Design of a Computer-based Approach for Analyzing Consumer Demands Method in Electronic Word of Mouth

Ching-Yuan Huang[a] Mao-Yuan Pai[a*] Shu-Yi Liaw[b] Yuh-Min Chen[a]
[a]Institute of Manufacturing Information and Systems National Cheng Kung University
[a]Department of Business Administration, National Pingtung University Science and Technology, Pingtung, ROC
* Corresponding author E-mail: maoyuanpai@gmail.com

ABSTRACT

Consumer opinions are one of the most valuable assets that enterprises have, and many consumers share their appraisals of products or services through electronic word-of-mouth (eWOM). Since these consumer appraisals usually reflect consumer needs, and thus their demands, collecting and analyzing eWOM data has become a key task for many businesses. In order to effectively collect and analyze eWOM data, this study proposes a computer-based approach for analyzing consumer demands. It is anticipated that companies will be able to improve their products and services through the application of this computer-based approach, thus enhancing their competitiveness.

Keywords: Electronic Word-of-Mouth (eWOM), Knowledge Management, Consumer Demands, Sentiment Analysis

I. INTRODUCTION

In this service economy era, enterprises' primary goals include rapid responses to consumer requirements, shortening the time needed to launch new products or services, and improving the quality of their products [1][2][3]. To meet these goals, more and more companies are putting efforts into discovering consumer demands with regard to their brands, products, and services. This information can then be used as a reference when in designing products or making innovations. Therefore, the efficient gathering and analysis of consumer demands can help companies to remain competitive [4][5].

With the development of new technologies, more and more consumers post their opinions about products or services on blogs, web forums, and bulletin board systems (BBS), as well as in e-mails. This phenomenon is called, "electronic word-of-mouth" (eWOM) or "online word-of-mouth" [6]. eWOM are consumer product or service appraisals, they also include important consumer demand. In order to understand consumer demands, enterprises are investing a substantial amount of money in the collection and analysis of eWOM content. However, the tremendous amount of such information that is available online often overwhelms these collection and analysis efforts, and results in eWOM information overload [7].

However, few studies have been conducted to investigate methods for the semantic analysis of eWOM. One of the advantages of analyzing eWOM content is that it directly reveals consumer opinions, which can enable enterprises to monitor market changes in real time. Nevertheless, the difficulty of analyzing eWOM lies in the complexity of the semantic content. Since consumers often use "textspeak" to post their opinions online, and because there are no standard writing formats and unified descriptive words for online reviews, the difficulty of analysis is increased. These problems also prevent computer programs from analyzing the text, and thus extracting the meanings of these terms is even more difficult.

Drawing upon the abovementioned difficulties of eWOM analysis, this study thus proposes a computer-based approach for the analysis of consumer demands. Specifically, this study has the following objectives: (i) constructing a consumer demands analysis model in eWOM; (ii) designing a consumer demands analysis procedure; and (iii) developing a system for analyzing consumer demands.

II. CONSUMER DEMANDS ANALYSIS MODEL USING EWOM.

A. *Traditional Approaches for Analyzing Consumer Demands*

In order to gain satisfaction and support from customers, businesses have long developed their products and services based on consumer opinions, with many enterprises and scholars using questionnaires to identify key consumer demands. The various methods that are used in the studies examined in this literature review to investigate consumer demands are shown in Figure 1:

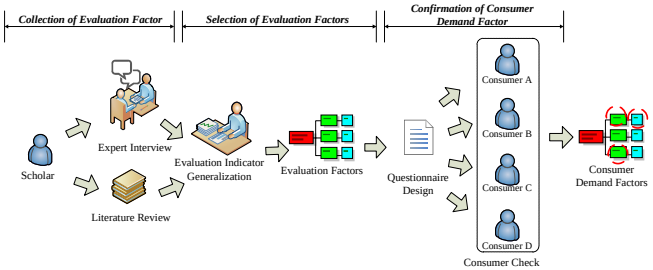

Figure 1. Traditional approaches for analyzing consumer demands

(1). Collection of Evaluation Factors*:* A researcher first collects the critical factors that influence purchasing decisions by employing two methods: First, experts are interviewed to identify evaluation factors. Second, a literature review is undertaken to identify evaluation factors.

