addition, the measurement of CTQ is practiced based on sentiment analysis related to keywords from topics in contrast to typical CTQ measurement. Finally, it examined what CTQ measured influences on the hotel business performance through the linear regression.

Keywords: *CTQ, LDA, Sentiment Analysis, Service Quality Management, Text Mining, VOC.*

1. INTRODUCTION

Recently, with the spread of smart devices and the introduction of various information channels, the amount of information generated, distributed, and stored is exponentially increasing. Thus, the interest and the research in the big data analysis technology[1] are growing. Big data is characterized by the use of existing methods and tools, which means regular or unstructured data that are difficult to collect, store, search, analyze, and visualize[2].

The VOC(Voice of the Customer) analysis of social media in the midst of big data analysis has the advantages of being able to proactively cope with the changeable customers' needs in that they can grasp more objective opinions in more real time in comparison with traditional questionnaires and interviews. In addition, it can be said that the utility value is very high because it can collect key information at a lower cost. Therefore, it is natural that demand and interest are concentrated on text mining to analyze the opinions of users in a huge amount of data.

This study deals with the task of CTQ(Critical to Quality) which is applied to service process design which is an area of Six Sigma management. In order to overcome the practical difficulties arising from the process of deriving CTQ, which is essentially designed from CCR(Critical Customer Requirement), text mining techniques were utilized. The purpose of this study is to extract CTQ required for solving the problems of service companies from VOC of social media and grasp how CTQ measured influences on the hotel business performance. In this study, in order to solve the service problems to the customer, KJ(Kawakita Jiro) method(affinity diagram) applied by text mining technique, LDA(Latent Dirichlet Allocation) for topic modeling, is examined. LDA is a topic extracting technique, which is used to translate raw VOC data into CTQ by analyzing the extracted topic words and compensate the defect of KJ method.

This research examined by applying the positive and negative sentiment intensity of the keywords related to CTQ without using traditional methods such as benchmarking or mystery customer technique in order to measure CTQ. This suggests that this aspect can be a basis for judging whether or not to implement CTQ, because if the implementation of CTQ is smooth, the amount of positive emotions will increase in the opinions of customers at that moment, or if not, negative emotions will increase.

This paper is organized as follows. In section 2, general CTQ and application to service industry are described. In section 3, methodologies for this research are introduced based on the research flow. The section 4 demonstrates the experimental results of applying the methodologies. Finally, the managerial implications and limitations of this study are suggested in the last section.

2. GENERAL CTQ & APPLICATION TO SERVICE FIELD

The key to the quality management process is the transition from customer needs to technical quality characteristics. This is a viewpoint of what quality characteristics should be designed by the customer in terms of product and process design, and also whether the design quality characteristics are achieved at the time of design completion. In addition, when the project is completed and these quality characteristics reach the required level, the requirements of the customer must be clearly met. Therefore, if the CTQ is not explicitly derived from the CCR, the content of the future design will be meaningless, and even if CTQ is clearly derived, if CTQ is not able to be measured, it is difficult to confirm success after completion of the project.

In the process of development of service and product, in particular, it is more difficult to derive the CTQ from the CCR in the service area than manufacturing due to the characteristics of the service sector. The first is relatively easy to choose because there are a number of measurement items and methods for pre-selected quality attributes - such as tensile strength, tear strength, viscosity- that are pre-selected in the manufacturing sector. In the service sector, however, not only is there a lack of awareness on the importance of service measurement, but also new quality characteristics need to be derived frequently and deriving is difficult. The second is that the use of quantitative and statistical methods to evaluate and verify the degree of correlation between CCR and quality characteristics is lacking, even though the quality characteristics corresponding to CCR are derived. From this point of view, it is difficult to develop CTQ from the CCR, which is useful in the design sector of the service company like a hotel.

In spite of the importance of deriving CTQ in the service industry, the research on CTQ is still lacking. It was practiced to derive CTQ for DFSS (Design for Six Sigma) in service areas by using factor analysis with interview data[3]. Interviews were conducted to identify customer satisfaction requirements, using CTQ tree[4]. In order to select the main CTQ data which become the subject of quality management in the organization and manage them systematically, management method of integrated CTQ data was presented in detail, which establish the criteria of CTQ data in terms of performance and IT and quantify the CTQ data according to the criteria.

3. METHODOLOGIES

This study is conducted following research flow as Figure 1. Schematic contents present as follows.

