Winter 12-6-2015

A Review System Based On Product Features In A Mobile Environment

Jaehyeon Ju

Dongyeon Kim

Kyuhong Park

Yongjin Park

Dong-Joo Lee

A REVIEW SYSTEM BASED ON PRODUCT FEATURES IN A MOBILE ENVIRONMENT

Jaehyeon Ju, KAIST, slamking@business.kaist.ac.kr
Dongyeon Kim, KAIST, South Korea, dykim88@business.kaist.ac.kr
Kyuhong Park, KAIST, South Korea, x64576@business.kaist.ac.kr
Yongjin Park, KAIST, South Korea, livediary@business.kaist.ac.kr
Jae-Hyeon Ahn, KAIST, South Korea, jahn@business.kaist.ac.kr
Dong-Joo Lee, Hansung University, South Korea, djlee@hansung.ac.kr

ABSTRACT

With the rapid growth of the mobile commerce, firms have been trying to get their online channels optimized for the mobile devices. However, many contents on online shopping sites are still focused on a desktop PC environment. Especially, consumer reviews are difficult to browse and grasp via a mobile device. Usually, it is not helpful to simply reduce the size of fonts or photos to fit to mobile devices without a fundamental transformation of the review presentation. In this study, we suggest a feature-based summarization process of consumer reviews in mobile environment. Further, we illustrate an implementation of the process by applying opinion mining techniques to product reviews crawled from a major shopping site in Korean. Finally, a plan for a controlled laboratory experiment is proposed to validate the effectiveness of the suggested review framework in this study.

Keywords: Review, Mobile, Usability, Opinion mining, Product features, Summarization, E-commerce.

INTRODUCTION

Mobile commerce continues its rapid growth with a great impact on the global e-commerce environment. According to Criteo report [4], the mobile channel accounts for 34% of e-commerce transactions globally in the fourth quarter of 2015. Mobile share of e-commerce is expected to reach 40% globally by the end of 2015. Especially, mobiles in South Korea and Japan hold over 50% of e-commerce transactions. The increase of e-commerce sites and payment methods that are optimized for mobile devices may be key drivers in the two countries. About 60% of Japanese e-commerce sites support a mobile version [1]. More and more sites become optimized for a smartphone or a tablet globally. However, several e-commerce features such as consumer reviews are still focused on a desktop PC and e-commerce sites are required to change them. They are poorly suited for a smartphone, making contents difficult to navigate. It is unhelpful to simply reduce the size of fonts to fit to small screen size without a fundamental transformation.

Time spent on mobiles exceeds desktop PCs [16]. Mobile users spend 2.8 hours, 51% of total 5.6 hours per day with digital media. Mobile computing is now a part of most people's daily lives and is complementary to a desktop PC. The numbers of online shoppers using smartphones and desktop PCs are 291.1 million and 333.1 million, respectively [22]. In addition, the number of products which consumers browse via smartphones when going shopping is the same as that via desktop PCs [4]. Consumers use their mobile devices to explore and purchase products even when a desktop PC is nearby [22]. Smartphones are used constantly regardless of where they are. Over 40% of consumers think a smartphone is an important resource for a purchase and about 60% of consumers have used a mobile exclusively when deciding to purchase products [22]. Browsing products from a smartphone has become increasingly common. However, there are disadvantages in the use of a smartphone on e-commerce. Because of physical constraints of a mobile such as small size, navigating via a smartphone can be a huge pain. Although average page views for a mobile and a desktop PC are 8.2 and 9.10, respectively [23], there can be a difference between the amounts of gathered information via a mobile and a desktop PC. For example, suppose that consumers read reviews of a product before making a purchase decision on Amazon.com (Figure 1). Consumers navigating via a mobile spend more time to get the same amount of information compared to consumers using a desktop PC. There are 15 lines vertically in left side of Figure 1 whereas review in right side has just 8 lines and fewer than 8 characters per line. Consumers can see a full review by touching "Read full review" and have to do up and down scrolling frequently. It would be a bothersome task given a great number of reviews. According to Monetate report [23], add-to-cart rate and sales conversion rate for mobile (5.41%, 0.96%) are much lower than for desktop (7.75%, 2.71%). The content form which is not suited for a mobile can be one of the reasons. Thus, the design should be visually simpler and very easy to navigate, effectively providing the information that consumers want to see.

