Information Processing and Management 53 (2017) 436–449

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/infoproman

Opinion mining from online hotel reviews – A text summarization approach

Ya-Han Hu[a,*], Yen-Liang Chen[b], Hui-Ling Chou[b]

[a] *Department of Information Management, National Chung Cheng University, No. 168, University Rd., Chiayi 62102, Taiwan, ROC*
[b] *Department of Information Management, National Central University, No. 300, Jhongda Rd., Chung-Li 32001, Taiwan, ROC*

ARTICLE INFO

Article history:
Received 12 January 2016
Revised 22 November 2016
Accepted 2 December 2016
Available online 10 December 2016

Keywords:
Online hotel review
Multi-text summarization
Sentiment analysis
Sentence clustering
k-medoids

ABSTRACT

Online travel forums and social networks have become the most popular platform for sharing travel information, with enormous numbers of reviews posted daily. Automatically generated hotel summaries could aid travelers in selecting hotels. This study proposes a novel multi-text summarization technique for identifying the top-k most informative sentences of hotel reviews. Previous studies on review summarization have primarily examined content analysis, which disregards critical factors like author credibility and conflicting opinions. We considered such factors and developed a new sentence importance metric. Both the content and sentiment similarities were used to determine the similarity of two sentences. To identify the top-k sentences, the k-medoids clustering algorithm was used to partition sentences into k groups. The medoids from these groups were then selected as the final summarization results. To evaluate the performance of the proposed method, we collected two sets of reviews for the two hotels posted on TripAdvisor.com. A total of 20 subjects were invited to review the text summarization results from the proposed approach and two conventional approaches for the two hotels. The results indicate that the proposed approach outperforms the other two, and most of the subjects believed that the proposed approach can provide more comprehensive hotel information.

1. Introduction

Advances in Internet technology and the vigorous development of Web 2.0 applications have caused substantial change in the tourism industry (Litvin, Goldsmith, & Pan, 2008). The rise of online tourism forums has rendered the Internet the primary means of seeking travel information (Chung & Koo, 2015; Jannach, Zanker, & Fuchs, 2014; Li, Law, Vu, Rong, & Zhao, 2015; Liu & Park, 2015). Travelers communicate with each other and share their perspectives and experiences on social networking sites, generating numerous reviews daily (Cantallops & Salvi, 2014; Chung, Han, & Koo, 2015; Ye, Law, Gu, & Chen, 2011). For example, TripAdvisor.com, one of the most widely used travel sites, provides a platform for sharing reviews and opinions on various hotels and restaurants. In addition, members can rate reviews according to their perceived usefulness. According to a survey conducted by Ady and Quadri-Felitti (2015), nearly 95% of travelers read online hotel reviews before making their booking decision, and more than one-third of travelers believe that the opinion expressed in reviews is the most critical factor in selecting a hotel online.

* Corresponding author.
E-mail address: yahan.hu@mis.ccu.edu.tw (Y.-H. Hu).

http://dx.doi.org/10.1016/j.ipm.2016.12.002
0306-4573/

Up to thousands of reviews for a single hotel can easily be accumulated from social media, but this causes the problem of information overload (Harer & Kadam, 2014; Liu, Hsaio, Lee, Lu, & Jou, 2012; Peetz, de Rijke, & Kaptein, 2016). Summarized review content can give readers the most essential information about a hotel and also save time during their purchasing decision process (Ady & Quadri-Felitti, 2015). Papathanassis and Knolle (2011) indicated that in the era of web 2.0, a novel content management process is required to control and maintain the user-generated content (UGC) in tourism websites. Therefore, how to automatically summarize online hotel reviews is an appealing research topic.

The *text summarization* technique is used to extract the most essential information from an original text and generate a simplified version of the text for users (Gambhir & Gupta, 2016; Gupta & Lehal, 2010; Kar, Nunes, & Ribeiro, 2015; Mani & Maybury, 1999). In the past decade, text summarization has been applied in various domains (Abdi, Idris, Alguliev, & Aliguliyev, 2015; Liu et al., 2012; Liu, Chen, & Tseng, 2015; Ly, Sugiyama, Lin, & Kan, 2011; Mehta, 2013; Sankarasubramaniam, Ramanathan, & Ghosh, 2014). For example, given a user query, Google can offer dozens of website links as well as a short paragraph of summarized text regarding the content of each website, which helps users judge the interestingness or usefulness of websites. For another example, a well-known application software, Summly, can automatically retrieve relevant news articles and then show the digests of each news article according to the news categories selected by users.

Generally, text summarization can be divided into single-text and multi-text summarizations (Gupta & Lehal, 2010). Single-text summarization addresses the problem of text summarization for a single document only. Because the author (or group of authors) of a document completed it according to a common consensus, the content does not exhibit inconsistency problems. In other words, the opinions in a single text do not conflict. Furthermore, because a single document is issued at a specified time point, single-text summarization does not need to consider the effect of content novelty. Comparably, multi-text summarization simultaneously processes multiple documents on the same subject (Heu, Qasim, & Lee, 2015). When handling multi-text summarizations, the problem of conflicting opinions raised by different authors should be resolved and the consistency of semantic expression in the summarization results must be ensured.

This paper proposes a novel multi-text summarization technique specifically designed for hotel review summarizations. Previous studies on review summarizations have primarily focused on using text processing techniques, such as bag-of-word and semantic approaches (Atkinson & Munoz, 2013; Jeong, Ko, & Seo, 2016; Meng & Wang, 2009; Wang, Zhu, & Li, 2013), disregarding other useful information that could be extracted from online social media, such as (a) author reliability, which implies that the reviews written by a famous author or blogger have high credibility; (b) review time, which implies that more recent reviews generally provide users with more up-to-date information; (c) review usefulness, which implies that reviews that receive higher ratings from other users typically provide more information; and (d) conflicting opinions, which implies that different reviewers have their own preferences and might not agree with one another. Thus, reviews expressing different sentiment polarities (i.e., either positive or negative comments) might contain information raised by different reviewers. To the best of our knowledge, no studies have considered these four factors collectively for the hotel review summarizations. In this manner, this study addressed the following research questions:

- How can the proposed multi-text summarization technique accurately generate a useful summarized review from a set of online hotel reviews?
- With the consideration of author reliability, review time, review usefulness, and conflicting opinions, can the proposed approach offer better hotel review summarization results than the conventional text summarization approaches (i.e., consider review text only)?

In experimental evaluation, the hotel reviews were collected from TripAdvisor.com and the abovementioned four factors were jointly considered. In particular, the first three factors were used to calculate the importance score of each sentence. To resolve conflict opinions, both content and sentiment analyses were performed to determine the similarity between two sentences. Based on the new similarity measure, the sentences can be clustered into k groups and the most representative sentence in each group can be utilized to form the final text summarization result (i.e., top-k informative sentences).

The remainder of this paper is organized as follows: Section 2 reviews both online hotel reviews and previous research on text summarization; Section 3 formally discusses the research method of this study; Section 4 presents the preparation of data, experimental setup, and the evaluation results; and finally, Section 5 concludes the paper.

2. Literature review

2.1. Online hotel reviews

UGC has become one of the main information sources on the Internet. Different types of electronic word-of-mouth (eWOM) such as online reviews and opinions have been recognized as the most influential communication channel among service providers and consumers as well as among consumers themselves (Casaló, Flavián, & Guinalíu, 2010).

