J Syst Sci Syst Eng (Oct 2018) 27(5):542-558
DOI: https://doi.org/10.1007/s11518-018-5388-2
ISSN: 1004-3756 (Paper) 1861-9576 (Online)
CN11-2983/N

PRODUCTS RANKING THROUGH ASPECT-BASED SENTIMENT ANALYSIS OF ONLINE HETEROGENEOUS REVIEWS

Chonghui Guo Zhonglian Du Xinyue Kou

Institute of Systems Engineering, Dalian University of Technology, Dalian 116024, China

dlutguo@dlut.edu.cn (✉); *duzhonglian@mail.dlut.edu.cn; xxyykou@mail.dlut.edu.cn*

Abstract

With the rapid growth of online shopping platforms, more and more customers intend to share their shopping experience and product reviews on the Internet. Both large quantity and various forms of online reviews bring difficulties for potential consumers to summary all the heterogenous reviews for reference. This paper proposes a new ranking method through online reviews based on different aspects of the alternative products, which combines both objective and subjective sentiment values. Firstly, weights of these aspects are determined with LDA topic model to calculate the objective sentiment value of the product. During this process, the realistic meaning of each aspect is also summarized. Then, consumers' personalized preferences are taken into consideration while calculating total scores of alternative products. Meanwhile, comparative superiority between every two products also contributes to their final scores. Therefore, a directed graph model is constructed and the final score of each product is computed by improved PageRank algorithm. Finally, a case study is given to illustrate the feasibility and effectiveness of the proposed method. The result demonstrates that while considering only objective sentiment values of the product, the ranking result obtained by our proposed method has a strong correlation with the actual sales orders. On the other hand, if consumers express subjective preferences towards a certain aspect, the final ranking is also consistent with the actual performance of alternative products. It provides a new research idea for online customer review mining and personalized recommendation.

Keywords: Online review mining, LDA topic model, improved PageRank algorithm, personalized recommendation

1. Introduction

With the rapid development of web 2.0, the role of Internet users has changed greatly from the previous information recipients to the information producers. Especially, the development of online shopping platforms has attracted more and more consumers to share their shopping experience and product reviews on the Internet. A survey found that 63% of online shoppers consistently search and read reviews before making a purchasing decision. Among these, 64% of the online shoppers spend 10 minutes reading reviews, whereas 33% spend 30 minutes or more (Marsden 2010). The forum data also attracted the attention of scholars (Wang and Tang 2016). Due to the large number of online reviews, how to quickly and effectively view the valid information and refer online reviews to assist in making more appropriate purchasing decisions has become a hot research issue (Guo et al. 2017).

Although many scholars have studied content-based recommendation, collaborative filtering (Yang et al. 2017) and knowledge-based recommendation methods (Park et al. 2015) for user recommendation, few user recommendations are based on online reviews. At the very beginning, scholars concentrate mostly on numerical scores to sort products. However, with the expansion of Internet shopping platforms, online textual reviews are gradually noticed. More and more scholars pay attention to the combination of numerical and textual information. But most studies take the whole review as a research object to judge the sentiment intendency of the product and do not consider that sentiments of different aspects concerning a certain product might be different. Moreover, data preprocessing method is quite simple in existing researches. Natural language processing and text mining methods are not employed for data processing, especially for textual data. Machine learning methods are rare to be found in most existing studies for products evaluation. What's more, weights of different aspects are mostly determined subjectively or by weighted average method directly, which is lacking in objectivity and accuracy.

In this paper, we combine textual reviews and numerical scores to extract customers' preferences concerning different aspects of alternative products by text mining technology. The reviews are firstly segmented into words of different aspects based on an external dictionary. The directed graph model more intuitively shows the comparison network and the relative competitive advantage among items. To define the weight coefficients of all aspects, a method combining objective weight and subjective weight is put forward. The LDA (Latent Dirichlet Allocation) topic model is used to determine the objective weights of aspects about the product. Therefore, the customer's personalized preferences are reflected and the massive online reviews can be maximally utilized. In addition, mobile phone, digital camera, lamps and other low-price products are selected in the previous studies. This paper takes automobile as the research object. Considering with the improvement of living standards, the demand for automobile is increasing. What's more, the price of automobile is expensive. So, making accurate personalized recommendation is of great practical significance for consumers.

The rest of this paper is organized as follows. In Section 2, we review related works about online review data selection and online products evaluation methods. In Section 3, it is proposed a method to mine and integrate textual reviews and numerical information. The directed graph model is used for information fusion and improved PageRank method is deployed to calculate the node value. In Section 4, a case study is given to evaluate the proposed ranking method. Section 5 presents conclusions as well as future work.

2. Related Work

There are different types of online reviews, such as textual reviews, numerical ranking, comparison information, etc. To rank products through online product reviews, it is firstly necessary to determine what kind of information should be chosen. What's more, methods to integrate personalized preferences into the evaluation model are also various. Thus, the related work can be further classified into online review data selection and online products evaluation methods.