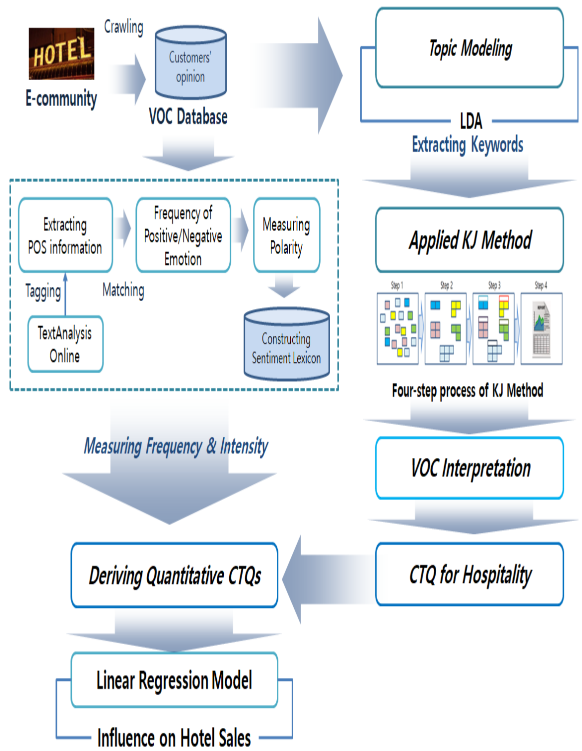

Figure 1: Research Flow

First of all, this research consists of two-track process. One process is LDA to make up for KJ Method so as to interpret VOC into CTQ, and the other is a sentiment analysis process, which constructs the sentiment lexicon from raw VOC crawled from social media, and measures the amount of sentiment intensity on the keywords related to CTQ. Finally, we examined how CTQ measured influences on the hotel business performance through the linear regression.

KJ Method also called the "Affinity Diagram" is a four-step process to capture, organize and summarize a large amount of data, such as ideas, issues, problems, and solutions into related groups after brainstorming in order to be able to understand the essence of a problem or a solution. The KJ method is useful for sorting and organizing many pieces of uncertain information into logically cohesive groups. The goal is to construct a limited number of groups. The result makes a better

Journal of Theoretical and Applied Information Technology
30th November 2018. Vol.96. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

understanding of the idea selection or problem. The following Figure 2 is the four-step process of KJ method[5].

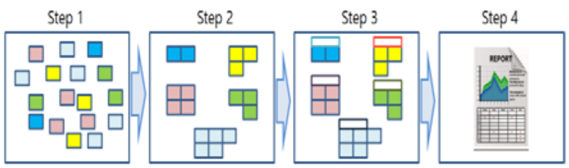

Figure 2: Four-step Process of KJ Method

Step 1 is a phase to collect data and prepare notes. In this step, a performer prepares some index cards or cohesive notes for use. After that, he or she writes down each idea for each card.

In Step 2, ideas are sorted into related groups. First of all, put all cards into one pile. Pick one card at a time from the pile. If the card contained an idea related to be already selected is picked up, lump it together into the group. Keep picking up the cards one at a time and examine that the card containing the idea is similar to an existing group. Keep doing this process until all cards are picked up and all of the ideas are classified into proper groups. After all cards are picked, if an idea is applicable equally to other groups, make a duplicate of the card and put one in each group. In addition, it is possible for one card to place alone and build a group. As a result of grouping, the ideas should be closely related and significant differences should exist between groups.

Step 3 is a phase of creating a header card for the group. A header is a title of a group which essentially links among the ideas included in a group of cards. The header should be the most appropriate word or phrase that depicts the meaning of each group. The meaning of the header should be clear to readers without reading the contents of the cards in the group. It is also possible that a high group has several small groups under different headers. It means that hierarchical or multilevel groups can be adopted.

In step 4, the performer writes reports and practices further analysis. Affinity diagram can help discover hidden groups and structures in huge amount of data composed of a variety of notes and words.

LDA for topic modeling is used to supplement the KJ method. The LDA algorithm automatically encapsulates large archives of documents by discovering hidden topics or themes found in a set of documents[6]. In the process of KJ method, analysts for VOC analysis naturally exchange their ideas while classifying ideas. In particular, they write down the names of the groups and exchange ideas about commonalities and differences between ideas. When excluding the ideas that they think are irrelevant to the subject, the criteria of classification may conflict in classifying the ideas. In order to mitigate this aspect, it is usually the case that a small number of attendees first classify the VOC by using the KJ method, and then the rest of them review the classification results and raise objections. However, in this process, conflicts of individual opinion may cause mis-classification of the data. In this study, LDA technique was used to prevent such mis-classification and integrate the process of the KJ method. The five topics required in the service quality management, which can be headers in KJ method, are selected by the precedented studies[7]. They are composed of "Operations, Sales, Amenities, Facilities and Experiences," important elements in assessing hotel service quality[7].