As the arising of online tourism services, the hotel industry is strongly affected by eWOM. Cantallops and Salvi (2014) reviewed the published articles regarding the influence of eWOM in the hotel industry. Two major lines of research have been identified: review-generating factor and the impacts of eWOM. The first line of research aim at discovering critical factors that motivate a user to write online hotel reviews (Casaló et al., 2010; Stringam & Gerdes Jr, 2010). A number of aspects have been investigated, including service quality and satisfaction, failure and recovery, and sense of community belonging

(Kim, Mattila, & Baloglu, 2011; Swanson & Hsu, 2009). The second line, the mainstream research nowadays, focuses on understanding the impact of online hotel reviews (Litvin et al., 2008; Vermeulen & Seegers, 2009). From customer's perspective, factors influencing online purchase and repurchase of hotel accommodation and the decision of selecting hotel online have been studied (Verna, 2010; Wen, 2009). Relatively few studies have analyzed the impact of eWOM from corporate perspective. Most of these studies addressed issues regarding the pricing of hotels, company's online reputations, interactions with online users, and the generation of customer loyalty (Loureiro & Kastenholz, 2011; Yacouel & Fleischer, 2012).

On investigating the impact of eWOM from either consumer or company perspectives, variables related to review helpfulness and trustworthiness were frequently considered in previous studies (Jun, Vogt, & MacKay, 2010; Spark & Browning, 2011; Ye, Law, & Gu, 2009). For example, Ye et al. (2009) empirically studied the impact of online hotel reviews on hotel room sales. Their results indicated a significant relationship between the polarities of hotel reviews and online hotel bookings. Verna (2010) conducted a detailed literature review on customer choice modeling in hospitality domain. The investigated literatures showed that UGC and professional ratings on hotel have considerable impact for travelers on the choice of hotel. The results of Jun et al. (2010) also have similar findings that exploring online hotel reviews is one of the main search strategies for online hotel booking.

Previous researchers have also demonstrated both the importance and the influence of online hotel reviews in the tourism industry (Litvin et al., 2008; Liu & Park, 2015; Spark & Browning, 2011). As more and more travelers contribute their travel experience on travel websites, a huge amount of hotel reviews are generated daily. As a result, it becomes a tedious task for travelers to identify helpful reviews in a reasonable time. To address this issue, previous studies have investigated the helpfulness of online hotel reviews (Hwang, Lai, Chang, & Jiang, 2014; O'Mahony & Smyth, 2010; Zhu, Yin, & He, 2014). The concept of these studies is that if the review helpfulness can be determined once the reviews have been posted, then the travel websites can help online viewers evaluate the review quality and thus reduce their time for searching travel information.

Many critical factors affecting the review helpfulness were identified. O'Mahony and Smyth (2010) considered four kinds of review features from hotel reviews to predict review helpfulness, including content, sentiment, reputation, and social features. The complete sets of reviews from two US cities, Chicago and Las Vegas, were extracted from TripAdvisor.com. Their results showed that both sentiment and author reputation features have a significant impact on identifying helpful hotel reviews.

Hwang et al. (2014) evaluated the effects of three kinds of content features on hotel review helpfulness prediction. Specifically, the term frequency - inverse document frequency, topic-model-based Latent Dirichlet Allocation (LDA), and semantic-based LDA features, were extracted from online hotel reviews. The results show that content features are the most important predictors; the prediction models using content features achieve higher performance than those using semantic or review quality features.

Zhu et al. (2014) investigated the moderation effects of review rating extremity and hotel price on the relationship between reviewer credibility and review helpfulness. The two review author features, reviewer expertise (i.e., the number of Elite badges) and online attractiveness (i.e., the number of friends), were selected as predictors in the study. Their results indicated that the influence of the two predictors was moderated by hotel price, and the influence of reviewers having high expertise and attractiveness was also negatively moderated by review rating extremity.

Although the helpfulness of online hotel reviews has become an important topic in tourism and information technology literature, no research has considered all of the aforementioned critical factors to summarize online hotel reviews. In addition, prior studies mainly focus on identifying helpful hotel information on a review basis. From technical point of view, these studies adopted document-level information filtering approaches to identify useful reviews. However, the hotel reviews having been tagged as non-helpful may also contain valuable information. Unfortunately, the non-helpful reviews will be neglected by the document-level approaches, resulting in information loss. Instead of using document-based approaches, our study considers sentence-based information filtering approach to identify informative sentences from online hotel reviews.

2.2. Text summarization applications

Text summarization techniques have been used in various application domains, such as summaries of patents, webpages, and news articles (Jeong et al., 2016; Silla, Kaestner, & Freitas, 2003). Table 1 lists recent literature related to the text summarization in these domains.

Tseng, Wang, Lin, Lin, and Juang (2007) used a single-text summarization technique to generate patent summaries. The proposed approach considered the position of sentences in a paragraph to determine sentence importance. In addition, cue phrases in the sentences were also considered to determine the aims, purpose, or functions of a patent. Finally, the summary was formed by extracting and judging keywords. Trappey, Trappey, and Wu (2009) applied ontology tree structure and term frequency – inverse document frequency techniques for identifying keywords and retrieving the core content of a patent document. A clustering technique was then used to group and integrate sentences to derive a summary.

Regarding webpage summaries, Vazhenin, Ishikawa, and Klyuev (2009) conducted a query expansion by using WordNet and entered the expanded query into the Google search engine to locate relevant documents. The sentences containing related keywords were selected as the final summary. Kar et al. (2015) addressed a new research question concerning how to generate summaries in dynamic text collections. The proposed approach can automatically provide a summary depicting the prominent changes made to a document over a specific period of time. Specifically, the LDA technique was used to

Table 1
Recent study on text summarization.

Work	Document source	Content	Factor				
			Author	Time	Comment helpfulness	Semantic	Sentiment
Hu and Liu (2004)	Amazon	Product review				V	V
Zhuang et al. (2006)	IMdb	Movie review				V	V
Tseng et al. (2007)		Patent				V	
Meng and Wang (2009)	ZOL	Product review				V	
Trappey et al. (2009)	US patent	Patent				V	
Vazhenin et al. (2009)		Web page				V	
Kallimani et al. (2012)		News					
Atkinson and Munoz (2013)	N/A	News				V	
Wang et al. (2013)	Amazon	Product review				V	
Kar et al. (2015)	Wikipedia	Web page		V		V	
Lloret et al. (2015)	Amazon WhatCar	Product review				V	
Jeong et al. (2016)	AbleNews KORDIC	News				V	
Qiang et al. (2016)	DUC-2004	News				V	
This study	Tripadvisor	Hotel review	V	V	V	V	V

identify different topic structures of the document sets collected from different time periods. To evaluate the proposed method, the articles and their complete revision history were collected from Wikipedia.com. The results indicated the LDA-based approach has superior performance to other baseline approaches.

Regarding news summaries, Kallimani, Srinivasa, and Eswara Reddy (2012) proposed a statistical method whereby the title and first sentence of a news article were assigned a high weight (score). Moreover, the location and length of a sentence, term frequency, proper nouns, and terms containing capital letters were also weighted differently. Finally, all of the sentence scores were obtained and significant sentences were selected to form a summary. Qiang, Chen, Ding, Xie, and Wu (2016) developed a closed sequential pattern-based method for multi-document summarization. The method first scored each sentence by accumulating the weights of its covering patterns. In the sentence selection phase, the method iteratively selected the sentence that has (1) the highest score among all candidate sentences and (2) low similarity to the previously chosen sentences. Experimental evaluations on the DUC-2004 news dataset demonstrated that the proposed method significantly outperform not only all participating approaches on DUC-2004 competition but also some term- and ontology-based approaches.