Nonetheless, these two data collection methods have their limitations. Insufficient information from the

experts may jeopardize the reliability of the study, and an incomprehensive literature review can also hinder reliability.

(2). Selection of Evaluation Factors: After extensive data collection, the evaluation factors are organized and classified. This requires researchers to sort out the data, and repetitions of similar evaluation factors are eliminated. However, this process is time consuming and laborious.

(3). Confirmation of Consumer Demand Factors: The generated evaluation factors are subsequently used to design a questionnaire. By this process, the exact consumer demand factors can be identified.

The procedures outlines above require a considerable amount of work to filter and organize the collected data. However, through a computer based approach, it is possible to use automation to reduce the complexity of this process and avoid errors, and thus improve the accuracy of consumer demands analysis.

B. Analysis of the Relationship Between Appraisal Words and Consumer Demands

Generally speaking, the amount of appraisal towards a product or service tends to be steady, and a sudden shift in the amount of appraisal would indicate a change in consumer demand. For example, certain aspects of a product would attract consumers' attention, or another product's lack of features would not receive much interest, thus leads to a fall in the desire to discuss this product. Therefore, when there is a significant change in the amount of appraisals, it is likely that consumers' demands have altered. Company's marketing department and management could utilize this rate of appraisal words as a basis for determining consumer demands, and to find out the reasons behind such shifts in demands.

Frequencies have been used in many researches as an indicator for future trends. For instance, Google Trends uses the frequency of searches for keywords to determine trends in these related areas. Figure 2 shows a progression graph for the keyword "KFC" between year 2005 to 2010, and the letters A to M indicate different events during this period. As shown by the graph, the amount of searches for KFC changes with each event, and this is especially true for events J and K. It could be then inferred that these two events attracted the most attention from consumers online. It could also be said that when there is a change in the frequency of appearance for a keyword, certain events could have happened to cause an increase in searches for this keyword. In this research, we focus on understanding consumer demands using appraisal words, and the consumers' perceptions are shown by significant upswing or downswing of appraisal words. Therefore, the frequencies of the appearance of appraisal words could be used as indicators for identifying consumer demands.

To effectively understand the relationship between appraisal words and consumer demands, we use one case for the purpose of demonstration. For example, eWOM authors use appraisal words, such as "delicious," "smells good," and "not bad," to express their feelings about a product or service. However, eWOM authors' use of such words varies due to their different demographic backgrounds, such as environment, education, and age. Different appraisal words from a variety of authors can cause difficulties in computerized content analysis. To clarify this information, appraisal words have been organized into the following characteristics:

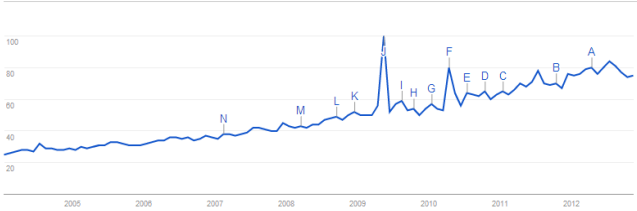

Figure 2. Keywords trends for "KFC"

(1). Synonyms: These are appraisal words that describe similar consumer demand factors. For example, for the consumer demand factor "flavor," "concentrated" is synonymous with "nice-smelling," while for the consumer demand factor "environment," "beautiful" is synonymous with "clean" (Figure 3).

Flavor	**Environment**
(**Synonymous**) -concentrate -nice-smelling -fragrant	(**Synonymous**) -beautiful -clean -comfortable

Figure 3. Examples of synonymous appraisal words

(2). Antonyms: Appraisal words that describe similar consumer demand factors but are opposite in meaning. For example, the appraisal word "nice-smelling" and "bad-smelling" express positive and negative flavor, respectively. Therefore, "nice-smelling" and "bad-smelling" are antonyms that describe the same consumer demand factor "flavor" (Figure 4).

Flavor (Positive Appraisal)		**Flavor (Negative Appraisal)**
-concentrate -nice-smelling -fragrant	←Antonym→	-dilute -bad-smelling -stink

Figure 4. Examples of antonym appraisal words

(3). Multi-Relational: Appraisal words that can be used to describe multiple consumer demand factors. For example, "not bad" and "very good" describe both the "flavor" and "environment" consumer demands factors.