In this study, the KJ method was conducted in reverse order of steps. This is able to be called the applied KJ Method. Five topics that become headers in KJ Method were pre-established as told above, key words were extracted using LDA, topic modeling technique, and sentences linked with the key words of five topics were also extracted in order to grasp the accurate intention of customers who posted the opinion VOC was converted to CTQ after the process of interpretation.

The LDA assumes the generative process for a corpus as follows[8]: 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 Multinomial(θ). (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 get over the limitations of local observation and the linear increase in the number of parameters[8]. 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[8]:

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

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

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.

For sentiment analysis, the database is constructed crawling customers' opinion information from the hotel social media, and the part of speech(POS) information about vocabulary in the customers' reviews is identified by utilizing a POS tagging function offered by TextAnalysisOnline. The negative and positive emotional vocabulary group, sentimental lexicon, was built. Based on the data, the polarity and intensity of keywords from CTQ are measured and used to grasp what CTQ measured influences on the hotel business performance through the linear regression.

4. EXPERIMENT & ANALYSIS

The subjects for analysis are hotel review data. They were collected from "Expedia and Tripadvisor(www.expedia.com,www.tripadvisor.com)." (Figure 3)

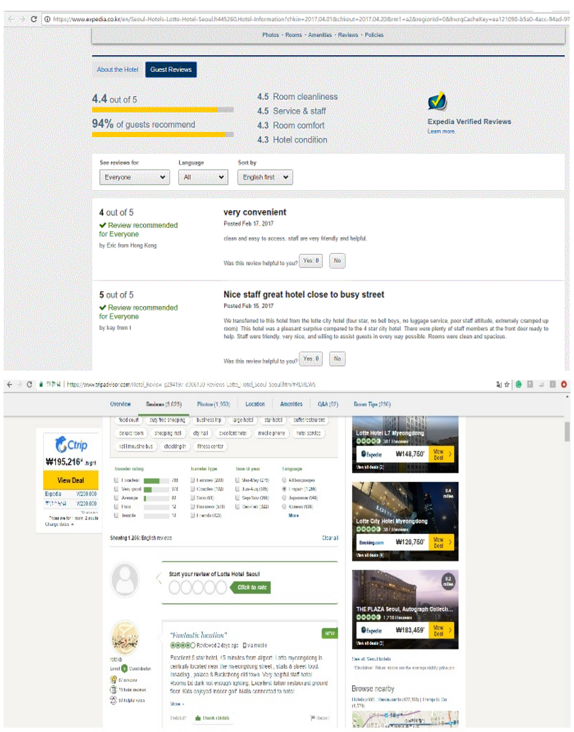

Figure 3: Sources of VOC

46,870 reviews written by foreign travelers covering 20 hotels in Seoul, Korea from Jan. 1. 2014 to Jun. 30. 2017 including titles and status information were used to conduct the topic analysis. In the middle of 20 hotels, the five hotels are five-star and the others are business hotel. The reason why fifteen business hotels are chosen is 34% of foreign travelers have stayed at the business hotels[9].

LDA topic modeling outputs posterior probabilities that capture the probability that a particular word belongs to a particular topic. Each topic presented probability of using more than 15 most likely keywords. Table 1 shows the probability that each word in the five topics (Operations, Sales, Amenities, Facilities and Experiences) through LDA belongs to the topic. "Topic Modeling Toolbox 0.4.0" for topic modeling was used and topic modeling guidelines on parameter values by Gruen and Hornick[10] and Gibbs sampling is applied to the main computation. The five topics are fixed for simplicity and interpretability improved.

Table 1: Extracted Keywords by Topic

Topic	Extracted Words
Operations	reservation, front, check-in, check-out, information, service, time, long, housekeeping, delayed, billing, waiting, error, comfortable, quick, deposit, clean.
Sales	website, rate, paid, coupon, sale, worth, price, cost, extra, charge, day, fee, cash, cheap, tip, reward, discount, code, night, credit, card, season.
Amenities	Microwave, tv, ice, room, refrigerator, machine, maker, iron, dryer, English, floor, bar, night, buffet, tea, coffee, smoking, non-smoking.
Facilities	wifi, room, internet, business, store, free, access, speed, wireless, high, computer, available, pool, fitness, spa, kids, noise, swimming, indoor, nice
Experience	service, good, food, street, enjoy, first, spacious, beautiful, perfect, play, outstanding, revisit, welcome, old, culture.