The multi-text summarization technique has also been widely used for review summarization on numerous electronic commerce or social networking websites. Hu and Liu (2004) generated summaries for consumer reviews from Amazon, and Zhuang, Jing, and Zhu (2006) summarized movie reviews from IMDb. In addition to applying text-mining techniques to judge the similarity among sentences, they also considered possible conflicting comments and thus conducted sentiment analyses. Although the author information for each review could be retrieved on the previously mentioned websites, as well as time information and the helpfulness of reviews, neither of the two aforementioned studies considered all three factors. Wang, Zhu, and Li (2013) proposed a personalized text summarization technique that enabled users to select factors for summarizations according to their interest. Rice cooker reviews on Amazon.com were used as data for analysis. Turney (2002) collected reviews from Epinions and calculated the difference between the point-wise mutual information (PMI) of adjectives and adverbs in a review containing two sentiment words, "excellent" and "poor." The results can be used to judge the sentimental tendency and thus make recommendations. Meng and Wang (2009) collected Chinese reviews from ZOL.com, the largest online electronics store in China. They first extracted product specifications from ZOL.com. A product specification tree was then constructed using a clustering technique. The tree was used as a regular expression for identifying features from the product reviews. Finally, for each specification in the tree, frequently used adjectives in the product reviews were selected to form a review summary.

Lloret et al. (2015) developed a novel concept-level text summarization approach for generating ultra-concise opinion summaries. In the process of opinion summarization, five phases were considered: (1) syntactic text simplification; (2) text analysis; (3) concept representation; (4) surface representation; and (5) sentence selection. In experimental evaluation, a total of 400 reviews were collected from both Amazon.com and WhatCar.com. The results showed that the proposed approach outperforms a number of baseline methods.

In addition, text sentiment was considered in some of prior studies on text summarization, especially for the UGC (Hu & Liu, 2004; Zhuang et al., 2006). Kabadjov, Balahur, and Boldrini (2009) investigated the problem of whether sentences with strong sentiment polarity are good ones for text summarization. Although the results did not provide significant evidence for the association between sentiment intensity and text summarization quality, their study indicated that content- and sentiment-based summarizations should be performed separately. Balahur, Kabadjov, Steinberger, Steinberger, and Montoyo (2012) considered a number of different approaches concerning opinion summarization from multiple documents. Their results showed that integrating both topic-opinion analysis and semantic information can yield satisfactory results in opinion summarization.

Most of online social media provide information related to the post histories of review authors, as well as the time of posting and the helpfulness of the reviews. However, as shown in Table 1, previous studies did not consider all of these

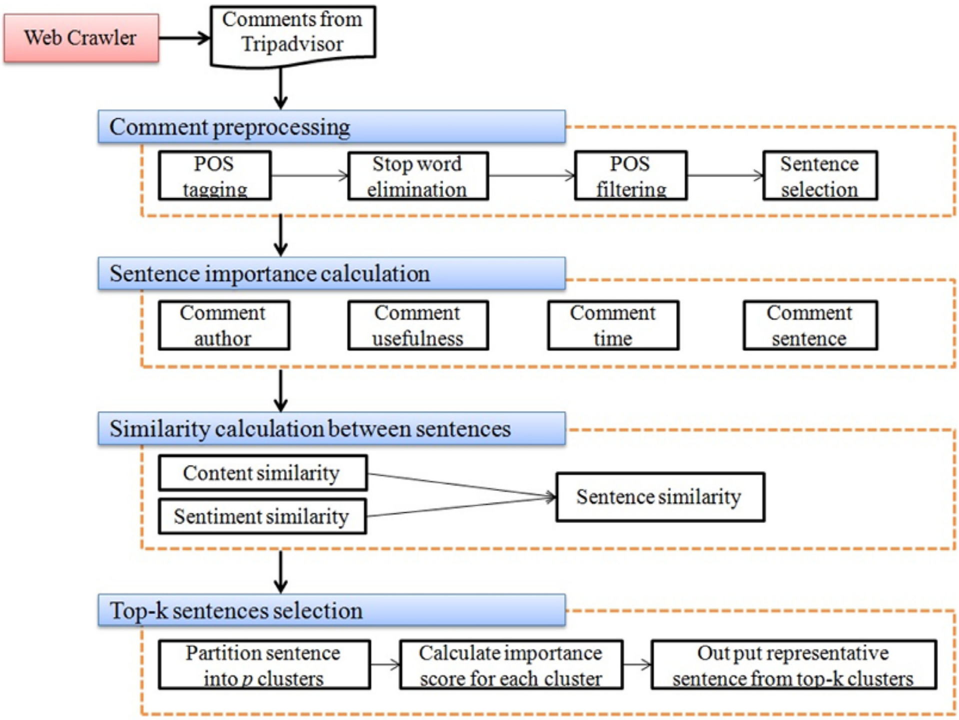

Fig. 1. Research process.

aspects. In addition, although text sentiment has been considered in the literatures, conflicting comments in the reviews posted by different authors were ignored. Because the aforementioned information can be critical factors to use for generating representative summarizations, we considered all of the factors, authors, review time, the helpfulness of reviews, and conflicting opinions, to develop a review summarization method.

3. Research method

Our research process (as shown in Fig. 1) can be divided into five principal steps: hotel review collection, review preprocessing, sentence importance calculation, sentence similarity calculation, and top-k sentence recommendations. First, hotel reviews were collected from TripAdvisor.com. The collected reviews were then preprocessed by removing all unnecessary information. The importance score of each sentence was then determined using the proposed approach. The conflicting review comments were also assessed by calculating similarities between two sentences. Finally, for each hotel, the k-medoids clustering technique was applied to discover the top-k most representative sentences for hotel information summarization.

3.1. Review preprocessing

We used sentences as the basic unit for analysis. The preprocessing tasks included part-of-speech (POS) tagging, stop-word elimination, POS filtering, and sentence selection. First, the Stanford Loglinear POS Tagger (Porter, 2006), developed by the Stanford Natural Language Processing Group, was adopted for the POS tagging. Each word in a sentence was allocated a POS tag, such as nouns (NN), adjectives (JJ), or adverbs (RB). The stop-word elimination step removed words that did not provide any valuable information in a sentence. In this study, we referred to the stop-word list provided by MySQL (http://dev.mysql.com/doc/refman/5.5/en/fulltext-stopwords.html). In the POS filtering step, this study retained only nouns, adjectives, and negative adverbs for each sentence. The primary reason is that most hotel reviews were expressed using nouns and adjectives, and negative adverbs substantially influence sentiment analysis. After the aforementioned preprocessing tasks, each sentence in a review must include at least one noun and one adjective.

3.2. Sentence importance calculation

Several factors that influence sentence importance of a review are considered in this study: (a) the representativeness of a review author, (b) helpfulness of a review, (c) review time, and (d) the content of sentences in a review.

The representativeness of a review author can be evaluated according to both author credibility and author recommendation score.

Definition 1. (Author credibility) Let r_h^a denote the rating of a hotel h given by a review author a, and ar_h the average rating of a hotel h. The author credibility of a review author a, AC_a, is defined using the following equation:

$$AC_a = 1 - \left(\frac{\sum_{h=1}^{H_a} \frac{\left| r_h^a - ar_h \right|}{5}}{H_a} \right), \tag{1}$$

where H_a is the total number of ratings of a review author a.

Author credibility is calculated based on the mean absolute error between the review author's ratings and the corresponding hotel's final ratings. The lower the error mean, the more the review author's ratings that agree with the majority viewpoint, and thus, the review author obtains high credibility. In this study, the minimum of H_a was set as 3; in other words, review authors who obtained at least three ratings were included.

Subsequently, the representativeness of a review author can be determined by calculating the total number of recommendations that authors obtained from their previous reviews. A high number of recommendations represents high representativeness.

Definition 2. (Author recommendation score) Let arn_a denote the average number of recommendations of a review author a. The author recommendation score of a, denoted as ARS_a, is defined as:

$$ARS_a = \begin{cases} 1, if \frac{\log_2(arn_a+1)}{2} \geq 1 \\ \frac{\log_2(arn_a+1)}{2}, \text{otherwise} \end{cases} \tag{2}$$

Based on Eq. (2), if arn_a is greater than or equal to 3, the reviews written by the review author would be deemed trustworthy and allotted a score of 1.