2.1 Online Review Data Selection

Numerical score is considered as a succinct summary of consumer's opinion. It is widely employed because of its simplicity and has attracted lots of attention because of its convenience for data processing. Eight million product reviews and 1.5 million business reviews were selected to evaluate quality and rank different products. They found scoring average indicators have a better performance (Mcglohon et al. 2010). However, due to the fuzziness of numerical score information, the overall score of a product can't well reflect the user's sentiment expression of the product's local characteristics. That is, the products with same score can't be distinguished from different aspects (Ghose and Ipeirotis 2011). Review texts is a kind of valuable user-generated content which can better reflect the user's emotion and provide more information to their readers. However, it has been neglected all the time. A new text sentiment analysis method called eSAP was proposed for decision support systems and carried out experiments in seven different fields of data sets to verify the effectiveness of the method (Khan et al. 2016). To solve the problem of unbalanced sentiment classification in Chinese product reviews, this paper proposed a new method based on topic sentence to improve the accuracy of classification (Tian et al. 2016). The study incorporated contextual features derived from replies and used a multi-view ensemble learning method to identify mentions of product defects from social media. A case study in the automotive industry demonstrated the utilities of both novelties in their method (Liu et al. 2017). However, textual reviews generally contain specific opinions based on a certain aspect or some aspects of the product, which means they can't cover comprehensive information. Recently, models integrating numerical ratings and review texts as training sources have attracted a lot of attention (Hu et al. 2014, Najmi et al. 2015). Consumers' opinions on some important aspects greatly influence their overall opinions on the product. Researches dedicate to the topic of aspect ranking, which aims to automatically identify important aspects of a product from consumer reviews (Yu et al. 2011). Reviews including the behavioral information and the textual review contents from B2C websites were extracted to explore reviewers' opinions. But they didn't take considers' preferences into consideration (Li et al. 2016).

Most of the data used in product ranking are numerical information. Although natural language processing technology has been developing in recent years, few researches employed machine learning methods to do text mining, such as topic model and hidden Markov model. On the other hand, most of existing researches focus on the whole review but ignore reviewers' sentiments towards different aspects of the product. Based on above analysis, this paper combines heterogeneous information to mine consumer's sentiments of a product from different aspects, and then use machine learning methods to determine objective and subjective weights of aspects respectively.

2.2 Online Products Evaluation Methods

Whether the experiment uses single data or heterogeneous data, the final evaluation system needs to be constructed. The method to rank multiple products through heterogeneous information was proposed, but the weights are determined only according to online reviews, which does not consider consumers' subjective preferences (Yang et al. 2016). The review's credibility and posting date are considered to measure the weight of each review. However, every aspect is given the same weight so that it cannot reflect the importance of different aspects for consumers (Zhang et al. 2011). Liu et al. (2017) proposed a method to rank products with sentiment analysis technique and intuitionistic fuzzy set theory. The weight of each aspect is determined according to the frequency of sentiment words appearing in Hownet dictionary. Although reflecting the personalized preference, the weights are determined subjectively (Liu et al. 2017). There are different types of online reviews with distinct features, such as numeric ratings and textual descriptions. Such heterogeneous information leads to more complexities for customers to make purchase decisions. To solve the problems of unstructured data and over-loaded information, researchers have applied machine learning techniques to build automated product defect identification models, which can help manufacturers reduce labor costs significantly (Zhang et al. 2016) (Abrahams et al. 2012). Topic model was used to mine product reviews and obtain product attributes that consumers care about. Then, products ranking is done through TOPSIS method. But this method (needs reviews classification in advance) must classify the reviews by sentiment in advance (Chen et al. 2015). A feature network-driven quadrant mapping is proposed that captures and incorporates opinions from customer reviews. Their focus is on construction of a feature network, which is based on co-occurrence and sematic similarities, and a quadrant display showing the opinions polarity of feature groups (Su et al. 2017).

In this paper, a directed graph model is proposed to integrate such rich and heterogeneous information. The modified PageRank method is used to calculate the value of each node and to rank the different products. To evaluate the relative advantages of the product the directed graph model was used, but there was no analysis of the network node weight in the network (Zhang et al. 2013). In this paper, the network node is the comprehensive sentiment value of the product, and the direction of the edge represents the relative advantage between the products. The weight of the side represents the relative advantage.

3. Research Methods

To make the ranking result of different products more accurate, the user-generated contents should be made full use of. This paper selects online textual reviews and numerical information to do analysis. The directed graph model is used for information fusion and the improved PageRank method is deployed to calculate the node values, which refer to the final scores of different products. The research framework is shown in Figure 1.

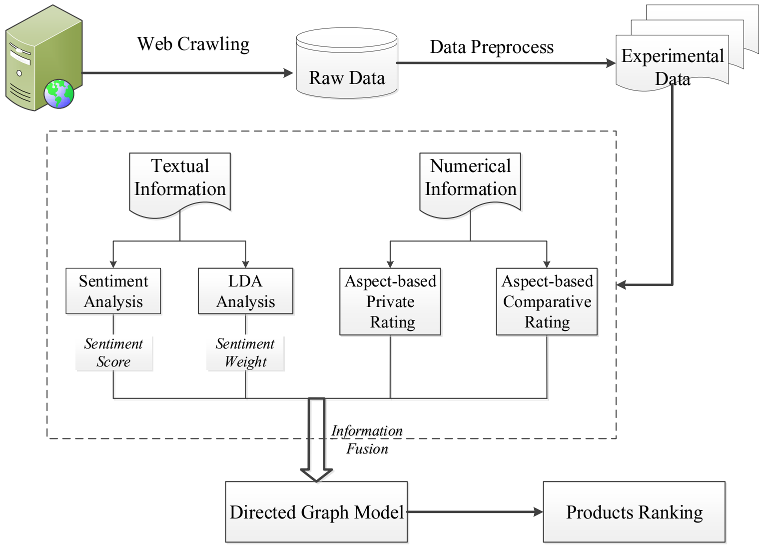

Figure 1 Research framework

3.1 Textual Reviews Processing Method

In this section, we firstly formulate the problem of aspect-based products ranking. Then, we describe the method to calculate sentiment intensity and weights determination of aspects.