Based on the results of LDA, in the next step, hotel "S" is selected in order to derive CTQs for a case study. It is a 5-star hotel in Seoul and selected for measuring CTQ and linear regression analysis of the next step. CTQ must be met clear, specific and quantitative requirements from the raw VOC data[11].

Journal of Theoretical and Applied Information Technology
30th November 2018. Vol.96. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

After LDA, the process of deriving CTQ is that there is an interpretation process in the middle of process to derive CTQ from raw VOC data. Table 2 shows an example of translating VOC to CTQ for hotel operations.

Table 2: Translating VOC to CTQ for Hotel Operation(Example)

VOC	Interpretation	CTQ
It took too long time to check-in.	Short Wait	Waiting Time
Hotel didn't respond quickly to my asking to get slippers.	Quick Response from housekeeping in the room.	Response Speed of Housekeeping
There were too many errors on my bill.	Billing Accuracy	Percentage of Erroneous Invoice

Since CTQ should have measurable characteristics, the qualitative characteristics of service-related CTQ should be converted to quantitative characteristics. In this study, sentiment analysis was conducted on key words related to CTQ to convert CTQ into quantitative and measurable characteristics. The relation between the change of emotion on the extracted words and the sales of hotel company is compared and analyzed.

A word has various meanings based on POS(Part-of-Speech) in the sentence. For example, as the word "good" is used, As we use it as an adjective or a noun, the polarity of the each positive or negative meaning is also different. Thus, it is necessary to extract the POS information in order to analyze the positive or negative polarity with respect to customers' reviews. In this study, POS tagging tool from "www.TextAnalysisOnline.com" was taken advantage of in order to extract the POS information of each word. TextAnalysis API(Application Programming Interface) provides customized Text Mining Services such as POS Tagging, Stemmer, Word Tokenize, Lemmatizer, Chunker, Parser, Key Phrase Extraction(Noun Phrase Extraction), Sentence Segmentation (Sentence Boundary Detection), Sentiment Analysis, Text Summarizer, Grammar Checker, Text Classifier and other Text Analysis Tasks. It is so helpful of NLP Tools, such as TextBlob, Pattern, NLTK, MBSP and so on. TextAnalysisOnline displays the extracted POS information. Figure 3 describes the analysis result of extracting POS information from the reviews posted by customers.