Finally, the representativeness of a review author is defined as the average of AC_a and ARS_a.

Definition 3. (Representativeness of review author) Given the author credibility $AC_a$and recommendation score ARS_a for the review author a, the representativeness of a, denoted as RCA_a, is defined using the following equation:

$$RCA_a = (AC_a + ARS_a)/2 \tag{3}$$

Definition 4. (Review helpfulness) Let crn_i denote the number of recommendations of the review i and let $\max(crn)$ denote the maximal number of recommendations among all of the reviews for a hotel. The review helpfulness, denoted as CH_i, is defined using Eq. (4):

$$CH_i = \frac{crn_i}{\max(crn)} \tag{4}$$

Generally, more recent reviews provide more accurate hotel information. Thus, the post time of a review is also a crucial factor.

Definition 5. (Review recency) Let t denote the time-interval between the review time and query time and let dm denote the time-interval between the first and final review for a hotel. The recency of a review i, denoted as CR_i, is defined using Eq. (5):

$$CR_i = \exp(-t/dm) \tag{5}$$

The factors determining the importance of each sentence in a document has been studied extensively in the past. First, if a sentence contains title words, capital letters (i.e., typically used for proper nouns or particularly emphasized words), or cue phrases (such as summary and results), its importance increases (McDonald & Chen, 2002). In addition, the position of the sentence is critical. For example, the first sentence in each paragraph typically involves more information regarding the paragraph (Edmundson, 1969), and longer sentences typically include more relevant information (Kupiec, Pedersen, & Chen, 1995).

This study defines the importance of each sentence in all reviews as follows. Three factors were considered in this study: sentence position, indicator phrase, and number of words and phrases in a sentence. First, according to Trappey et al. (2009), the first sentence or the title has high importance. Second, a sentence containing indicator phrases generally implies the sense of summarization, which increases the importance of the sentence. The indicator phrases considered in this study are listed in Table 2. Finally, the more words are contained in a sentence, the greater the possibility of containing a large amount of information. Thus, the importance of a sentence increases as the number of words increases.

Definition 6. (Review sentence score) Let $LOC(s_j)$ denote the function of measuring the importance of review location, where $LOC(s_j)=1$ if sentence s_j is the title or first sentence of a review, and otherwise, $LOC(_j)=0$. Let $IP(s_j)$ denote the function of identifying indicator phrases, where $IP(s_j)=1$ if s_j contains any indicator phrase, and otherwise, $IP(s_j)=0$. Let $NW(s_j)$ denote

Table 2
List of indicator phrases.

however	nevertheless	though	goal	summaries
although	summary	results	as a result	conclusion
invention	even though	intent	intention	discussion
conclusions	even if	purpose	in summary	all in all
discussions	objective	finally	Not with standing	result

the ratio of the number of words in s_j to the maximal number of words among all of the sentences. The review sentence score of s_j, denoted as CSS_j, is defined using Eq. (6):

$$CSS_j = w_1 \times LOC(s_j) + w_2 \times IP(s_j) + w_3 \times NW(s_j) \tag{6}$$

where w_1, w_2, and w_3 represent the weights of review location, indicator phrase, and number of words, respectively.

Finally, the importance of each sentence of a review can be obtained by combining the values of review author, review usefulness, review time, and review sentence.

Definition 7. (Sentence importance) Given a review i written by author a, the importance score of sentence s_j in review i, denoted as $SI_{a,\,i,\,j}$, is defined using Eq. (7):

$$SI_{a,i,j} = \frac{RCA_a + CH_i + CR_i}{3} \times CSS_j \tag{7}$$

3.3. Calculation of sentence similarity

Before we conducted sentence clustering for text summarization, defining the distance (or similarity) between two sentences was necessary. We considered two types of similarity, content and sentiment similarity, which used nouns and adjectives in similarity calculations, respectively.

To calculate content similarity, the distance (i.e., dissimilarity) of each noun-noun pair between two sentences was first calculated using Normalized Google Distance (NGD) (Cilibrasi & Vitanyi, 2007). NGD provides the number of hits (i.e., search results) for a set of keywords returned by the Google search engine, which can be used to define the semantic dissimilarity between two terms. Given two nouns n_x and n_y, the NGD between n_x and n_y is defined as follows:

$$NGD(n_x, n_y) = \frac{\max\{\log f(n_x), \log f(n_y)\} - \log f(n_x \& n_y)}{\log G - \min\{\log f(n_x), \log f(n_y)\}} \tag{8}$$

where G represents the total number of webpages in the Google search engine and $f(n_x)$, $f(n_y)$, and $f(n_x\&n_y)$ represent the number of hits returned by the Google search engine for n_x, n_y, and $(n_x\&n_y)$, respectively.

The similarity between two nouns can be easily transformed using the following equation:

$$Sim_{NGD}(n_x, n_y) = 1 - NGD(n_x, n_y) \tag{9}$$

Definition 8. (Content similarity) Assume that two sentences s_j and $s_{j'}$ have m and n nouns, respectively. Based on Eq. (9), a total of $(m \times n)$ Sim_{NGD} values can be obtained for each noun-noun pair between s_j and $s_{j'}$. Given a user-specified threshold β, the content similarity between s_j and $s_{j'}$ is defined as follows:

$$ContentSim(S_j, S_{j'}) = \frac{count_\beta(s_j, s_{j'})}{m \times n} \tag{10}$$

where $count_\beta(s_j, s_{j'})$ represents the total number of noun-noun pairs with Sim_{NGD} values satisfying β.

To select an optimal value for β, a preliminary experiment was conducted. First, we randomly selected 10 terms from each of the two topics, hotel location and facility. The 10 selected terms for location were *location, view, station, mall, distance, airport, bus, block, minute, walking*. The 10 selected terms for facility were *room, lobby, Internet, facility, pool, parking, bed, Wi-Fi, toilet, bathroom*. All of these terms generated 190 noun-noun pairs (100 intertopic and 90 intratopic pairs) and their corresponding Sim_{NGD} values. Subsequently, the 190 noun-noun pairs were sorted in descending order according to their Sim_{NGD} values. The optimal value for β was determined by selecting a Sim_{NGD} value from the ordered list that could minimize the possibility of intertopic pairs. According to the result, the optimal value for β was 0.65.

To calculate sentiment similarity, the Semantic Orientation-Pointwise Mutual Information (SOPMI) method (Turney & Littman, 2003) was used for calculating the positive and negative sentiment strength scores of a sentence. In this study, we considered the lists of positive (i.e., good, nice, excellent, positive, fortunate, correct, and superior) and negative (i.e., bad, nasty, poor, negative, unfortunate, wrong, and inferior) adjectives proposed by Missen, Boughanem, and Cabanac (2013) to determine the sentiment strength score for each adjective in a review. In particular, the sentiment strength score of an adjective a_x is defined as follows:

$$O(a_x) = \sum_{t_l \in A_{pos}} SOPMI(a_x, t_l) - \sum_{t_l \in A_{neg}} SOPMI(a_x, t_l) \tag{11}$$

$$SOPMI(a_x, t_l) = \log_2\left(\frac{p(a_x \& t_l)}{p(a_x) \times p(t_l)}\right) \tag{12}$$

where t_l is a term in the positive or negative adjective list; $p(a_x)$, $p(t_l)$, and $p(a_x\&t_l)$ are the number of observations of a_x, t_l, and $(a_x\&t_l)$ occurring in a corpus, respectively, normalized by the size of the corpus.

It is worth noting that sentiment polarity might be affected by negative adverbs. If a negative adverb occurs near an adjective, the sentiment strength score for the adjective is adjusted in the opposite direction (i.e., $-O(a_x)$).