3.1.1 Sentiment Intensity Calculation Method

Online aspect-based reviews can be crawled from the third-party website. Let $P=\{P_1,P_2,\ldots,P_i,...,P_I\}$ denotes the set of alternative products, where P_i is the ith alternative product, i=1, 2, …, I. Let $F=\{f_1,f_2,\ldots,f_k,...,f_K\}$, where f_k denotes the kth aspect, k =1, 2, …, K. Let $C=(c_1,c_2,...,c_k,...,c_K)$, where c_k is the weight of f_k, $c_k\geq 0$ and $\sum_{k=1}^{K}c_k=1$. Let $Q=\{q_1,q_2,\ldots,q_i,...,q_I\}$, where q_i denotes the number of online reviews about P_i. Let $W_{ij}=\{w_{ij}^{1},w_{ij}^{2},...,w_{ij}^{k},...,w_{ij}^{K}\}$, where w_{ij}^{k} is the set of words segmented from the jth online review concerning the kth aspect of P_i, $i=1,2,...,I$, $j=1,2,...,q_i$, $k=1,2,...,K$. If there is no sentence concerning the kth aspect in the jth review of P_i, $w_{ij}^{k}=\varnothing$. In this paper, the sentiment word lexicon established by the information retrieval team of Dalian University of Technology is used as the basis of text analysis (Xu et al. 2008). The sentiment word lexicon is denoted by $SW=\{sw_1,\ sw_2,\ \ldots,\ sw_q,\ \ldots,\ sw_Q\}$, where sw_q denotes the qth sentiment word, $q=1,2,...,Q$. The corresponding sentiment intensity is denoted by $SV=\{sv_1,\ sv_2,\ \ldots,\ sv_q,\ \ldots,\ sv_Q\}$, where sv_q denotes the sentiment intensity of sw_q. The sentiment polarity is denoted as $SP=\{sp_1,\ sp_2,\ \ldots,\ sp_q,\ \ldots,\ sp_Q\}$, where sp_q denotes the sentiment polarity of sw_q. If word sw_q is positive, then sp_q=1 ; if it is negative, $sp_q = -1$; otherwise, $sp_q = 0$. Let $S_{ij} = \{S_{ij}^{\ 1}, S_{ij}^{\ 2}, ..., S_{ij}^{\ k}, ..., S_{ij}^{\ K}\}$, where $S_{ij}^{\ k}$ denotes the sentiment intensity of the *k*th aspect in the *j*th review of P_i .The process to calculate $S_{ij}^{\ k}$ is given as below.

Step1: Let $w_{ij}^{\ kn}$ be the *n*th segmented word of $w_{ij}^{\ k}$. If $w_{ij}^{\ kn}$ appears in SW and its corresponding word in SW is sw_q , then the sentiment intensity sv_q and sentiment polarity sp_q can be obtained respectively.

Step2: Determine the modified sentiment intensity sv_q'. If there is no negation word around $w_{ij}^{\ kn}$, the sentiment intensity is calculated by $sv_q' - sv_q = 0$; otherwise, it is modified by $sv_q' + sv_q = 0$.

Step3: Determine the sentiment polarity sp_q . If $sp_q = 1$, which means the review is positive, then $S_{ij}^{\ k} = S_{ij}^{\ k} + sv_q'$; If $sp_q = -1$, $S_{ij}^{\ k} = S_{ij}^{\ k} - sv_q'$;otherwise, $S_{ij}^{\ k}$ will not change.

In this paper, we assume that weights of reviews from different customers for the *k*th aspect of P_i are equal. Therefore, the sentiment intensity of the *k*th aspect about alternative product P_i can be calculated as:

$$S_i^{\ k} = \frac{1}{q_i}\sum_{j=1}^{q_i} S_{ij}^{\ k} \tag{1}$$

3.1.2 Aspect-Based Weight Determination Method

To obtain the overall sentiment intensity of P_i , we need to calculate the weight of each aspect. LDA is a typical topic model and is mainly used to deal with discrete data. It has a wide range of applications in information retrieval, natural language processing and other fields.

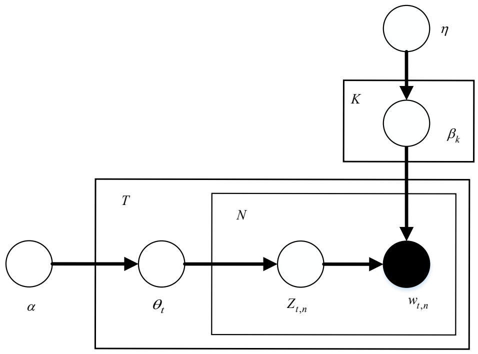

Figure 2 LDA model

The relationship of LDA variables is described in Figure 2, where hollow circles represent hidden variables and solid circles represent observable ones. Only $w_{t,n}$, the word frequency of word *n* in document *t*, is observable. It depends on $Z_{t,n}$, the topic of word *n* in document *t* and β_k , the word distribution of topic *k*. Meanwhile, $Z_{t,n}$ depends on θ_t , which refers to the topic distribution in document *t*. θ_t and β_k respectively are defined by Dirichlet Allocation with parameters α and η . Accordingly, the probability distribution of LDA is given as follows

$$p(W, z, \beta, \theta \mid \alpha, \eta) = \prod_{t=1}^{T} p(\theta_t \mid \alpha) \prod_{i=1}^{K} p(\beta_k \mid \eta)$$
$$\left(\prod_{n=1}^{N} P(w_{t,n} \mid z_{t,n}, \beta_k) P(z_{t,n} \mid \theta_t)\right) \tag{2}$$

where $p(\theta_t \mid \alpha)=\frac{\Gamma(\sum_k \alpha_k)}{\prod_k \Gamma(\alpha_k)}\prod_k \theta_{t,k}^{\alpha_k-1}$ and $p(\beta_k \mid \eta)$ usually obey K- dimensional and N- dimensional Dirichlet Allocation with parameters α and η. The process to calculate weight of each aspect is given as below.