Popular shopping sites such as Amazon.com support mobile version but these sites are focused on fitting to screen size of device. Original photos of desktop version are transformed into smaller photos or the latter part of original product descriptions is omitted according to screen size. This transformation can be applied to mobile version because the amounts of these contents are limited. However, product reviews are different. The length of review is long and the number of reviews is already uncountable on many sites. It is difficult to grasp the whole content via mobile browsing. Mobile consumers can face information overload [27] and browsing loss [8] because they are not able to process the information adequately or the information is not well-organized to understand [12]. Consumers sometimes suffer a trouble in product evaluation even when using a desktop PC [3]. Therefore, current review presentation form on a mobile has to be changed. In mobile environment, it is unsuitable to display detailed review as a desktop version. The appropriate form presenting product reviews is important because product reviews have a great impact on consumers' purchase decision [6][9][17].

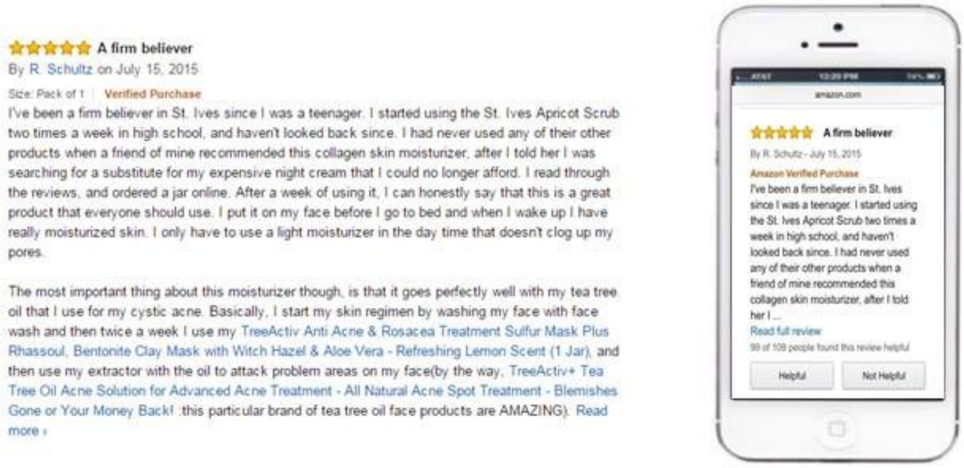


Figure 1. An illustration of the product review on Amazon.com: desktop (left) vs. mobile (right)

Website adaptation to alleviate the problems of frequent scrolling [27] and restricted navigation [24] has been focused on the desktop PC environment. However, content transformation or presentation for a mobile has not been extensively studied. Because of distinct characteristics of a mobile, research in desktop PC environment may no longer be applicable. This paper presents the process of opinion summarization of consumer reviews and explores how to adapt the reviews for mobiles devices effectively. We use a summarization technique to make the optimized form for a mobile. Previous studies report that text summarization in desktop PC environment enables users to find information faster and improve their satisfaction [11][19][20]. We develop a review summarization system which classifies reviews into two types, positive and negative reviews, for each product feature. It consists of feature selection and classifying sentence that contains features. This system provides an overview about product features and a representative sample of each feature as an ouput. Utilizing the output, we will implement two different review systems, an original review system and a feature-based review system, for mobile users. Finally, we propose an alternative to minimize usability issue for a mobile

PROPOSED FRAMEWORK

Opinion mining is the task of finding out user's opinion, attitude, and emotion towards particular topic [28]. It is performed in sentence [13] or document level [29] and determines whether user's opinion is positive, negative or neutral about specific subject. A review consists of a subjective sentence or an objective one. Consumers want to know the reasons why the existing consumers like/dislike the product, and sellers are basically interested in consumers' overall opinions about their products. Thus, most of the existing research on opinion mining are focused on product feature and adjective identification [5][13][26]. However, there are limitations of these existing systems. The accuracy of feature extraction and identification of opinion words is not satisfactory. Thus, the results of classification are not reliable. Identification of product features and opinion words are important tasks in opinion mining. By adding domain knowledge to the existing works, we classify product reviews in sentence level.