Finally, the sentiment strength score of sentence s_j, denoted as $O(s_j)$, is equal to the average of the sentiment strength score of every adjective in s_j, as expressed by the following formula:

$$O(s_j) = \frac{\sum O(a_x)}{\text{\# of adjetives in } s_j} \tag{13}$$

Based on Eq. (13), the sentence polarity of s_j is defined using Eq. (14):

$$SP(s_j) = \begin{cases} 1, & \text{if } O(s_j) > r \\ 0.5, & \text{if } \left|O(s_j)\right| \le r \\ 0, & \text{if } O(s_j) < -r \end{cases} \tag{14}$$

where r is a user-defined threshold.

As shown in Eq. (14), the sentiment polarity is positive if $SP(s_j)=1$, neutral if $SP(s_j)=0.5$, and negative if $SP(s_j)=0$.

The setting of threshold r is based on an experiment. First, 100 sentences were randomly selected and manually assigned a sentiment polarity (i.e., positive, neutral, or negative) as the answer set. The sentiment scores for these sentences were then calculated using SOPMI, and the varying threshold r generated different sentiment polarity results for comparison with the answer set. Thus, the optimal r value was determined by selecting the result exhibiting the lowest error rate. According to the result, the optimal value for r was 0.2.

Definition 9. (Sentiment similarity) Given two sentences s_j and $s_{j'}$, the sentiment similarity between these two sentences is defined as follows:

$$SentiSim(O(s_j), O(s_{j'})) = \begin{cases} 1, & \text{if } SP(s_j) = SP(s_{j'}) \\ 0.5, & \text{if } SP(s_j) = 0.5 \text{ } or \text{ } SP(s_{j'}) = 0.5 \\ 0, & \text{if } \left|SP(s_j) - O(s_{j'})\right| = 1 \end{cases} \tag{15}$$

Definition 10. (Sentence similarity) Following Definitions 8 and 9, the sentence similarity between two sentences s_j and $s_{j'}$, denoted as $SenSim(s_j, s_{j'})$, is defined using Eq. (16):

$$SenSim(s_j, s_{j'}) = ContentSim(s_j, s_{j'}) \times SentiSim(s_j, s_{j'}) \tag{16}$$

3.4. Selection of the top-k sentences

To select the top-k sentences, this study involved first partitioning all of the sentences into k clusters by using the k-medoids algorithm. For each cluster, the sentence nearest to its centroid was then used as the representative sentence. However, the collected sentences might have contained trivial information (i.e., noise data), which influenced the clustering result of the k-medoids algorithm. To solve this problem, this study first used the k-medoids algorithm to partition sentences into p clusters ($p > k$).

Provided with a set of sentences in which each sentence was represented as a vector, and a user-specified number of p clusters, the k-medoids algorithm begins by randomly selecting p sentences as an initial set of medoids (i.e., the representative objects in a cluster). The cosine similarity between each sentence and each medoid was then calculated, and each sentence was assigned to its nearest medoid. After each sentence was assigned to a specific medoid to form a cluster, the algorithm calculated the average similarity of each sentence to other sentences in the same cluster. The sentence exhibiting the highest average similarity was then used as a new medoid. The procedure was iterated until all medoids were stable or the criterion function converged.

Once the p sentence clusters are determined using the k-medoids algorithm, the importance score of a cluster can be calculated by summing the importance score of each sentence (i.e., $SI_{a,\,i,\,j}$). Finally, the top-k sentences can be determined by collecting representative sentences from the top-k clusters exhibiting high importance scores.

Table 3
Hotel information.

Hotel	City	Star	Average rating	# of sentences
Gansevoort Meatpacking NYC	New York	5	4	474
Red Roof Inn	Chicago	2	3	481

Table 4
Weight information.

	1	2	3	4	5	6	7	8	9	10
w_1	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.3	0.4
w_2	0.1	0.2	0.3	0.1	0.2	0.3	0.1	0.2	0.3	0.4
w_3	0.8	0.7	0.6	0.7	0.6	0.5	0.6	0.5	0.4	0.2

4. Experimental evaluation

4.1. Data collection

In this study, the two hotels, Red Roof Inn and Gansevoort Meatpacking Hotel, were selected from TripAdvisor.com. All of the reviews regarding the two investigated hotels were collected from January 1, 2012 to March 31, 2013. However, the reviews that we analyzed presented the following limitations: (a) TripAdvisor includes reviews in several languages, such as Chinese, Japanese, and French. This study performed analyses of the English reviews only. (b) TripAdvisor is a cross- platform website. In addition to the reviews posted by members, the website also includes reviews posted by members from other travel websites, such as Hotel.com and Expedia.com. In this study, the collected reviews were all posted by TripAdvisor members. (c) The reviews that we collected were written by review authors who had reviewed on more than two hotels.

Table 3 shows the hotel information and the number of sentences that we collected. From the collected reviews, we randomly selected 50 reviews for each hotel in the experiments.

4.2. Experimental setup

We considered three possible approaches for determining top-k sentences in the experimental study. The first approach (A) treated all reviews as a single document. Only the importance score of review sentences (i.e., $SI_{a,\,i,\,j}$ defined in Section 3.2) was considered and all other factors were ignored. The second approach (B) treated all reviews as a single document. The sentences in this document were partitioned into k clusters by using the k-medoids algorithm. For each cluster, the sentence that exhibited minimal distance to the centroid was used as a representative sentence. The third approach (C) considered all aspects that we defined in Section 3 (i.e., the proposed method).

The parameter settings primarily consisted of two parts: three weights in the calculation of review sentence scores and a number of clusters (i.e., parameters p and k in Section 3.4) during the sentence clustering phase.

For determining the three weight values, a few experts were invited to review and rank 10 randomly selected sentences based on the importance of the sentences. For the same sentence, 10 possible weight settings, shown in Table 4, were considered to calculate the scores. Then, the sentences were ranked based on the proposed approach. The ranking list that yielded the highest similarity with the experts' ranking was used as the final weight. Based on the results, the three weight values, w_1, w_2, and w_3, were 0.3, 0.1, and 0.6, respectively. For determining parameters p and k, we considered the following two settings: ($k=5,\ p=8$) and ($k=10,\ p=15$), which represented the short and long summaries, respectively.

A total of 20 participants were invited to first review all of the selected reviews (50 reviews) for each hotel. The sample included 20 graduate students majored in information systems and had experience on browsing hotel reviews in TripAdvisor.com. Of the students who participated in the text summarization usability survey, 13 (65%) were female and 7 (35%) were male.

To fairly evaluate the performance of the three approaches, we adopted Latin square design. Specifically, the three top-k sentence lists generated using the three approaches were provided, and then participants blindly accessed the three sentence sets. After the evaluation, each participant was asked to rank the three approaches based on the usefulness of the three sentence sets. The most useful approach received 3 points, followed by 2 points, and the worst approach received 1 point.

As mentioned in Section 2.1, except the review content, the review helpfulness also has strong relationships with author reliability, review time, and review polarity features. Therefore, this study develops a text summarization method (i.e., method C) with the consideration of all the aforementioned features. The two conventional methods (i.e., methods A and B) were implemented based on content features only. To examine the effects of these features on review summarization, the following hypotheses are then postulated and tested:

H1. The sentences selected by method C are more useful than those selected by method A;

H2. The sentences selected by method C are more useful than those selected by method B.

Table 5
Summary statistics.

Hotel	Top-k	Method	# of subjects			Mean	Std. dev.
			3 points	2 points	1 point		
Red Roof Inn	5	A	0	10	10	1.5	.51299
		B	2	10	8	1.7	.65695
		C	18	0	2	2.8	.61559
	10	A	0	3	17	1.15	.36635
		B	5	12	3	2.1	.64072
		C	15	5	0	2.75	.44426
Gansevoort Meatpacking	5	A	4	2	14	1.5	.82717
		B	5	10	5	2.0	.72548
		C	14	5	1	2.5	.60698
	10	A	0	15	5	1.75	.44426
		B	3	4	13	1.5	.76089
		C	17	1	2	2.75	.63867

Table 6
Results of significance tests.