Analysis Result

favorite|JJ doc|NN author|NN car|NN great|NN love|IN vehicle|NN seats|NNS good|JJ power|NN chrysler|NN ride|NN interior|NN mileage|NN drive|VBP gas|NN mpg|NN comfortable|JJ miles|NNS system|NN driving|VBG engine|NN nice|NN pacifica|NN radio|NN bought|VBD room|NN highway|NN smooth|NN rear|NN hemi|NN seat|NN suv|NN comfort|NN features|NNS fun|VBP wife|NN handling|NN price|NN style|NN van|NN performance|NN row|NN feel|NN speed|NN exterior|NN steering|VBG wheel|NN front|NN handles|NNS lot|NN dvd|NN heated|VBD problems|NNS road|NN owned|VBN pt|NN styling|NN buy|NN purchased|VBD sound|NN years|NNS auto|NN cars|NNS design|NN fuel|NN quiet|NN controls|NNS easy|JJ kids|NNS trip|NN cruiser|NN dealer|NN seating|VBG big|JJ model|NN control|NN touring|VBG inside|NN leather|NN plenty|NN quality|NN transmission|NN bad|JJ fit|NN dash|NN drove|VBD excellent|NN people|NNS problem|NN stereo|NN top|NN town|NN awd|NN warranty|NN door|NN drives|NNS happy|VBP lots|NNS luxury|NN months|NNS sebring|VBG city|NN loved|VBD mph|NN aspen|VBN size|NN year|NN camera|NN replaced|VBD slow|NN tires|NNS limited|VBN options|NNS side|VBP air|JJ driver|NN feature|NN find|NN lifetime|NN money|NN roomy|NN wheels|NNS american|JJ driven|NN finish|JJ high|NN motor|NN nav|NN satellite|NN small|JJ times|NNS wanted|VBD automatic|JJ family|NN feels|NNS real|JJ rides|NNS traded|VBN trips|NNS body|RB chrome|VBP economy|NN low|JJ player|NN purchase|NN sedan|NN snow|NN solid|VBD trunk|NN brakes|NNS cargo|VBP convertible|JJ head|NN large|JJ needed|VBN package|NN safety|NN sirius|NN vehicles|NNS adults|NNS color|VBP doors|NNS light|JJ loaded|VBN noise|NN space|NN thing|NN thought|VBD acceleration|NN dream|NN liftgate|NN pretty|RB bit|NN cool|NN cylinder|NN day|NN hard|RB loves|VBZ minivan|JJ option|NN recommend|NN roof|IN start|NN windows|VBZ works|NNS awesome|VBP backup|NN base|NN boston|NN console|NN country|NN cruise|NN docno|NN ease|NN extremely|RB lease|VB series|NNS tank|NN test|NN things|NNS wonderful|JJ absolutely|RB average|JJ compliments|NNS crossfire|VBP cup|NN ford|NN lexus|VBZ mirrors|NNS passenger|JJR st|NN work|NN wrong|IN beautiful|JJ days|NNS dealership|NN decent|NN decided|VBD expected|VBN extra|JJ fold|NN friends|NNS gps|NNS lbs|NNS making|VBG pleased|VBN red|VBN remote|NN replace|NN service|NN signature|NN standard|NN storage|NN worth|NN dodge|NN durango|NN easily|RB fantastic|JJ fast|NN forward|NN inch|NN jeep|NN job|NN leg|NN mini|NN poor|NN read|NN set|NN started|VBD stylish|JJ suspension|NN turn|NN adjustable|JJ beat|NN blue|NN brake|NN buying|VBG classy|NN close|NN cost|NN fine|NN firm|NN free|JJ hit|NN holders|NNS hp|VBP liter|NN market|NN oil|NN paid|VBD perfect|NN pickup|NN powerful|JJ reliable|JJ sporty|NN stick|NN sunroof|IN touch|JJ towing|NN toyota|NN truck|NN true|NN weeks|NNS wood|VBD belts|NNS capacity|NN center|NN check|NN climate|NN cyl|NN designed|VBN dual|JJ edition|NN flat|JJ gate|NN gear|NN hatch|NN

Figure 3: An Example of POS tagging

NN (Noun) / W (Word) / VG (Verb Group) presents an attribute of POS classification. There are 30 sorts of the POS classification like PRP(Personal Pronoun), JJ(Adjective), NNS(Noun Plural), VBN(Verb, Past, Participle). The POS information identified in this way from each customer's opinion was saved in the database. The POS tagging classifies POS of words, and extracts noun and adjective vocabulary presented NN and JJ in the middle of the words arranged.

Lexicon of positive and negative emotion based on customers' reviews is able to be established. By applying POS information extracted from words related to CTQ and extracting synonyms of words with WordNet, which offers the relation between words such as synonym and antonym, we built seed words and expanded them(Table 3). As the next step, the polarity of the expanded seed words from SentiWordnet 3.0 is collected and the sentimental intensity of each word is computed. Thus, the sentiment lexicon could be built by selecting 2,815 negative sentiment words and 1,021 positive sentiment words. Table 3 displays a part of lexicon expanded from seed words.

Journal of Theoretical and Applied Information Technology
30th November 2018. Vol.96. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 3: A Part of Lexicon of Negative & Positive Domain.

Domain	Seed words	Lexicon expanded
Negative	sadness, loneliness, unhappiness, depression, shame, regret and etc.	sad, blue, grief, sorrowful, mournful, lonely, lonesome, lone, depressed, unhappy, humiliation, humiliated, mortified, ashamed, and etc.
Positive	joy, pleasure, delight, satisfaction, happiness, fulfillment, and etc.	enthusiastic, pleased, classy, pleasant, enjoyable, awesome, happy, ecstatic, balanced, spirited, comfy, agile, and etc.

Based on the lexicon, even if they are classified into the same range of sentiment, the sentiment polarity of each word is different from one another. SentiWordNet 3.0 provides the polarity information indicating the sentimental intensity for each word according to POS and usage, and also offers an index information indicating the frequency of use[12].

Table 4: Polarity Information of "sad."