C. Design of Consumer Demands Analysis Model in eWOM

Based on the analysis of the relationship between appraisal words and consumer demands, this study designed a consumer demands analysis model. It may help enterprises to

have a better understanding of consumer demands, as shown in Figure 5. The consumer demands analysis model mainly consists of eWOM collection, appraisal extraction, and appraisal word classification, as in explained in more detail below:

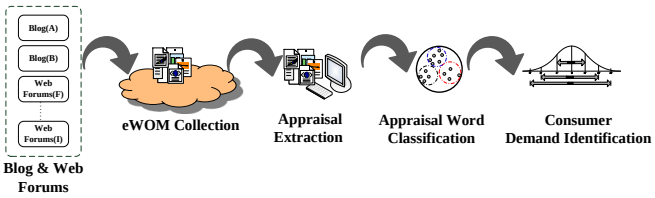

Figure 5. Consumer demands analysis model in eWOM

III. METHODS AND TECHNOLOGIES

A. *Procedure for Analyzing Consumer Demands in eWOM*

Based on the design of consumer demands analysis model in eWOM, this study develops procedures for consumer demands analysis. In this way, the enterprises may be able to better understand consumer demands, as displayed in Figure 6. The proposed consumer demands analysis in eWOM procedure consists mainly of eWOM collection, appraisal extraction, appraisal word classification and consumer demands identification, as described in the sections that follow.

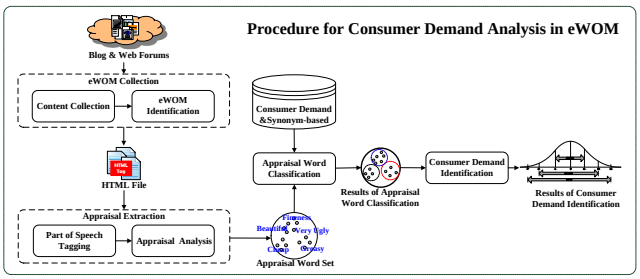

Figure 6. The Procedure for consumer demand analysis in eWOM

(1). eWOM Collection: eWOM content is collected from the Internet with content collection and identification technologies. The eWOM collection process includes content collection and eWOM identification, as described in detail as follow:
  (a). Content Collection: The names of products, services, and companies are the keywords that are searched for in the web content.
  (b). eWOM Identification: The collected web content is filtered for appraisal words.

(2). Appraisal Extraction: Semantic analysis and content extraction are conducted using the eWOM content to identify appraisal words related to the target products or services:
  (a). Parts of Speech Tagging: All of the parts of speech in the text are identified.
  (b). Appraisal Analysis: Sentiment analysis technologies monitor unknown appraisal words.

(3). Appraisal Word Classification: Appraisal words include consumer demands. Therefore, this study uses calculations of similarity to determine the relationship between appraisal words and consumer demands. Then, the classifications for appraisal word are used to calculate the frequencies of the appearance of appraisal words in consumer demand factors. The number of the appearance of consumer demand factor are saved month-by-month.

(4). Consumer Demand Identification: Analyzing the frequency for appraisal words in consumer demand factors, consumer demands identification is conducted to identify changes in consumer demands.

B. *eWOM collection*

In order to effectively collect eWOM content, a collection method was developed to extract specific content related to eWOM communication and to exclude non-related content (Figure 7).

The eWOM collection processes mainly includes “content collection” and “eWOM identification”, as described below:

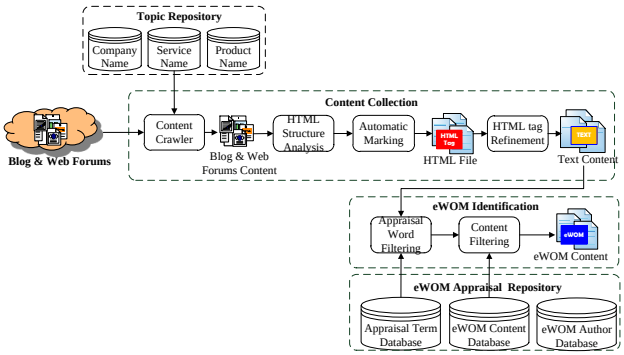

Figure 7. eWOM collection process

C. *Appraisal extraction*

Despite applying the above-mentioned filtering processes, the content still included some unidentified appraisal words. To determine the appraisal words that are commonly used by consumers, the following techniques were applied, as shown in Figure 8.