Step1: Parameter K, the number of topics, can be determined with prior knowledge. The topic-word distribution can be obtained from β_k, based on which the word cloud for each topic can be generated. With textual visualization techniques, we can display each topic more intuitively and summarize realistic meaning of each topic with consideration of realistic background;

Step2 : From θ_t, we can know the document-topic distribution. Then $\alpha\theta_t$, the distribution of each topic in document *t*, can be determined. The weight vector $C=(c_1,c_2,...,c_k,...,c_K)$ can be calculated as

$$C=\frac{1}{q_i}\sum_{t=1}^{q_i}\alpha\theta_t \quad (3)$$

Step3: Let TS_i be the total sentiment value of P_i. It is determined by both objective and subjective sentiment values. The objective sentiment value is calculated as

$$S_i=\sum_{k=1}^{k}c_k S_i^k \quad (4)$$

The subjective sentiment value is related to consumers' individual preferences. Let γ_k be the weight of the *k*th aspect given by a potential consumer. His/her subjective sentiment value is

$$S_i'=\sum_{k=1}^{k}\gamma_k S_i^k \quad (5)$$

Therefore, the total sentiment value TS_i can be computed as

$$TS_i=\delta S_i+(1\text{-}\delta)S_i' = \delta\sum_{k=1}^{K}c_k S_i^k+(1\text{-}\delta)\sum_{k=1}^{K}\gamma_k S_i^k \quad (6)$$

where δ denotes the weight of the consumer's objective sentiment value. δ and $1-\delta$ respectively reflect to what extent the customer depends on objective reviews and subjective preference.

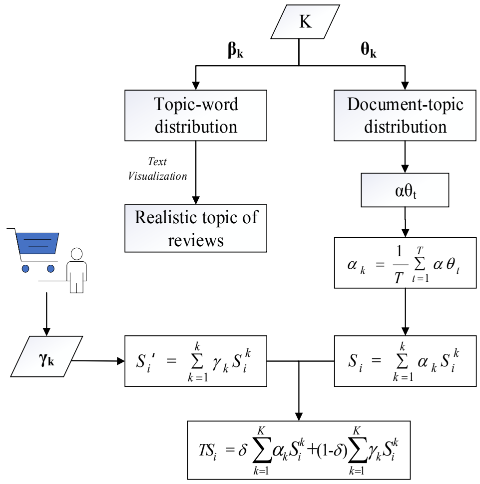

Figure 3 Flowchart of calculating comprehensive sentiment value

3.2 Numerical Information Processing Method

Numerical information is adopted by many scholars because of its simplicity and convenience for processing. Numerical score represents the satisfaction degree of the consumer on a certain product. This paper will analyze it from two aspects: numerical rating based on itself and the numerical rating based on the comparison.

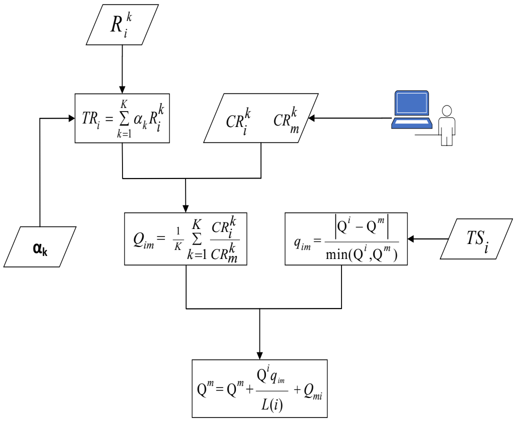

Figure 4 Flowchart of calculating products' importance

a. The Numerical Rating Based on Itself

Let R_{ij}^k be the numerical rating of alternative product P_i concerning the kth aspect in the jth online review. Since difference customers are of same importance, the numerical rating of P_i concerning the kth aspect can be computed as

$$R_i^k = \frac{1}{q_i} \sum_{j=1}^{q_i} R_{ij}^k \tag{7}$$

Therefore, the overall numerical rating of P_i can be computed as

$$TR_i = \sum_{k=1}^{K} c_k R_i^k \tag{8}$$

b. The Numerical Rating Based on the Comparison

Nowadays, many third-party platforms allow customers to compare homogeneous products side by side. Let CR_i^k be objective comparative rating concerning the kth aspect of alternative product P_i. It can be crawled from online platforms directly. So, the relative comparative superiority of product P_i over P_m can be computed as:

$$Q_{im} = \frac{1}{K} \sum_{k=1}^{K} \frac{CR_i^k}{CR_m^k} \tag{9}$$

3.3 Information Fusion Network Construction Method

Construct a directed graph model $G(V,E,Q^V,Q^E)$. Node V denotes the product. Edge E denotes the directed connection between products, and Q^v denotes weight of node V, which refers to the sum of sentiment value TS_v and the numerical score TR_v. q_{im} denotes the weight of edge E. that is, the relative comparative superiority, which can be computed as:

$$q_{im} = \frac{|Q^i - Q^m|}{\min(Q^i, Q^m)} \tag{10}$$

If $q_{im} > 1$, the edge is directed from P_m to P_i. If $q_{im} < 1$, the direction is reversed. If $q_{im} = 1$, there is no connection between these two nodes.