Word	POS	Index Value	Positivity	Negativity
sad	Adj.	1	0.125	0.75
		2	0	0.25
		3	0	1

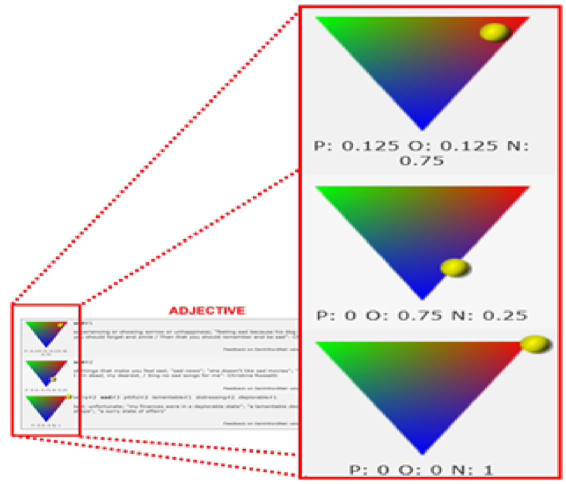

Figure 4: Polarity Information from SentiWord Net 3.0.

First of all, the difference between positivity and negativity of each meaning is calculated respectively, and then the index values given to its meaning are multiplied. The higher the index value is, the greater weight is assigned. The total sum of each word is normalized. That is, for instance, in the case of "sad", the value of the negativity was extracted using this example of the formula[13].:

{(0.125-0.75)*1+(-0.25)*(1/2)+(-1)*(1/3)}/{(1+(1/2)+(1/3)} (2)

By conducting text matching both the extracted values of negativity and positivity and lexicon built by the process above, sentiment intensity of CTQ is measured. Based on this process, the sentiment lexicon was constructed selecting 2,815 negative and 1,021 positive sentiment words and built the sentiment domains. It is meaningful that the sentiment lexicon related to the hotel industry is constructed.

In the midst of hotels, "S" hotel was chosen for a case study. "S" hotel is located in Seoul, Korea as a five-star hotel and listed on the stock market, KOSPI. 1007 reviews were crawled and analyzed from "Expedia and Tripadvisor" from Jan. 1. 2014 to Jun. 30. 2017.

Table 5: Customers' Reviews of "S" Hotel.

Company	2014	2015	2016	2017	Total
"S" Hotel	224	355	303	125	1007

In the table 6, it presents basic descriptive statistics on the amounts of one month sentiment

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

change during three years and six months. The amount of one month sentiment change (S_t) is measured according to this following formula[14].:

$$S_t = \ln(E_t / E_{t-1}) \qquad (3)$$

In this formula, E_t is the frequency of this month emotion and $E_{t-1}$is the frequency of the last month's emotion. The reason of using the amount of sentiment change as a variable is that the distribution of sentiment shows a similar aspect[15]. The similar distribution of sentiment responds the change varied such as sales. Furthermore, because the number of posting presents a large difference depending on the product or brand, the amount of sentiment change is proper for variable.

The average of negativity of hotel review is having a positive value, while that of positivity has a negative value. It is able to be ascertained that the sentiment appropriately divided as the average of sentiment of other propensity has the opposite sign to each other (Table. 6).

Table 6: Basic Descriptive Statistics.

Variables	Min.	Max	Aver.	St. Dev
Negativity	-1.7064	1.1536	0.001241	0.1632425
Positivity	-0.8562	0.7188	-0.000759	0.1214652

Based on the materials above, linear regression was conducted. In the research model, what VOC, which is consumer reviews on the hotel services, has an impact on the hotel sales was examined through the linear regression. The independent variables were converted to the quantified sentiment of texts of five areas from CTQ of "S" hotel by applying the opinion mining technique. The regression model is graphically represented to explain influence on the hotel sales on VOC digitizing sensibility(Figure 5).

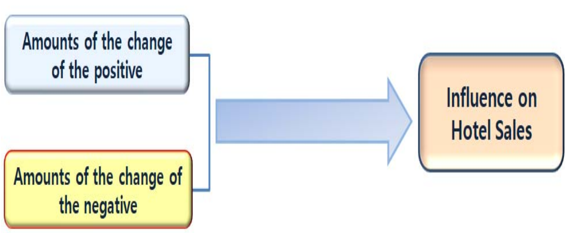

Figure 5: Linear Regression Model.

In case of "S" hotel, it has been one of the top hotel organization according to the Korea Hotel Association. The results from regression investigating the impact on the hotel sales in accordance with the change of sentiment from the words related to CTQs are stated as follows.