Through the results of classification, we aim to investigate the effect of content transformation in mobile environment. Mockup webpages for a mobile can be utilized to explore the impact of content transformation instead of using the actual result of complicated opinion mining. However, mockup webpages are limited to implement actual mobile environment. Limited reviews which are classified manually may not authentically observe user's actual perception or performance on review summarization for a mobile. Also, most of the limited empirical studies on content transformation on mobiles used emulators of mobiles on desktop PC. Some technical challenges that mobile users suffer can be overlooked when performing tasks. It would be easier to browse using a mouse on an emulator than touching on a mobile. The results of the studies may not truly reflect actual perception and performance on mobile [30]. Therefore, we will use actual review data which are classified by review summarization system and implement a new review system for mobile.

The architectural overview of our review summarization system includes five steps, pre-processing, feature extraction, opinion extraction, opinion classification, and visualization (Figure 2). When users write reviews, they may make errors inadvertently. These errors generally include grammatical, spelling, and punctuation errors. The errors have a negative influence on the performance of the opinion mining because the technique is based on word extraction and linguistic analysis. Opinion mining tasks such as splitting sentence or part-of-speech (POS) tagging [18] are influenced by these errors. Statistical analyses relying on term frequencies are also likely to provide wrong conclusions. Thus, pre-processing crawled data is required prior to beginning review summarization.