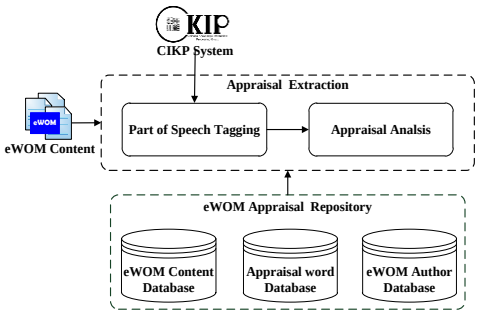

Figure 8. Appraisal extraction process

(1).Parts of Speech Tagging: The Chinese Knowledge Information Processing (CKIP) system, developed by Academia Sinica in Taiwan, is used to tag parts of speech.

(2). Appraisal Analysis: Consumer opinions have been frequently analyzed with sentiment analysis technologies [8]. In analyzing Chinese text, the Chinese sentiment analysis method proposed by Ku et al. [9] is used in this work for the analysis of unknown appraisal words, suing the following operational steps:

First and foremost, this study conducted a calculation of appraisal character weight, which indicates positive or negative values, as calculated with *equations* (1) and (2).

$$P_{ci} = \frac{fp_{ci}}{fp_{ci} + fn_{ci}} \tag{1}$$

$$N_{ci} = \frac{fn_{ci}}{fp_{ci} + fn_{ci}} \tag{2}$$

where ci is the i^{th} character in the appraisal character set, fp_{ci} is the number of times appraisal character ci appears in the positive appraisal word database; fn_{ci} is the number of times appraisal character ci appears in the negative appraisal word database; P_{ci} is the positive weight of the appraisal character ci, *and* N_{ci} is the negative weight of the appraisal character ci.

Equations (1) and (2) denote the number of times an appraisal character appears in the positive and negative appraisal word databases. *Equations* (3) and (4) regularize *equations* (1) and (2), respectively.

$$NP_{ci} = \frac{fp_{ci} / \sum_{j=1}^{n} fp_{cj}}{fp_{ci} / \sum_{j=1}^{n} fp_{cj} + fp_{ci} / \sum_{j=1}^{m} fn_{cj}} \tag{3}$$

$$NN_{ci} = \frac{fn_{ci} / \sum_{j=1}^{m} fn_{cj}}{fp_{ci} / \sum_{j=1}^{n} fp_{cj} + fp_{ci} / \sum_{j=1}^{m} fn_{cj}} \tag{4}$$

where fp_{ci} is the number of times appraisal character ci appears in the positive appraisal word database; fn_{ci} is the number of times appraisal character ci appears in the negative appraisal word database; n is the total number of appraisal words in the positive appraisal word database, and m is the total number of appraisal words in the negative appraisal word database; NP_{ci} is the positive weight of appraisal character ci in the positive appraisal word database, and NN_{ci} is the negative weight of appraisal character ci in the negative appraisal word database.

Based on the abovementioned results, we carried out a calculation of appraisal word weight. In *equation* (5), S_{cj} is the appraisal character weight of each appraisal character. NP_{ci} and NN_{ci} are the positive and negative weights of the appraisal character, respectively, and are used to assess the meanings of these characters. For example, if NP_{ci} is greater than NN_{ci}, the appraisal character is positive, and vice versa.

$$S_{ci} = NP_{ci} - NN_{ci} \tag{5}$$

$$S_w = \frac{1}{p} \times \sum_{j=1}^{p} S_{cj} \tag{6}$$

Equation (6) primarily evaluates whether the appraisal word is positive or negative. S_w is the weight of the appraisal word, and is either positive or negative, p is the number of appraisal characters in the appraisal word, and S_{cj} is the appraisal character weight of each appraisal character.

D. Appraisal word classification

Based on the analysis of the relationship between appraisal words and consumer demands carried out above, processes were designed to analyze the relationship between appraisal words and consumer demand factors by calculating the frequencies of appraisal words found in consumer demand occurrences. This study analyzed the words with similar meanings in both appraisal words and consumer demand factors, with the results used as the basis for the categorization of the former. In a review of the literature [10][11][12][13][14][15][16][17][18], this study found the following 21 demand factors: reliability, responsiveness, competence, access, courtesy, communication, credibility, security, understanding/ knowing, tangibles, food quality, food price, food innovation, food rarity, natural food, food flavor, food mouthfeel, serving size, convenience of location, restaurant environment and restaurant style. This study then constructed a database containing synonyms for all of these demand factors.