There are many ways to calculate the importance of network nodes. The classical PageRank algorithm is based on the idea that pages from high-quality web pages must be of high quality (Page et al. 1999). Similarly, one product with comparative advantage from another high evaluation product will surely be of high evaluation. We use improved PageRank algorithm to calculate importance of each node. Assume that there are n nodes in a directed graph, n=1,2, 3, ..., N, the number of output edges of node n is denoted as $L(n)$. Therefore, the importance of node m can be computed as:

$$TQ^m = Q^m + \frac{\sum_{i \in D_m} Q^i q_{im}}{L(i)} + Q_{mi} \tag{11}$$

where D_m is the set of nodes that have an input edge to node m. For example, in Figure5, $D_1 = \{2,3\}$. The importance of node 1 can be computed as

$$TQ^1 = Q^1 + \frac{Q^2 q_{21}}{L(2)} + \frac{Q^3 q_{31}}{L(3)} + Q_{11} + Q_{12} + Q_{13} \tag{12}$$

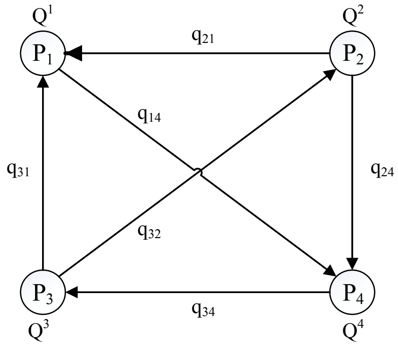

Figure 5 Directed graph example

4. Experimental Study

With the rapid development of China's economy, people's living standards have been improved significantly. Automobile has become a necessity in people's life. Because of its durability and high value, consumers are discreet while selecting a satisfied one among many automobiles with different brands. To provide an information source while they making purchase decisions, internal and external WOM(word-of-mouth) websites have been developed for consumers to share their reviews about products they have bought and have a significant impact on retailer's sales, of which the external WOMs, also known as third-party platforms, play a more important role in consumers' information search process than those internal ones while customers intend to buy high-involvement products such as digital cameras(Gu et al., 2012). Therefore, in this paper, we crawled online reviews from two third party platforms, Autohome (http://www.autohome.com.cn/) and PCauto (http://www.pcauto.com.cn/), to do analysis. Each review published by a customer includes numerical scores and textual information concerning each aspect of the product.

Consider a consumer who wants to buy a compact SUV automobile with price in the range of 100,000 and 300,000 yuan. There are four alternative brands he is interested in and eight features he concerns about, which are listed as follows.

P_1: Mazda CX-5 2015, 2.5L (Deluxe Edition automatic four-wheel drive);

P_2 : Trumpchi GS4 2016 (Deluxe Edition 235T G-DCT);

P_3 : Roewe RX5 2016 (3.0T automatic four-wheel drive version of the Internet);

P_4: Honda CR-V 2015, 2.4L (Deluxe Edition two-wheel drive).

f_1 : appearance

f_2 : car interior

f_3 : space

f_4 : configuration

f_5 : power

f_6 : cross-country

f_7 : oil consumption

f_8 : comfort

The web crawler technique is used to obtain online reviews from Autohome and PCauto about automobiles of the four brands. The detailed information about reviews crawled is shown in Table 1.

Stanford Parser is applied to preprocess the textual data. An external dictionary is added for word segmentation to avoid ambiguities because of separation of field words. After word segmentation, 307 comprehensive reviews of P_1, 282 of P_2 , 179 of P_3 and 316 of P_4 are decomposed into aspect-based reviews. The sentiment intensity and polarity of each word can be found in SW and SV, with which the sentiment intensity of each review about each aspect concerning product P_i can be obtained. For example, $S_{12}{}^1 = 0$, $S_{12}{}^2 = 0$,…,

$S_{12}^{8}=-3$. The sentiment intensities of the first eight reviews about Mazda CX-5 concerning eight aspects are listed in Table 2. Based on Eq. (1), the sentiment intensities S_i^k of product P_i towards the kth aspect can be calculated, $i=1,2,3,4$, $k=1,2,\ldots,8$. The detailed result is shown in Table 3.

Table 1 Data description

	Total number of aspect-based reviews	Total number of comprehensive reviews	Number of reviews about f_i							
			f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8
P_1	2338	307	297	280	287	298	301	294	291	290
P_2	2140	282	264	263	260	273	274	269	276	261
P_3	1384	179	175	171	174	172	172	174	176	170
P_4	2381	316	297	296	302	294	301	293	300	298

Table 2 Aspect-based sentiment values of Mazda CX-5 from different customers

	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8
$j=1$	5	0	3	3	-5	3	0	0
$j=2$	0	0	5	0	5	5	-7	-3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$j=8$	-5	3	5	0	11	0	0	5

Table 3 Aspect-based sentiment values of different products

	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8
P_1	4.77	3.37	2.29	2.84	3.78	7.00	3.16	2.51
P_2	4.95	2.91	2.97	2.22	3.13	5.18	2.57	1.94
P_3	6.80	5.88	4.33	3.97	3.70	4.99	2.92	2.68
P_4	4.57	2.48	3.00	2.63	3.90	4.02	3.55	2.68

According to the method proposed in Sec 3.1.2, the preprocessed textual reviews are used to determine weights of aspects based on LDA topic model. For example, with online reviews of Mazda CX-5, given the parameter K=8, we can obtain the topic distribution in each document (each online text review namely), which is denoted as a 307 × 8 document-topic probability matrix θ_t.

$$\theta_t = \begin{array}{c} \\ Doc\_1 \\ Doc\_2 \\ \cdot \\ \cdot \\ \cdot \\ Doc\_307 \end{array} \begin{array}{c} \text{Topic\_1 Topic\_2} \quad \ldots \quad \text{Topic\_8} \\ \begin{bmatrix} 0 & 0 & 0.345 & \ldots & 0.552 & 0 & 0 \\ 0 & 0 & 0.878 & \ldots & 0 & 0 & 0 \\ \cdot & \cdot & & & & \cdot & \cdot \\ \cdot & \cdot & & & & \cdot & \cdot \\ \cdot & \cdot & & & & \cdot & \cdot \\ 0 & 0.562 & 0 & \ldots & 0 & 0 & 0.02 \end{bmatrix} \end{array}$$

Meanwhile, with LDA topic model, the distribution of keywords under each topic can be also determined and then value of word frequency f can be set. It is demonstrated a negative correlation between the number of keywords and word frequency.