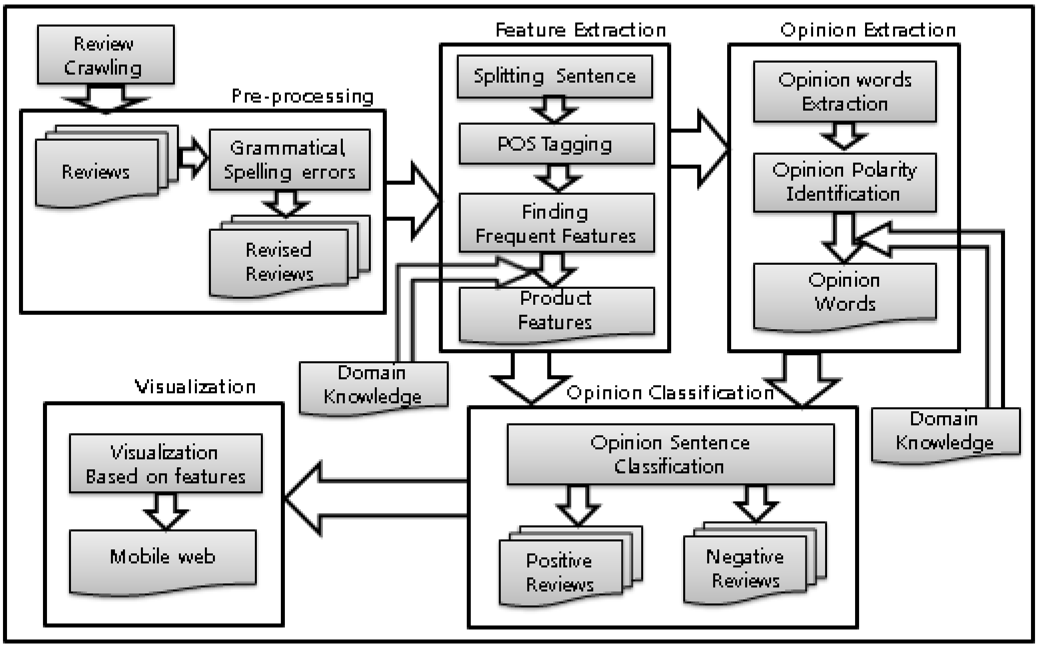


Figure 2. The process of feature-based review summarization

DATA COLLECTION & PRE-PROCESSING

Data were acquired in the NAVER shopping site using crawlers. NAVER shopping is the biggest product review website in Korea, collecting all consumer reviews from the representative online shopping sites in Korea such as Auction, G-market and 11st. The acquired data contain product category, product name, price, review, etc. The number of collected reviews is 16,083,512 and we choose a specific category, cream in cosmetics, which has many reviews (394,852). Because of the negative effects of errors mentioned above, it is necessary to eliminate noise of review data as much as possible. Pre-processing comprises sub tasks like spelling check, grammar correction, etc. The goal of pre-processing is not to correct all the errors but to minimize the number of errors. We used spell-check program that is provided by NAVER. About 2.2% of 394,852 reviews is removed because these reviews have encoding errors (e.g., “%amp;”) or meaningless words (e.g., “ekvnwdkenvirsmut”).

DATA PROCESSING

Feature Extraction

Splitting sentence

A product review consists of several sentences. It contains evaluation of a shopping site, a product, a seller, or delivery. Other factors except product evaluations are external factors because these factors depend on sites selling a product. We focus on the consumer’s product evaluation based on product features. To sort out unnecessary sentences, we need to split user’s review into sentences. After splitting each review, we got 1,039,766 sentences. We parsed each review utilizing linguistic parser which is used in natural language processing (NLP).

Part-of-Speech (POS) tagging

The most fundamental part of the linguistic analysis is part-of-speech tagging. POS aims to label each word with a distinct tag that represents its syntactic position such as noun and verb. Product features exist as a noun or noun phrases in a sentence. POS tagger is important to extract noun or noun phrases. We use Kokoma Korean morpheme analyzer (http://kkma.snu.ac.kr) on each sentence. A revised data file contains a sentence and POS tag information of each word in each sentence. An example is as follows (Table 1).

Table 1. The example of data file

Sentence	POS tagging
This cream product is good for dry skin	This[DT] cream[NN] product[NN] is[VBZ] good[JJ] for[IN] dry[JJ] skin[NN]

Finding frequent features

This step identifies product features on sentence that consumers wrote. In our framework, the process of feature extraction is designed to operate in semi-automation. Some studies try to use a statistical approach to find frequent features [13][31]. The drawback of this approach is that it may extract words which are not related to product attributes. We have proposed a framework which utilizes domain knowledge to frequent features list. There are a lot of nouns or noun phrases in whole sentences. A small part of nouns or noun phrases appears commonly in many sentences and the rest of them are only in one sentence or two sentences.

The Pareto principle (known as the 80-20 rule) may be applied. The number of initial nouns in a cream category was 31,445. We extracted nouns or noun phrases which are high on the list according to Pareto principle. The number of extracted nouns is 204. Especially in Korean, demonstrative, interrogative, personal or reflexive pronouns are classified as nouns. Nouns such as "one", "two" or "three" are removed. Because consumers mention the name of product category and brand many times, these are also removed. There are still many nouns unrelated with product features in the list. After elimination, the list of nouns left has 75, which is small enough compared to the initial number of nouns and noun phrases. Four persons who have domain knowledge chose nouns or noun phrases which are really related to product attributes out of the 75 nouns or noun phrases. The final list contains 21 nouns or noun phrases.