Hotel	# of sentences	Methods	Z	P
Red Roof Inn	Top 5	C – A	−3.615	.000***
		C – B	−2.846	.004**
	Top 10	C – A	−4.053	.000***
		C – B	−2.419	.016*
Gansevoort Meatpacking hotel	Top 5	C – A	−2.807	.005**
		C – B	−1.966	.049*
	Top 10	C – A	−3.503	.000***
		C – B	−2.992	.003**

Note: *Significance: $p < 0.05$. **Significance: $p < 0.01$. ***Significance: $p < 0.001$.

4.3. Results

Table 5 presents the evaluation results of three summarization methods for the two hotels. Regarding the Red Roof Inn, method C exhibited scores of 2.8 and 2.75 in $k = 5$ and $k = 10$, respectively. Regarding the Gansevoort Meatpacking hotel, method C exhibited scores of 2.5 and 2.75 in $k = 5$ and $k = 10$, respectively. Due to the small sample sizes, the nonparametric Wilcoxon sign rank test was performed to separately test both H1 and H2, using the four data sets. As shown in Table 6, the results indicate that the usefulness score of the sentences generated by method C was significantly higher than those generated by methods A and B at the 0.05 significance level ($p < 0.05$). In summary, the results indicate that method C, proposed in this study, significantly outperformed the other two conventional methods.

All of the participants were invited for a brief interview after the experiments. Most of the participants reported that method C provided highly comprehensive information that covered several aspects, such as information on the hotel location, room environment, and service attitudes. The sentences selected using method C were specific and provided substantial information, and thus, met the expectations of most of the participants.

We further investigated the sentences selected using the three methods. As shown in Table 7, we observed that the sentences selected using method A primarily contained information on location and price, but rarely included room and service information. For example, both the third sentence, "*great location for the price–great staff*," and the seventh sentence, "*good location and price average place*," mentioned the "*good*" location of the hotel, but these sentences did not specify why the location of the hotel was good (e.g., near a shopping mall or airport), producing poor information content. A possible reason is that Method A involved selecting sentences according to the importance score of review sentences, which considered only the explicit features of a sentence (i.e., sentence position, whether it contained indicator phrases, and the number of keywords). Thus, the sentence title, first sentence of a review, and sentences expressing a sense of conclusion exhibit a strong possibility of being selected. Conversely, method B involved using the k-medoids algorithm to cluster sentences before sentence selection, and thus, the generated sentence list covered additional factors. However, because Method B did not consider the importance of a review sentence, the possibility of selecting meaningless sentences (i.e., noise) increased. For example, the 10th sentence, "*my head is shaved so there is no doubt they were not mine*," cannot provide any useful information to readers.

Regarding method C, we considered both the sentence importance score and sentence clustering methods. Thus, the sentences selected using method C not only contained information on additional aspects (i.e., location, price, room, environment, and service) but also reduced the chances of meaningless sentences being included. As shown in Table 7, we observed that the first sentence, "*most hotel staff are professional and many are even polite, but few are also kind and personable*," specif-

Table 7
Top-10 sentences for Red Roof Inn.

Chicago Red Roof Inn ($p=15$, $k=10$)	
Method A	
1.	The walk from union station to red roof is pleasant
2.	We checked-in last month with my family at the red roof inn to take advantage of their 'red deals' to save some money
3.	Great location for the price–great staff
4.	Just had a wonderful few days in the red roof inn Chicago
5.	The hotel is indeed in a good location, but that is the only positive thing i can say about it
6.	Beware of vermin!
7.	Good location and price average place
8.	A group of us booked a hand full of room here for labor day weekend
9.	I was looking for a clean safe place for a couple of nights
10.	We stayed at the red roof inn on east ontario because of the location, and the rates were reasonable
Method B	
1.	Beware of vermin!
2.	Good location and price average place
3.	The building is old, the elevators are slow and the service is sometimes slow and inadequate
4.	The other complaints made here by other users are true
5.	This was a good choice for us because it was relatively inexpensive for the location, only blocks away from the miracle mile, and walking distance to the navy pier and various other attractions
6.	If you're not fussed about an older building and dated decor (carpets, etc), you'll likely find red roof to be value for money
7.	It was also very gross to find random hairs stuck to the walls and chairs
8.	Could not have a better trip
9.	We were so close to all the sights and enjoyed how easy it was to get around
10.	My head is shaved so there is no doubt they were not mine
Method C	
1.	Most hotel staff are professional and many are even polite, but few are also kind and personable
2.	We stayed at the red roof inn on east ontario because of the location, and the rates were reasonable
3.	The money you save you can spend on all the nice restaurants and bars in the area
4.	On rare occasions, there are some noisy guests, and someone left some beer cans on the elevator, the front desk took care of it immediately
5.	The tv stopped working randomly saturday morning, but the staff seemed so busy i did not ask for it to be fixed until sunday morning
6.	Rooms are disorganized and not as clean and fresh you would expect from a hotel especially in that part of town
7.	Good location and price average place
8.	The building is old, the elevators are slow and the service is sometimes slow and inadequate
9.	It is only a 10 min walk from water tower place and about 20 min walk to downtown
10.	rooms: dated and a little dark, but fine for a 2

ically described service attitude, and the ninth sentence, "*it is only a 10 minute walk from water tower place and about 20 minutes' walk to downtown*," detailed why the location of the hotel is favorable.

Similar results were also yielded for Gansevoort Meatpacking hotel. As shown in Table 8, the sentences selected using method A primarily described location information; those using method B included room, environment, and facility information; and those using method C covered the most substantial information, including facilities, prices, and room and service information, among all three methods.

5. Practical and theoretical implications

A number of practical implications may be derived from this study. First, the developed review summarization technique can be served as a guideline to develop a smart tourism information system (STIS) for travel websites. Specifically, when a user browses the reviews of a hotel in a travel website, the STIS can directly generate a summarized hotel review from thousands of reviews of the target hotel. Because online viewers usually have limited time to handle a large amount of online hotel reviews, this system can help users quickly grasp important information of the selected hotels and thus save time during their online booking process.

Second, most of the current travel websites allow online viewers to sort the hotel reviews by date, review rating, or author reputation. They basically did not provide an intelligent system to help online viewers to filter out less informative reviews. For example, the hotel reviews in TripAdvisor.com can be sorted only by the review posted date and review rating, but these two features cannot truly reflect the helpfulness of reviews. With the use of the developed STIS, travel websites can provide additional query functions to filter and rank reviews, including the functions like (a) grouping reviews by review polarity, (b) filtering out the reviews with similar content, and (c) sorting reviews by the number of informative sentences in a review.

Table 8
Top-5 sentences for Gansevoort Meatpacking.