Let f=40 and then 163 keywords are kept, from which a 8×163 topic-keyword probability matrix β_k can be generated, i.e.

$$\beta_k = \begin{array}{c} \\ Topic\_1 \\ Topic\_2 \\ \vdots \\ Topic\_8 \end{array} \begin{array}{c} \begin{array}{cccc} Word\_1 & Word\_2 & \cdots & Word\_163 \end{array} \\ \begin{bmatrix} 0.0034 & 0.0146 & \ldots & 0.0023 \\ 0.0043 & 0.0092 & \ldots & 0.0022 \\ \vdots & \vdots & & \vdots \\ 0.0047 & 0.0104 & \ldots & 0.0026 \end{bmatrix} \end{array}$$

With β_k, the word cloud can be generated to visualize keywords of each topic, from which we can get the realistic meaning of each one. The greater the probability value of a word in a topic is, the higher the correlation between this word and this topic is. As a result, the size of this word will be larger in the word cloud. As is seen in Figure 6, gearbox, acceleration and motive relate much closer to this topic than other words. Therefore, combined with practical experience, we can draw a conclusion the realistic meaning of this topic is power.

(a) Chinese word cloud **(b)** English word cloud

Figure 6 Word cloud about power aspect

The objective sentiment values of the four products based on online reviews can be obtained. The results are shown in Table 4.

Table 4 Topic-product correspondence and sentiment value analysis

		Topic1	Topic2	Topic3	Topic4	Topic5	Topic6	Topic7	Topic8	
		f_1	f_2	f_7	f_6	f_3	f_5	f_8	f_4	S_1
P_1	α_k	0.012	0.305	0.119	0.317	0.021	0.017	0.074	0.135	3.3311
	$S_1^{\ k}$	4.772	2.071	3.160	4.997	2.292	3.788	2.508	2.844	
		f_3	f_6	f_4	f_8	f_1	f_7	f_2	f_5	S_2
P_2	α_k	0.088	0.041	0.031	0.18	0.048	0.246	0.338	0.028	4.1359
	$S_2^{\ k}$	2.968	5.184	2.794	3.943	4.950	3.574	4.915	3.128	
		f_4	f_7	f_5	f_8	f_6	f_2	f_3	f_1	S_3
P_3	α_k	0.023	0.057	0.009	0.065	0.047	0.461	0.281	0.046	4.9414
	$S_3^{\ k}$	3.972	2.916	3.704	2.682	4.994	5.883	4.330	6.804	
		f_1	f_6	f_2	f_4	f_3	f_7	f_5	f_8	S_4
P_4	α_k	0.025	0.102	0.326	0.241	0.017	0.164	0.044	0.081	3.8804
	$S_4^{\ k}$	4.557	4.016	4.481	3.627	2.997	3.551	3.899	2.684	

Based on the analysis in Section 3.3, we can construct a directed graph model which is shown in Figure 8. In this figure, the value of each node represents the objective sentiment value of each alternative product. The direction between two nodes represents that one product has a competitive advantage over the other and the number beside the line denotes the degree of relative comparative superiority. For example, the line directed from node P_2 to node P_1

indicates that P_1 has a competitive advantage over P_2. As is shown in Figure 7, P_1 has competitive advantage over P_2.

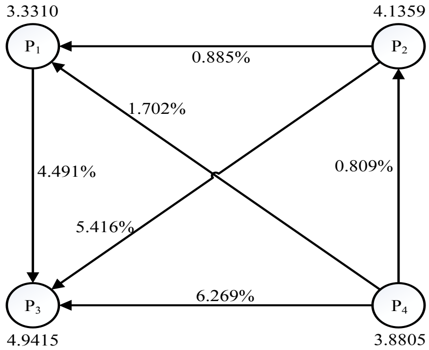

Figure 7 Directed graph model of four products

Based on the improved PageRank algorithm, the final scores for the four products without consumers' subjective preference can be obtained. By Eq. (12), the final scores of the four products are 3.3713, 4.1464, 5.2842 and 3.8805. Therefore, the final ranking result is $P_3 > P_2 > P_4 > P_1$. Competition car ranking information provided by Pacific automotive network shows the actual ranking result is $P_3 > P_1 > P_2 > P_4$, that is, Roewe RX5 and Trumpchi GS4 are the first and fourth respectively. It means the result calculated by the proposed method in this paper is relatively consistent with the actual performance of each product. However, the calculated result is not totally consistent with the actual ranking. It is because the textual reviews are considered in this paper which were not used to get the result provided by Pcauto Therefore, the result still verifies the effectiveness of the proposed method without considering the subjective preferences of customers.

To further verify the accuracy of the method, three more kinds of automobiles were added into the alternative ones, which are RAV4, 2016, 2.0L (CVT four-wheel drive new Edition), X-Trail 2016, 2.0L (CVT intellectual Edition 4WD) and Kuga 2015(2.0 GTDi, Deluve Edition four-wheel drive). The reviews of these three products are also crawled from Autohome. According to the method proposed in section 3, the directed graph network model is shown in Figure 8.

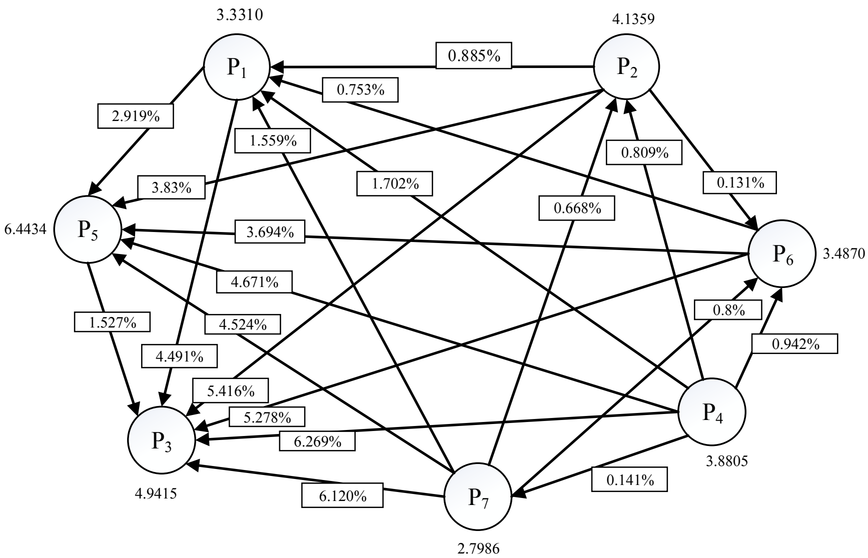

Figure 8 Directed graph model of seven products

Spearman correlation coefficient is used to verify the validity of the ranking method proposed in this paper, whose equation is