Opinion Extraction

Opinion words extraction

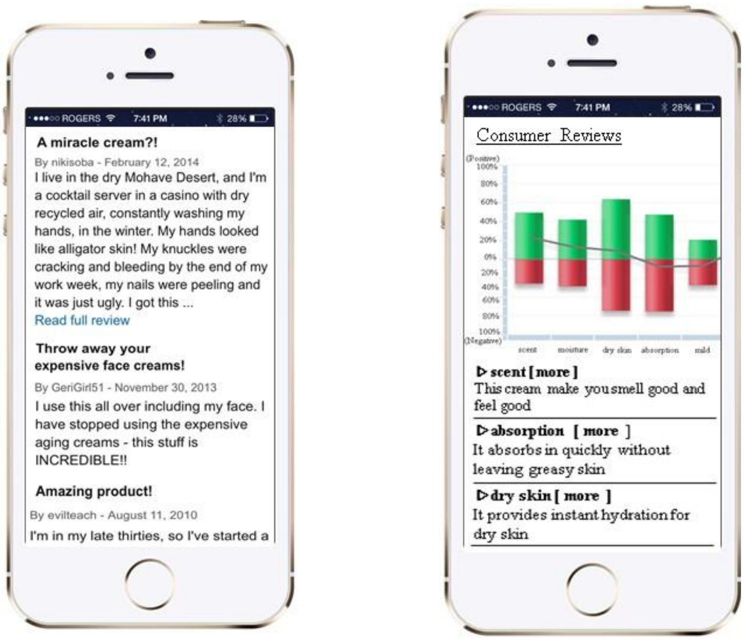
We limit to sentences containing product attributes in the final list above because we make a summary of consumer's opinion on product attributes. The number of sentences that include product attributes is 58,654 among 1,039,766. This step identifies opinion words that consumer expresses a positive, negative, or neutral opinion. Previous research reports that an opinion word is close to a product feature [13]. Also, the presence of adjectives is significantly correlated with subjectivity of a sentence [2]. Thus, we use adjectives and verbs which are near product attributes as opinion words. In total, 61,339 opinion words are selected.

Opinion polarity identification

In this step, we identify opinion word's polarity based on SO-PMI algorithm [29]. This algorithm estimates the semantic orientation by measuring the similarity of pairs of words. When the reference words exists "good" and "bad", the assumption is that a phrase has a positive semantic orientation when a phrase is more strongly associated with "good". Choosing the reference words is important when using SO-PMI algorithm. The Dictionaries like WordNet [21] do not include polarity information for Korean words. Thus, we used Korean sentiment dictionary, OpenHangul (www.openhangul.com) based on collective intelligence. According to the frequency of opinion words, we limit to opinion words that appear more than 100 to choose candidates of reference words. In total, 410 opinion words are selected. We determine the polarity of candidates using OpenHangul. There is limited in determining accurate polarity because OpenHangul does not decide polarity with domain knowledge. The four persons who have domain knowledge about e-commerce extracted the most related words because 410 words are small enough to examine manually. After extracting, the positive group has 68 words and the negative group has 38 words. These words are used as a seed list. The goal of this research is not to automate entire processes but to make review pages for mobile in semi-automation. Using SO-PMI algorithm and the seed list, we determined the polarity of other adjectives and verbs.

Opinion Classification

Opinion sentence classification



After constructing feature and opinion dictionary, we classify each sentence into a positive sentence or a negative sentence. If positive (negative) opinion word appears, the sentence is regarded as a positive (negative) one. However, there may be more than one feature words and opinion words in a sentence. When we make a feature-opinion pair, it is important to know whether this pair is valid or not. We use a naïve method that finds the closest opinion word for a product feature. In this research, we do not consider a "but" clause and a negation word.

Figure 3. Traditional review (left) and feature-based review page (right) for a mobile

Visualization

Our research aims to investigate the effect of mobile webpage that applies feature-based review classification. Figure 3 gives an example of traditional review page and feature-based review page for mobile. Traditional review page shows each review that existing consumers wrote without any transformation. Feature-based review page provides positive/negative review ratio by product features and shows a representative review about each product feature (Figure 4). If consumers want to navigate more reviews about the product feature, they touch "[more]" next to each product feature.

Figure 4. A new system: summary by features (left) and list by features (right)

EXPERIMENT

This study aims to investigate the effectiveness of the suggested review framework through a controlled laboratory experiment with two different mobile webpages. We implement a new mobile website for consumer review system (e.g., Figure 3). The laboratory test minimizes distraction from other factors such as user mobility. Participants are asked to do product purchasing tasks on two different systems via mobile. One of two systems is randomly selected and each participant is randomly assigned to one system.