R^2 means statistics, which presents how much the independent variables explain the dependent variable. In table 7, since R^2 is 0.374, two sentiment domains can explain the hotel sales approximately 37.4%. As R^2 adjusted is 0.367, it shows a similar level to R^2. As Durbin-Watson(DW) value presents from 1 to 3, which is unproblematic to independency of residual. As the DW value is 1.423, the condition of the independency of residual is met. In ANOVA, investigating P-value(significant probability) on F-value, as P-value is 0.018, it is appropriate for the regression model.

Table 7: The Result of Linear Regression Model.

R^2	R^2 Adj.	DW	ANOVA F	ANOVA P-v.	Variables	Stdzd. Coeffi B	t	sig.
					(constant)		12.819	0.000
0.374	0.367	1.432	3.119	0.018	Negativity	**-0.482**	-2.425	**0.016**
					Positivity	**0.216**	-2.430	**0.043**

As the result of significancy of each variable, P-value is shown to respectively 0.016 in Negativity, and 0.043 in positivity. The results represent that the aspects of two sentiment domains have a significant influence on "S" hotel's business performance. In addition, as seeing the standardized coefficient, The sentiment of negativity had negative (-) influence on hotel sales and that of positivity had a positive (+) impact on hotel performance, but it is a little meager level(0.216).

Finally, the comparison between occupancy rate and sentiment was conducted. As shown as Figure 6, as a result of this analysis, the emotion of the VOC reviews related to CTQs in the last quarter had an effect on the occupancy rate of hotels in the next quarter. Especially, the intensity of negative emotion has more influence on the occupancy rate than that of positive emotion. As a result of the linear regression above, the change of negative emotion has a greater influence on the hotel business performance was confirmed.

Journal of Theoretical and Applied Information Technology
30th November 2018. Vol.96. No 22

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

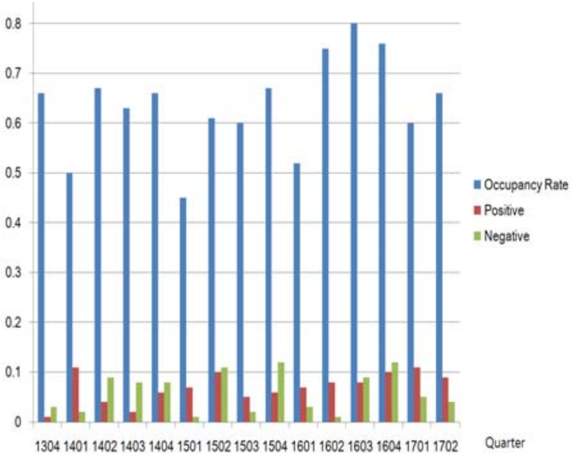

Figure 6: Comparison between Occuoancy Rate & Sentiment by Quarter

The results of the regression analysis present that negative sentiment is affecting the hotel's management performance more. However, as the results of the comparison analysis between occupancy rate the positive and negative sentiment change, it shows the aspect of that when positive feedbacks were growing(Q1 2014 and Q2 2014), the occupancy rate was increased by a larger margin. In addition, as both positive and negative emotions are increasing simultaneously(Q1 2015 and Q2 2015), the occupancy rate tends to decrease. Therefore, although positive opinions also have a strong influence on customers' choice of hotel, in case that positive and negative emotions interact with, negative emotions more strongly affect them.

5. CONCLUSIONS

The purpose of this study is to derive CTQs for service quality management strategy from the web-based VOC, transform into measurable CTQs by utilizing sentiment analysis, and finally conduct a case study to verify that it is possible sentiment values to be used as quantitative CTQs.

In order to derive CTQs for hotel service quality management and measure CTQ, the research was practiced with two-track flow. One is the process for deriving the CTQs, the other is for measuring them. In the process for deriving the CTQs, applied KJ Method was utilized. This method originates from KJ Method composed of four-step process; Step 1: Collecting data and note, Step 2: Sorting idea into groups, Step 3: Header card for discrete group, Step 4: Report and further analysis.

In this study, the KJ method was conducted in reverse order of steps. This is able to be called the applied KJ Method. Five topics("Operations, Sales, Amenities, Facilities and Experiences") as important elements in assessing hotel service quality became headers in the applied KJ Method and were pre-established. Key words were extracted using LDA, topic modeling technique, and sentences linked with the key words of five topics were also extracted in order to grasp the accurate intention of customers who posted opinions. The VOC was converted to CTQ after the process of interpretation.