$$\rho(\vec{R}_{x_i} - \vec{R}_{y_i}) = 1 - \frac{6\sum_{i=1}^{n}(R_{x_i} - R_{y_i})^2}{n(n^2-1)} \quad (13)$$

where $\vec{R}_{x_i}$ denotes the ranking result obtained by the proposed method, while $\vec{R}_{y_i}$ represents the actual sales ranking in April, 2017. Values of $\vec{R}_{x_i}$ and $\vec{R}_{y_i}$ are shown in Table 5. The result is $\rho(\vec{R}_{x_i} - \vec{R}_{y_i})$ =0.607>0.5 , which demonstrates that the calculated result has a strong correlation with the actual sales ranking. What's more, the newly obtained order of the seven products is $P_5 > P_3 > P_2 > P_4 > P_6 > P_1 > P_7$, from which we can see the ranking result of P_1, P_2, P_3 and P_4 does not change although three more products are added into alternative list.

As far as we know, different customers have specific individual needs and they may concern about different aspects of a product. As for automobiles, for example, people with lower income may pay more attention to cost performance, while others who like cross-country driving is more concerned about control power and off-road performance of the automobile. Therefore, consumers' subjective preferences should be taken into consideration while ranking alternative products. Based on Eq. (6), if the consumer concerns more about aspect f_k, its weight γ_k will be of a higher value. If he/she relies more on online reviews to make a purchase decision, δ will be higher. To analyze the relationship between ranking result and these two coefficients, we set different values of δ, γ_k and the results are shown in Table 6.

Table 5 Experimental results about seven products

Product	TQ^m	Ranking result	Sales list in April
P1(Mazda CX-5)	3.369	6	6
P_2(Trumpchi GS4)	4.145	3	1
P_3(Roewe RX5)	5.307	2	2
P_4(Honda CR-V)	3.881	4	3
P_5(RAV4)	6.630	1	5
P6(X-Trail)	3.499	5	4
P7(Kuga)	2.800	7	7

Table 6 Experimental results based on different δ and γ_k

	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7	γ_8	Ranking Result
δ=0.2	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	$P_5 > P_3 > P_2 > P_4 > P_6 > P_1 > P_7$
	0.3	0.1	0.1	0.1	0.1	0.1	0.1	0.1	$P_3 > P_5 > P_2 > P_1 > P_6 > P_4 > P_7$
	0.1	0.1	0.1	0.1	0.3	0.1	0.1	0.1	$P_3 > P_7 > P_1 > P_5 > P_2 > P_4 > P_6$
δ=0.5	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	$P_5 > P_3 > P_2 > P_4 > P_6 > P_1 > P_7$
	0.3	0.1	0.1	0.1	0.1	0.1	0.1	0.1	$P_5 > P_3 > P_2 > P_6 > P_1 > P_4 > P_7$
	0.1	0.1	0.1	0.1	0.3	0.1	0.1	0.1	$P_3 > P_5 > P_4 > P_7 > P_2 > P_1 > P_6$
δ=0.8	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125	$P_5 > P_3 > P_2 > P_4 > P_6 > P_1 > P_7$
	0.3	0.1	0.1	0.1	0.1	0.1	0.1	0.1	$P_5 > P_3 > P_2 > P_4 > P_6 > P_1 > P_7$
	0.1	0.1	0.1	0.1	0.3	0.1	0.1	0.1	$P_5 > P_3 > P_2 > P_4 > P_1 > P_7 > P_6$

where γ_k is the subjective preference coefficient of different aspects f_k.

From Table 6 we can find that the final ranking result will not change if weights of different aspects are the same even if the value of δ is different. However, if consumers express their subjective preferences towards a certain aspect f_k, the result will be different. For example, if f_1=0.3, $f_2 = f_3 = \ldots = f_8$=0.1, that is, the customer concerns more about aspect f_1 (i.e., appearance) and other aspects are of no differences for him, the ranking result is $P_3 > P_5 > P_2 > P_1 > P_6 > P_4 > P_7$. The top two products are P_3 (Roewe RX5) and P_5 (RAV4). From the comparative result provided by PCauto, we know that Roewe RX5 and RAV4 behave 6% and 3% better than those same level products in aspect f_1 respectively, while the last product P7 (Kuga) behaves 6% worse than others in aspect f_1, from which we can see the ranking result obtained by our method is consistent with the actual ranking result. To further demonstrate the effectiveness of our proposed method, Spearman coefficients are used to calculate the ranking results obtained by our method and the online ranking results about different aspects. We can get the comprehensive ranking aspect-based provide by PCauto, in which the rank of appearance and power are $P_3 > P_5 > P_2 > P_1 > P_6 > P_4 > P_7$, $P_3 > P_7 > P_1 > P_5 > P_4 > P_2 > P_6$ respectively. While setting different δ are calculated concerning aspects of appearance and power, which is shown in Figure 9.