To evaluate two different systems, we employ three constructs: user perception, information quality and memorability. (1) Perceived ease of use and perceived usefulness are used as measurement of user perception. These are useful to evaluate information systems and user's intention to adopt them [7][15]. The instruments for perceived ease of use and usefulness are proposed in the technology acceptance model and we adapt from those. Perceived ease of use refers to degree to which consumers think that given system can be used effortlessly and perceived usefulness refers to the degree to which consumers think that given system can improve their task performance. (2) Information quality consists of conforming to specifications and meeting consumer expectations and it is measured by completeness and appropriate amount of information [14]. Completeness refers to the degree to which information is not missing and is of sufficient breadth and depth for the task. Appropriate amount of information refers to the degree to which the volume of information is appropriate for the task. (3) Memorability is a typical dimension of usability [10]. It is used in examining the terms used, labels of displays, etc. When participants complete tasks, they perform a recall memory test. Later, we will find and add other constructs or instruments to evaluate our system more appropriately.

We are currently doing data processing making a new review system for a mobile. After completing data processing, we will implement the mobile web for consumer reviews which works actually. Thus, the experiment will be conducted later.

DISCUSSION

The number of product reviews grows rapidly and there are hundreds of reviews in popular products. A potential consumer may suffer difficulty to decide whether a product is worth purchasing. If consumers purchase a product after reading a few reviews, they may purchase with a biased view. It is also difficult for sellers to keep tracking consumers' opinion. As the number and the length of review change, a new review system is required. Because a mobile has been widely popularized, accessing the review from smartphones has become increasingly common. Thus, a new mobile version of a review system is necessary. In that sense, our research is timely and pertinent. We propose feature-based review summarization as an alternative to the traditional review system for mobile. We expect to help a potential consumer to make a purchase decision with an unbiased view. Effective presentation of information is crucial for improving consumer experience in mobile environment [25]. Moreover, as a consumer becomes more comfortable with mobile, his/her satisfaction increases [22]. In the perspective of purchase immediacy, about 49% of smartphone consumers want to make a retail purchase within an hour and 67% want to purchase within a day [22]. A potential consumer using a mobile do not want spend much time in shopping. Administrators who are operating online shopping sites should consider how to provide consumers enough information to purchase within a short time. The lack of average page view for mobile shopping has been partly attributed to its usability. Offering a helpful review system to consumers can raise the sales conversion rate in mobile environment. Our research can be extended to easily other contexts in mobile environment.

ACKNOWLEDGEMENT

This research was financially supported by Hansung University.

REFERENCES

[1] Atlas21 (2015) 'Domestic EC site survey about smartphone/tablet', *Atlas21(Web solution company)*, available at http://at21.jp/web/topic/topic03.html (accessed 13 October 2015).

[2] Bruce, R. F., & Wiebe, J. M. (1999) 'Recognizing subjectivity: A case study in manual tagging', *Natural Language Engineering*, Vol. 5, No. 2, pp. 187-205.

[3] Cao, Q., Duan, W., & Gan, Q. (2011) 'Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach', *Decision Support Systems*, Vol. 50, No. 2, pp. 511-521.

[4] Criteo (2015) 'State of mobile commerce Q1 2015', *Criteo(digital performance advertising firm)*, available at http://www.Criteo.com/resources/mobile-commerce-q1-2015/ (accessed 13 October 2015).

[5] Cui, H., Mittal, V., & Datar, M. (2006). 'Comparative experiments on sentiment classification for online product reviews',*In AAAI*, July, Vol. 6, pp. 1265-1270.

[6] Dellarocas, C., Gao, G., & Narayan, R. (2010) 'Are consumers more likely to contribute online reviews for hit or niche products?', *Journal of Management Information Systems*, Vol.27, No. 2, pp. 127-158.

[7] Davis, F. D. (1989) 'Perceived usefulness, perceived ease of use, and user acceptance of information technology', *MIS quarterly*, Vol. 13, No. 3, pp. 319-340.

[8] Furnas, G. W. (1997) 'Effective view navigation', In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, March, pp. 367-374.

[9] Godes, D., & Mayzlin, D. (2004) 'Using online conversations to study word-of-mouth communication', *Marketing science*, Vol. 23, No. 4, pp. 545-560.

[10] Han, S. H., Yun, M. H., Kwahk, J., & Hong, S. W. (2001) 'Usability of consumer electronic products', *International Journal of Industrial Ergonomics*, Vol. 28, No. 3, pp. 143-151.

[11] Harper, S., & Patel, N. (2005) 'Gist summaries for visually impaired surfers', In *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility*, October, pp. 90-97.