The Jaccard coefficient was applied to calculate the similarity between appraisal words and the synonym corpus for the consumer demand factors. Their similarity was analyzed, and the results used as the basis for categorizing the appraisal words, using *equation* (7).

$$sim(a_k, d_{ij}) = \left| \frac{a_k \cap d_{ij}}{a_k \cup d_{ij}} \right| \tag{7}$$

Based on *equation* (7), the results of appraisal word classification are listed in Table 4. a_k is the appraisal word set, $a_k \in \{a_1, a_2, a_3, \ldots, a_E\}$, E indicates the number of appraisal word sets (k=1~Z). As for "d_{ij}", i is the number of consumer demand factors, j is the ranking of the synonym in the corpus. This study uses appraisal word sets and the synonym corpus of consumer demand factors to construct a binary matrix, as indicated in *equation* (8).

$$sim(a_k, d_{ij}) = \begin{bmatrix} Max[sim(a_1, d_{1j})] & Max[sim(a_1, d_{2j})], & \cdots, & Max[sim(a_1, di_j)] \\ Max[sim(a_2, d_{1j})], & Max[sim(a_2, d_{2j})], & \cdots, & Max[sim(d_2, d_{ij})] \\ \cdots\cdots & \cdots\cdots & \ddots & \cdots\cdots \\ Max[sim(a_k, d_{1j})], & Max[sim(a_k, d_{2j})], & \cdots, & Max[sim(a_k, d_{ij})] \end{bmatrix} \tag{8}$$

E. Consumer Demands Identification

From the result of appraisal word classification, the frequency of appraisal word found in consumer demand factor per month could be calculated. To translate these numbers into

tangible enterprise knowledge, this research designs a consumer demands identification technique which analyzes and identifies changes in consumer demands.

As shown in Figure 9, this research assumes that consumer demand in each factor is normally distributed, and under the 95% confidence the frequencies of consumer demand occurrences would fall into the range of [μ+2σ, μ-2σ]. Therefore, when the frequency of consumer demand found in consumer demand occurrences is out of the 95% confidence interval range, it is defined as an obvious change in consumer demand factors.

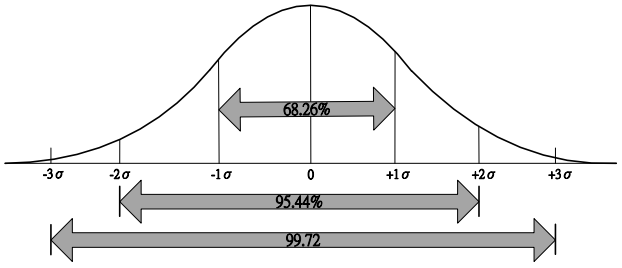

Figure 9. Estimation and Confidence Intervals

Equation (9) calculates the mean value of frequency of consumer demand occurrences per month, where t is the time set ($t \in \{t_j,\ t=1,2,3,\ldots,n-1\}$). *Equation* (10) calculates the standard deviation for frequency of consumer demand occurrences, where x_t stands for the frequency of consumer demand occurrences for each month.

$$\mu = \frac{1}{n-1}\sum_{t=1}^{n-1} X_t \tag{9}$$

$$\sigma = \sqrt{\frac{1}{n-2}\sum_{t=1}^{n-1}\left(X_t - \mu\right)^2} \tag{10}$$

IV. PROTOTYPE

The research set-up includes an Intel Core2 Duo E6420-2.13G PC, Microsoft Windows XP Professional, Internet Information Services (IIS), Microsoft SQL Server 2008, and Microsoft Visual Studio 2010.

These implementation displays are presented in Figures 10-14. Figure 10 shows the user interface screen for content collection. Figure 11 shows the semantic analysis and extraction of the content, while Figure 12 shows the classification of the appraisal words. Figure 13 shows the frequency results of consumer demand occurrence. Figure 14 presents the results of consumer demands identification.