New York Gansevoort Meatpacking ($p=8$, $k=5$)	
Method A	
1.	We stayed here in a back in November 2012, smack in the middle of the meat packing district with fantastic places to eat at anytime of the day on your doorstep
2.	Hotel gansevoort is a modern hotel in the meatpacking district - a convenient walk, metro, or taxi to just about anywhere in nyc
3.	We were a few years behind the curve in terms of visiting this place, but its hip credentials still seemed to be well in place, right down to a group of beautiful people, dressed completely in black, swaying out the entrance as we lurched, all too colorfully, in
4.	First of all the location of this hotel is perfect-meat packing district is really great part of the city to stay in
5.	Nice hotel in “cool” district. beware that people party late
Method B	
1.	The cutler bath amenities were lovely and plentiful
2.	bc the valet was overwhelmed with clubgoers and cabs
3.	If you need at least a few hours of sleep each night and do not sleep like the dead then choose somewhere a little quieter
4.	For you party people: there is a nice rooftop bar on 14th floor which gets quite packed on weekends
5.	They should either have them wait or designate a separate elevator for guests and for nonguests
Method C	
1.	The roof top bar (plunge) and its adjoining swimming pool affords nice views of the city, but the place swarms with the urgently hip on the weekends, making good time hard to come by
2.	Great hotel with reasonable price clean and comfortable room
3.	On our first day there the guy that was ’manning’ the pool area brought out freezies for the kids (and parents) and it was such a nice thing to do, we were all in shock
4.	Its location in the meatpacking district is ideal for people aged 20 to 40 i’d say
5.	The cutler bath amenities were lovely and plentiful

The results of the current research can also contribute to the literature by proposing a new research direction for online hotel reviews. First, as we mentioned in Section 2.1, previous studies mainly investigate critical factors associated with hotel review helpfulness. Their aim is to remove less informative reviews and identify a set of helpful reviews for a travel website. From the technical point of view, this approach deals with information filtering problem on a document basis. In this study, we further address the problem of information filtering from document-level to sentence-level granularity, which may shed some light on the use of text mining methodology to identify valuable information from online hotel reviews.

Our findings may have an important implication for other researchers by providing a new perspective on determining the importance of the sentence from the UGC. This study is the first to simultaneously consider author reliability, review time, review rating, and review polarity features on review summarizations. Moreover, to extract the most informative sentences from online reviews, the proposed approach uses a modified k-medoids clustering technique. The results of this study indicate that conventional content-based summarization techniques are less efficient to generate a useful summary from online hotel reviews. Except the review content, all the selected features in this study are also important to summarize the hotel reviews.

6. Conclusions

This research proposes a novel text summarization technique to determine the top-k most informative sentences from online hotel reviews. Most previous studies on review summarizations have focused on text preprocessing, which disregards critical factors such as the credibility of review authors, review time, review usefulness, and conflicting opinions. In addition to the information (i.e., keywords and key phrases) extracted from the hotel reviews, this study also considered all the above-mentioned factors. In particular, we used the review authors’ historical data to calculate their credibility, the review time to determine the value of a hotel review, and the number of recommendations to define the usefulness of a hotel review. Sentence positions, sentence length, whether a sentence contained indicator phrases, and all of the aforementioned factors were then combined to obtain the sentence importance score. Furthermore, we considered text preprocessing techniques, such as NGD and PMI, to manage conflicting opinions among the sentences. Subsequently, the k-medoids clustering technique was used to filter noisy sentences and then identify the top-k sentences as the review summarization result.

In our experiments, a total of 20 participants were invited to empirically compare the performance of the proposed research method and two other conventional methods. The results indicated that the sentences summarized using the proposed method can provide comprehensive information on both positive and negative aspects.

This study has a number of limitations. This study included a fairly small sample of participants in the experiments and all the participants have similar background (i.e., graduate students). Using such a homogeneous group of participants may limit the generalizability of this study. Future research can utilize a larger number of subjects that can truly represent the population. Also, this study utilized a limited number of reviews for two hotels. It is more desirable to include more hotel reviews in the experiments to verify the effectiveness of the proposed text summarization technique. In addition, an assumption of 'equal weight' was implicitly adopted for the variables used in sentence importance calculation, which is not often the case in practice. When adopting the proposed text summarization technique, one can employ the analytic hierarchy process or conduct supervised learning techniques to determine the weight of each variable.

Several extensions related to the development of text summarization system in travel website are also possible. First, this study considered only nouns, adjectives, and negative adverbs for sentiment analysis. Future studies can include both verbs and degree adverbs to facilitate assessing sentiment polarity. Because sentiment polarity can be evaluated using certain verbs, such as *like, hate*, or *love*, degree adverbs can represent the degree of sentiment polarity. Additionally, users might have their own preferences for selecting hotels. One possible research direction is to construct personalized recommender systems based on user profiles. In calculating the sentence importance scores, the weight of each factor can be customized according to each user's preference. Thus, different top-*k* sentences can be summarized for different users. Finally, this study focused on developing an extractive summarization technique, which did not consider the context of the sentences. To improve the readability of summarization results, future studies can involve developing abstractive summarization methods.

References

Abdi, A., Idris, N., Alguliev, R. M., & Aliguliyev, R. M. (2015). Automatic summarization assessment through a combination of semantic and syntactic information for intelligent educational systems. *Information Processing & Management, 51*(4), 340–358.

Ady, M., & Quadri-Felitti, D.. *Consumer research identifies how to present travel review content for more bookings* (Accessed 11.04.15) http://www.trustyou.com/press/study-reveals-travelers-prefer-summarized-review-content-full-text-reviews.

Atkinson, J., & Munoz, R. (2013). Rhetorics-based multi-document summarization. *Expert Systems with Applications, 40*(11), 4346–4352.

Balahur, A., Kabadjov, M., Steinberger, J., Steinberger, R., & Montoyo, A. (2012). Challenges and solutions in the opinion summarization of user-generated content. *Journal of Intelligent Information Systems, 39*(2), 375–398.

Cantallops, A. S., & Salvi, F. (2014). New consumer behavior: A review of research on eWOM and hotels. *International Journal of Hospitality Management, 36*, 41–51.

Casaló, L. V., Flavián, C., & Guinalíu, M. (2010). Determinants of the intention to participate in firm-hosted online travel communities and effects on consumer behavioral intentions. *Tourism Management, 31*(6), 898–911.

Chung, N., & Koo, C. (2015). The use of social media in travel information search. *Telematics and Informatics, 32*(2), 215–229.

Chung, N., Han, H., & Koo, C. (2015). Adoption of travel information in user-generated content on social media: The moderating effect of social presence. *Behaviour & Information Technology, 34*(9), 902–919.

Cilibrasi, R. L., & Vitanyi, P. M. (2007). The google similarity distance. *IEEE Transactions on Knowledge and Data Engineering, 19*(3), 370–383.

Edmundson, H. P. (1969). New methods in automatic extracting. *Journal of the ACM, 16*(2), 264–285.

Gambhir, M., & Gupta, V. (2016). Recent automatic text summarization techniques: a survey. *Artificial Intelligence Review*. doi:10.1007/s10462-016-9475-9.

Gupta, V., & Lehal, G. S. (2010). A survey of text summarization extractive techniques. *Journal of Emerging Technologies in Web Intelligence, 2*(3), 258–268.

Harer, S., & Kadam, S. (2014). Mining and summarizing movie reviews in mobile environment. *International Journal of Computer Science & Information Technologies, 5*(3), 3912–3916.

Heu, J. U., Qasim, I., & Lee, D. H. (2015). FoDoSu: Multi-document summarization exploiting semantic analysis based on social Folksonomy. *Information Processing & Management, 51*(1), 212–225.

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* (pp. 168–177). ACM.

Hwang, S. Y., Lai, C. Y., Chang, S., & Jiang, J. J. (2014). The Identification of Noteworthy Hotel Reviews for Hotel Management. *Pacific Asia Journal of the Association for Information Systems, 6*(4), 1–17.

Jannach, D., Zanker, M., & Fuchs, M. (2014). Leveraging multi-criteria customer feedback for satisfaction analysis and improved recommendations. *Information Technology & Tourism, 14*(2), 119–149.

Jeong, H., Ko, Y., & Seo, J. (2016). How to improve text summarization and classification by mutual cooperation on an integrated framework. *Expert Systems with Applications, 60*, 222–233.

Jun, S. H., Vogt, C. A., & MacKay, K. J. (2010). Online information search strategies: a focus on flights and accommodations. *Journal of Travel & Tourism Marketing, 27*(6), 579–595.