[12] Hiltz, S. R., & Turoff, M. (1985) 'Structuring computer-mediated communication systems to avoid information overload',*Communications of the ACM*, Vol. 28, No. 7, pp. 680-689.

[13] Hu, M., & Liu, B. (2004) 'Mining and summarizing customer reviews', In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, August, pp. 168-177.

[14] Kahn, B. K., Strong, D. M., & Wang, R. Y. (2002) 'Information quality benchmarks: product and service performance',*Communications of the ACM*, Vol. 45, No. 4, pp. 184-192.

[15] Karahanna, E., & Straub, D. W. (1999) 'The psychological origins of perceived usefulness and ease-of-use', *Information & Management*, Vol. 35, No. 4, pp. 237-250.

[16] KPCB (2015) 'Internet Trends 2015', *KPCB (Kleiner Perkins Caufield & Byers)*, available at http://www.kpcb.com/ blog/2015-internet-trends (accessed 13 October 2015).

[17] Liu, Y. (2006) 'Word of mouth for movies: Its dynamics and impact on box office revenue', *Journal of marketing*, Vol. 70, No. 3, pp. 74-89.

[18] Manning, C. D., & Schütze, H. (1999) *Foundations of statistical natural language processing*, MIT press.

[19] McDonald, D. M., & Chen, H. (2006) 'Summary in context: Searching versus browsing', *ACM Transactions on Information Systems (TOIS)*, Vol. 24, No.1, pp. 111-141.

[20] McKeown, K., Passonneau, R. J., Elson, D. K., Nenkova, A., & Hirschberg, J. (2005) 'Do Summaries Help? A task-based evaluation of multi-document summarization', *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Salvador, Brazil, August 15-19, pp. 210-217.

[21] Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D., & Miller, K. J. (1990) 'Introduction to wordnet: An on-line lexical database*', *International journal of lexicography*, Vol. 3, No. 4, pp. 235-244.

[22] MMAglobal (2014) 'Mobile path to purchase', *MMA(Mobile Marketing Association)*, available at http://www. mmaglobal.com/files/documents/mobile_path_to_purchase_2014_overview_us_whitepaper.pdf (accessed 13 October 2015).

[23] Monetate (2015) 'Monetate Q1 2015 ecommerce quarterly', *Monetate (A provider of personalization for web)*, available at http://www.monetate.com/resources/research/ (accessed 13 October 2015).

[24] Pang, B., Lee, L., & Vaithyanathan, S. (2002) 'Thumbs up?: Sentiment classification using machine learning techniques', In*Proceedings of the ACL-02 conference on Empirical methods in natural language processing*, July, pp. 79-86.

[25] Qiu, M., Zhang, K., & Huang, M. (2006) 'Usability in mobile interface browsing', *Web Intelligence and Agent Systems*, Vol.4, No. 1, pp. 43-59.

[26] Somprasertsri, G., & Lalitrojwong, P. (2008) 'A maximum entropy model for product feature extraction in online customer reviews', *In Cybernetics and Intelligent Systems of 2008 IEEE Conference*, September, pp. 575-580.

[27] Tsandilas, T., & Schraefel, M. C. (2004) 'Usable adaptive hypermedia systems', *New Review of Hypermedia and Multimedia*, Vol. 10, No. 1, pp. 5-29.

[28] Tsytsarau, M., & Palpanas, T. (2012) 'Survey on mining subjective data on the web', *Data Mining and Knowledge Discovery*, Vol. 24, No. 3, pp. 478-514.

[29] Turney, P. D. (2002) 'Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews', In *Proceedings of the 40th annual meeting on association for computational linguistics*, July, pp. 417-424.

[30] Zhang, D., & Adipat, B. (2005) 'Challenges, methodologies, and issues in the usability testing of mobile applications', *International Journal of Human-Computer Interaction*, Vol. 18, No. 3, pp. 293-308.
[31] Zhuang, L., Jing, F., & Zhu, X. Y. (2006) 'Movie review mining and summarization', In *Proceedings of the 15th ACM international conference on Information and knowledge management*, November, pp. 43-50.