V. CONCLUSIONS

A computer-based approach was developed in this study to analyze consumer demands, and this work also identified key consumer demands that can help enterprises develop their product lines and improve their services.

This work provides primary findings and contributions to the literature. First of all, user satisfaction survey has been conducted to reveal the effectiveness of this system, and the result proves that it is capable of analyzing consumer demands, as well as providing markets and managers valuable market information. Secondly, a new computer-based approach was proposed to improve traditional approaches for analyzing consumer demands.

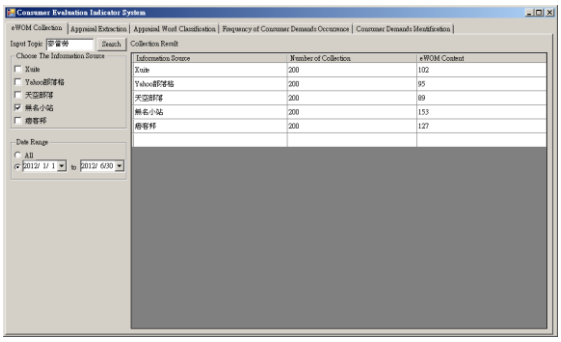

Figure 10. Collection of eWOM content

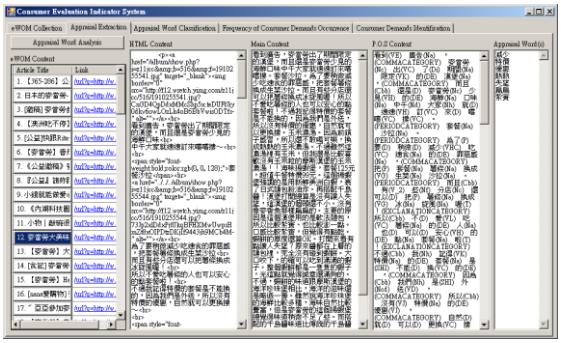

Figure 11. Semantic analysis and extraction of eWOM content

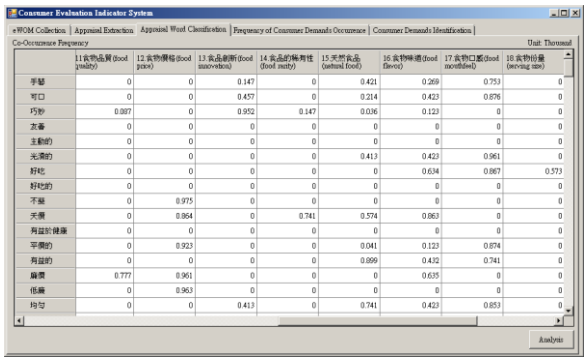

Figure 12. Classification of appraisal words

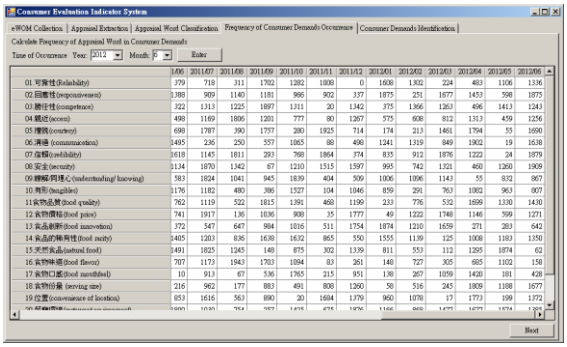

Figure 13. Frequency results of consumer demand occurrence

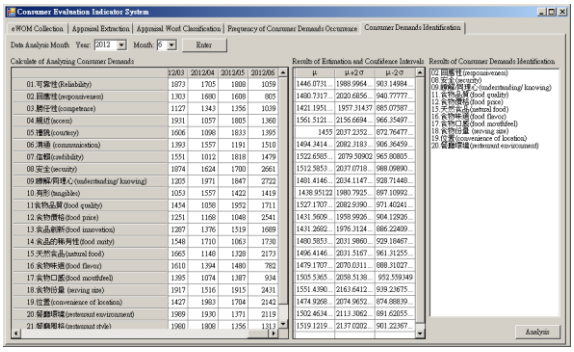

Figure 14. Results of consumer demands identification

REFERENCES

[1] C. W. D. Chen, and C. Y. J. Cheng, Understanding Consumer Intention in Online Shopping: A Respecification and Validation of the Delone and Mclean Model, *Behaviour & Information Technology*, vol.28, no.4, pp. 335-345, 2009.