Sentiment analysis was conducted on key words related to CTQ to convert CTQ into quantitative and measurable characteristics. The relation between the change of emotion on the extracted key words and the sales of hotel company was compared and analyzed. In the midst of the process, sentiment lexicon related to hotel service industry was constructed. Based on the lexicon, linear regression was conducted in order to identify how much sentiment associated with CTQ had an impact on "S" hotel's performance. As a result of the experiment, all of negative and positive emotion has an influence on the hotel performance and in particular in case that positive and negative emotions interact with, the negative emotion has a greater impact than positive on it through comparison between occupancy rate and sentiment change although positive opinions also have a strong influence on customers' choice of hotel.

It is certain that CTQ is one of the most crucial elements of quality management. However, it is still more lacking the study of CTQ in the field of service industry than manufacturing as ever. Existing CTQ extraction methods tended to be done through intuition by experiences. In fact, even though the method by which CTQ is measured quantitatively is ambiguous. In this study, however, LDA, topic modeling was used to extract CTQs through scientific and empirical experiments. The researcher can assure the process of deriving CTQs from raw data for the experiment by suggesting the methods, text mining techniques; LDA, and do scoring through sentiment analysis to grasp the influence on hotel business performance. So this study analyzed the impact of CTQ on business performance by measuring the sentiment of key words related to CTQ based on frequency. This could be an alternative to overcome the ambiguity of CTQ measurement. This study also has great significance in making CTQ for hotel business concrete, using a sentiment analysis as a

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

method to measure CTQ, and constructing sentiment lexicon for the hotel industry.

In spite of these managerial implications, there are limitations. It is necessary to study sentiment analysis on context, nuance and sentences in order to specify in detail and precisely measure the emotion and derive more meaningful experimental results by additionally applying various variables.

REFRENCES:

[1] O'Reilly Radar, *"Team, Big data now: Current perspectives from O'Reilly Radar"*, O'Reilly, 2011.

[2] McKinsey Global Institute, *"Big data: The Next Frontier for Innovation,Competition, and Productivity"*, McKinsey and Company, 2011.

[3] K. Min, and S. Kang, "Driving CTQ's for DFSS in the Service Area", *DAEHAN Association of Business Administration*, 2005, pp. 67-76.

[4] M. Ross and A. Radcliffe, "Making Six Sigma Applicable to Manufacturing, Service and Computing: The Southhampton Solent University Experiences", *Proceedings of the 12th International Conference on Manufacturing Research*, 2014, pp. 245-253.

[5] J. Kawakita, *"The Original KJ Method"*, Kawakita Research Institute, 1991.

[6] 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.

[7] D. Kim, S. J. Yu, "Hotel Review Mining for Targeting Stategy: Focussing on Chinese Free Independent Traveler", *Journal of Theoretical and Applied Information Technology*, VoI. 95, No.18, 2017, pp. 4436-4445.

[8] T. Kakkonen, N. Myller, and E. Sutinen, "Applying Latent Dirichlet Allocation to Automatic Essay Grading", *FinTAL 2006, LNAI 4139,* 2006, pp.110-120.

[9] www.hotelskorea.or.kr.

[10] B. Gruen, and K. Hornik, "Topic Models: An R Package for Fitting Topic Models", *Journal of Statistical Software*,Vol. 40, No. 13, 2011, pp.1-30.

[11] S. Kim, and W. Kim. "Study on the Selection Model CTQ data", *Journal of The Korea Society of Computer and Information,* Vol. 18, No. 4, 2013, pp 97-112.

[12] E. Diener, H. Smith, and F. Fujita, "The Personality Structure of Affect", *Journal of Personality and Social Psychology*, Vol. 69, No. 1, 1995, pp. 130-141.

[13] Jung, "The Influence of Negative Emotions on Customer Contribution to Organizational Innovation in an Online Brand Community", *Journal of Korean Society for Internet Information*, Vol. 14, No. 4, 2013, pp. 91-100.

[14] D. E. O'Leary, "Blog Mining-review and Extensions: From Each According to His Opinion", *Decision Support Systems,* Vol. 51, No. 4, 2011, pp. 821-830.

[15] H. Jung, and K. Nah, "A Study on the Meaning of Sensibility and Vocabulary System for Sensibility Evaluation", *Journal of the Ergonomics Society of Korea*, Vol. 26, No. 3, 2007, pp. 17-25.