Kabadjov, M., Balahur, A., & Boldrini, E. (2009). Sentiment intensity: is it a good summary indicator? In *Language and Technology Conference* (pp. 203–212). Springer Berlin Heidelberg.

Kar, M., Nunes, S., & Ribeiro, C. (2015). Summarization of changes in dynamic text collections using Latent Dirichlet Allocation model. *Information Processing & Management, 51*(6), 809–833.

Kallimani, J. S., Srinivasa, K. G., & Eswara Reddy, B. (2012). Summarizing news paper articles: experiments with ontology-based, customized, extractive text summary and word scoring. *Cybernetics and Information Technologies, 12*(2), 34–50.

Kim, E. E. K., Mattila, A. S., & Baloglu, S. (2011). Effects of gender and expertise on consumers' motivation to read online hotel reviews. *Cornell Hospitality Quarterly, 52*(4), 399–406.

Kupiec, J., Pedersen, J., & Chen, F. (1995). A trainable document summarizer. In *Proceedings of the 18th annual international ACM SIGIR conference on research and development in information retrieval* (pp. 68–73). ACM.

Li, G., Law, R., Vu, H. Q., Rong, J., & Zhao, X. R. (2015). Identifying emerging hotel preferences using emerging pattern mining technique. *Tourism Management, 46*, 311–321.

Litvin, S. W., Goldsmith, R. E., & Pan, B. (2008). Electronic word-of-mouth in hospitality and tourism management. *Tourism management, 29*(3), 458–468.

Liu, C. L., Hsaio, W. H., Lee, C. H., Lu, G. C., & Jou, E. (2012). Movie rating and review summarization in mobile environment. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 42*(3), 397–407.

Liu, C. Y., Chen, M. S., & Tseng, C. Y. (2015). Incrests: Towards real-time incremental short text summarization on comment streams from social network services. *IEEE Transactions on Knowledge and Data Engineering, 27*(11), 2986–3000.

Liu, Z., & Park, S. (2015). What makes a useful online review? Implication for travel product websites. *Tourism Management, 47*, 140–151.

Lloret, E., Boldrini, E., Vodolazova, T., Martínez-Barco, P., Muñoz, R., & Palomar, M. (2015). A novel concept-level approach for ultra-concise opinion summarization. *Expert Systems with Applications, 42*(20), 7148–7156.

Loureiro, S. M. C., & Kastenholz, E. (2011). Corporate reputation, satisfaction, delight, and loyalty towards rural lodging units in Portugal. *International Journal of Hospitality Management, 30*(3), 575–583.

Ly, D. K., Sugiyama, K., Lin, Z., & Kan, M. Y. (2011, June). Product review summarization from a deeper perspective. In *Proceedings of the 11th annual international ACM/IEEE joint conference on Digital libraries* (pp. 311–314). ACM.

Mani, I., & Maybury, M. T. (1999). Advances in automatic text summarization *(*: Vol. 293. Cambridge, MA: MIT press.

McDonald, D., & Chen, H. (2002, July). Using sentence-selection heuristics to rank text segments in TXTRACTOR. In *Proceedings of the 2nd ACM/IEEE-CS joint conference on digital libraries* (pp. 28–35). ACM.

Mehta, P. (2013). Survey on movie rating and review summarization in mobile environment. *International Journal of Engineering Research and Technology, 2*(March (3)) ESRSA Publications.
Meng, X., & Wang, H. (2009). Mining user reviews: from specification to summarization. In *Proceedings of the ACL-IJCNLP 2009 conference short papers* (pp. 177–180). Association for Computational Linguistics.
Missen, M. M. S., Boughanem, M., & Cabanac, G. (2013). Opinion mining: Reviewed from word to document level. *Social Network Analysis and Mining, 3*(1), 107–125.
O'Mahony, M. P., & Smyth, B. (2010). A classification-based review recommender. *Knowledge-Based Systems, 23*(4), 323–329.
Qiang, J. P., Chen, P., Ding, W., Xie, F., & Wu, X. (2016). Multi-document summarization using closed patterns. *Knowledge-Based Systems, 99*, 28–38.
Papathanassis, A., & Knolle, F. (2011). Exploring the adoption and processing of online holiday reviews: A grounded theory approach. *Tourism Management, 32*(2), 215–224.
Peetz, M. H., de Rijke, M., & Kaptein, R. (2016). Estimating reputation polarity on microblog posts. *Information Processing & Management, 52*(2), 193–216.
Porter, M. F. (2006). An algorithm for suffix stripping. *Program, 40*(3), 211–218.
Sankarasubramaniam, Y., Ramanathan, K., & Ghosh, S. (2014). Text summarization using Wikipedia. *Information Processing & Management, 50*(3), 443–461.
Silla, C. N., Jr, Kaestner, C. A., & Freitas, A. A. (2003). A non-linear topic detection method for text summarization using wordnet. *Workshop of technology information language human (TIL'2003)*: Vol. 24.
Sparks, B. A., & Browning, V. (2011). The impact of online reviews on hotel booking intentions and perception of trust. *Tourism Management, 32*(6), 1310–1323.
Stringam, B. B., & Gerdes, J., Jr (2010). An analysis of word-of-mouse ratings and guest comments of online hotel distribution sites. *Journal of Hospitality Marketing & Management, 19*(7), 773–796.
Swanson, S. R., & Hsu, M. K. (2009). Critical incidents in tourism: Failure, recovery, customer switching, and word-of-mouth behaviors. *Journal of Travel & Tourism Marketing, 26*(2), 180–194.
Tseng, Y. H., Wang, Y. M., Lin, Y. I., Lin, C. J., & Juang, D. W. (2007). Patent surrogate extraction and evaluation in the context of patent mapping. *Journal of Information Science, 33*(6), 718–736.
Trappey, A. J., Trappey, C. V., & Wu, C. Y. (2009). Automatic patent document summarization for collaborative knowledge systems and services. *Journal of Systems Science and Systems Engineering, 18*(1), 71–94.
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* (pp. 417–424). Association for Computational Linguistics.
Turney, P. D., & Littman, M. L. (2003). Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems, 21*(4), 315–346.
Vazhenin, D., Ishikawa, S., & Klyuev, V. (2009). A user-oriented web retrieval summarization tool. In *Proceedings of the 2009 Second international conference on advances in human-oriented and personalized mechanisms, technologies, and services* (pp. 73–78). IEEE Computer Society.
Verma, R. (2010). Customer-choice modeling in hospitality services: A review of past research and discussion of some new applications. *Cornell Hospitality Quarterly, 51*(4), 470–478.
Vermeulen, I. E., & Seegers, D. (2009). Tried and tested: The impact of online hotel reviews on consumer consideration. *Tourism Management, 30*(1), 123–127.
Wang, D., Zhu, S., & Li, T. (2013). SumView: A Web-based engine for summarizing product reviews and customer opinions. *Expert Systems with Applications, 40*(1), 27–33.
Wen, I. (2009). Factors affecting the online travel buying decision: A review. *International Journal of Contemporary Hospitality Management, 21*(6), 752–765.
Yacouel, N., & Fleischer, A. (2012). The role of cybermediaries in reputation building and price premiums in the online hotel market. *Journal of Travel Research, 51*(2), 219–226.
Ye, Q., Law, R., & Gu, B. (2009). The impact of online user reviews on hotel room sales. *International Journal of Hospitality Management, 28*(1), 180–182.
Ye, Q., Law, R., Gu, B., & Chen, W. (2011). The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e–word-of-mouth to hotel online bookings. *Computers in Human Behavior, 27*(2), 634–639.
Zhu, L., Yin, G., & He, W. (2014). Is this opinion leader's review useful? Peripheral cues for online review helpfulness. *Journal of Electronic Commerce Research, 15*(4), 267.
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* (pp. 43–50). ACM.