[2] J. Y. Choi, J. Shin, and J. Lee, Strategic Demand Forecasts for the Tablet Pc Market Using the Bayesian Mixed Logit Model and Market Share Simulations, Online publication, *Behaviour & Information Technology*, pp.1-14, 2011.

[3] W. B. Walstad, and K. Rebeck, Assessing the Economic Knowledge and Economic Opinions of Adults, *The quarterly review of economics and finance*, vol.42, no.5, pp.921-935, 2002.

[4] L. T. Bei, and C. F. Shang, Building Marketing Strategies for State-Owned Enterprises against Private Ones Based on the Perspectives of Customer Satisfaction and Service Quality, *Journal of Retailing and Consumer Services*, vol.13, no.1, pp.1-13, 2006.

[5] Y. F. Kuo, C. M. Wu, and W. J. Deng, The Relationships among Service Quality, Perceived Value, Customer Satisfaction, and Post-Purchase Intention in Mobile Value-Added Services, *Computers in Human Behavior*, vol.25, no4, pp.887-96, 2009.

[6] T. Hennig-Thurau, and G. Walsh, Electronic Word-of-Mouth: Motives for and Consequences of Reading Customer Articulations on the Internet, *International Journal of Electronic Commerce*, vol.8, no.2, pp.51-74, 2003.

[7] D. H. Park, and J. Lee, Ewom Overload and Its Effect on Consumer Behavioral Intention Depending on Consumer Involvement, *Electronic Commerce Research and Applications*, vol.7, no.4, pp.386-398, 2009.

[8] T. K. Fan, and C. H. Chang, Sentiment-Oriented Contextual Advertising, *Knowledge and information systems*, vol.23, no.3, pp.321-344, 2010.

[9] L. W. Ku, T. H. Wu, L. Y. Lee, and H. H. Chen, Construction of an Evaluation Corpus for Opinion Extraction, *In: Proceedings of the Fifth NTCIR Wksp. on Evaluation of Information Access Technologies: Information Retrieval, Question Answering, and Cross-Lingual Information Access*. Tokyo, Japan, 2005.

[10] J. Currie, S. DellaVigna, E. Moretti, and V. Pathania, The Effect of Fast Food Restaurants on Obesity and Weight Gain, National Bureau of Economic Research, 2009.

[11] M Earle, Innovation in the Food Industry, *Trends in Food Science & Technology*, vol.8, no.4, pp.166-175, 1997.

[12] S. A. French, L. Harnack, and R. W. Jeffery, Fast Food Restaurant Use among Women in the Pound of Prevention Study: Dietary, Behavioral and Demographic Correlates, *International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity*, vol.24, no.10, pp.1353-1359, 2000.

[13] J. Kivela, R. Inbakaran, and J. Reece, Consumer Research in the Restaurant Environment, Part 1: A Conceptual Model of Dining Satisfaction and Return Patronage, *International Journal of Contemporary Hospitality Management*, vol.11, no.5, pp.205-222, 1999.

[14] M. Lee, and F. M. Ulgado, Consumer Evaluations of Fast-Food Services: A Cross-National Comparison, *Journal of Services Marketing*, vol.11, no.1, pp.39-52, 1997.

[15] Y. Liu, and S. C. S. Jang, Perceptions of Chinese Restaurants in the Us: What Affects Customer Satisfaction and Behavioral Intentions?, *International Journal of Hospitality Management*, vol.28, no.3, pp.338-348, 2009.

[16] 28-A. Parasuraman, V. A. Zeithaml, and L. L. Berry, A Conceptual Model of Service Quality and Its Implications for Future Research, *The Journal of Marketing*, vol.49, no.4, pp.41-50, 1985.

[17] V. A. Zeithaml, L. L. Berry, and A. Parasuraman, Communication and Control Processes in the Delivery of Service Quality, *The Journal of Marketing*, vol.52, no.2, pp.35-48, 1988.

[18] V. A. Ziethmal, and M. J. Bitner, 'Service Marketing: Integrating Customer Focus across the Firm', McGraw Hill, 2003.