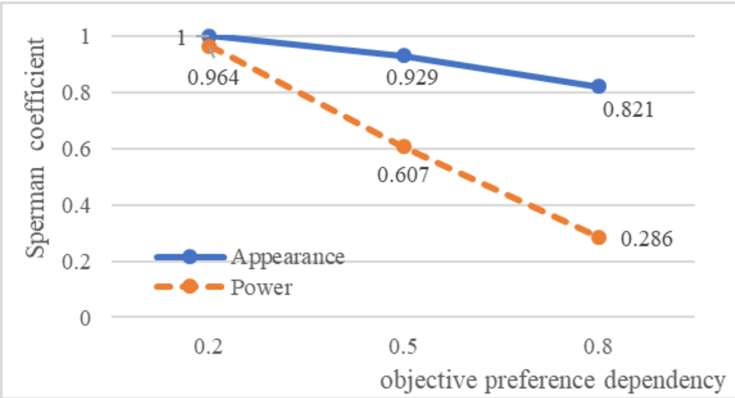

Figure 9 Objective preference dependency δ and Spearman coefficient ρ

If spearman correlation coefficient $\rho \geq 0.5$, it is shown that the experimental results have a strong correlation with the actual sales ranking. From Figure10, we can know that if consumers see objective reviews and subjective preferences as the same or rely more on objective preference, all Spearman coefficients ρ are bigger than 0.5, which testifies the effectiveness of our method. On the contrary, if we can know when potential consumers pay more attention on subjective preference, that is, the value of δ is small, the experimental ranking result will have strong correlation with aspect-based ranking. With the increase of objective preference dependency, the Spearman coefficient turns smaller and the correlation is reduced. It also indicates that the effect of objective preference dependency δ is more powerful than subjective preference weight γ_k.

5. Conclusions

In this paper, a new method of mining aspect-based online reviews is proposed. Firstly, different from other studies studying single kind of online data, this paper forms a directed graph model to fuse heterogeneous online reviews. Secondly, text mining techniques including LDA topic model and textual visualization are adapted to obtain the value of nodes in directed graph. The ranking result obtained from the improved PageRank algorithm considers not only the information mining from the massive online reviews but also consumer's preference. So, the ranking is more targeted to make personalized recommendation.

The experimental results demonstrate the feasibility and effectiveness of the proposed method. The Spearman coefficient is used to demonstrate that our experimental results have a strong correlation with actual sales ranking. It should be noted that the conventional evaluation measures (e.g., precision, recall, and F-measure) were not used in this study because the main intent of the proposed approach is not to provide customers with the precise answer. Instead, we would provide the customers with the most suitable products.

Although the current study focuses primarily on existing products, this idea can be extended to forecast the sales of new products. What's more, the posted time of online reviews should be taken into consideration. Newer online reviews should be assigned greater importance degrees. Furthermore, we believe that developing decision support systems for dealers is of great practical significance.

Acknowledgements

The authors would like to thank the anonymous referees for their insightful and constructive comments and suggestions that have led to an improved version of this paper. This research was supported in part by the National Natural Science Foundation of China [Grant Numbers 71771034, 71421001] and the Scientific and Technological Innovation Foundat ion of Dalian (2018J11CY009). This paper is a significantly extended and revised version of the conference paper presented at KSS-2017 (Guo, Du & Kou, 2017).

References

[1] Abrahams, A. S., Jiao, J., Wang, G. A., & Fan, W. (2012). Vehicle defect discovery from social media. Decision Support Systems, 54(1), 87-97.

[2] Chen, K., Kou, G., Shang, J., & Chen, Y. (2015). Visualizing market structure through online product reviews. Electronic Commerce Research & Applications, 14(1), 58-74.

[3] Ghose, A., & Ipeirotis, P. G. (2011). Estimat-ing the helpfulness and economic impact of product reviews: mining text and reviewer characteristics. IEEE Transactions on Knowledge & Data Engineering, 23(10), 1498-1512.

[4] Gu, B., Park, J., & Konana, P. (2012). Re-search note: the impact of external word-of-mouth sources on retailer sales of high-involvement products. Information Systems Research, 23(1), 182-196.

[5] Guo, C., Du, Z., & Kou, X. (2017). Mining Online Customer Reviews for Products Aspect-Based Ranking. International Symposium on Knowledge and Systems Sciences (pp.150-161). Springer, Singapore.

[6] Hu, N., Koh, N. S., & Reddy, S. K. (2014).

Ratings lead you to the product, reviews help you clinch it? the mediating role of online review sentiments on product sales. Decision Support Systems, 57, 42-53.

[7] Khan, F. H., Qamar, U., & Bashir, S. (2016). Esap: a decision support framework for enhanced sentiment analysis and polarity classification. Information Sciences, 367-368, 862-873.

[8] Li, L., Yuan, H., Qian, Y., & Shao, P. (2016). Towards exploring when and what people reviewed for their online shopping experiences. Journal of Systems Science & Systems Engineering, 1-27.

[9] Liu, Y., Bi, J. W., & Fan, Z. P. (2017). Ranking products through online reviews: a method based on sentiment analysis technique and intuitionistic fuzzy set theory. Information Fusion, 36, 149-161.

[10] Liu, Y., Jiang, C., & Zhao, H. (2017). Using contextual features and multi-view ensemble learning in product defect identification from online discussion forums. Decision Support Systems, 105, 1-12.

[11] Marsden P. (2010). PowerReviews 2010 Social Shopping Study | Top Line Results. http://digitalintelligencetoday.com/2010-social-shopping-study-top-line-results/.Cited May 14, 2010.

[12] Mcglohon, M., Glance, N. S., & Reiter, Z. (2010). Star Quality: Aggregating Reviews to Rank Products and Merchants. Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media, 114-121.

[13] Najmi, E., Hashmi, K., Malik, Z., Rezgui, A., & Khan, H. U. (2015). Capra: a comprehensive approach to product ranking using customer reviews. Computing, 97(8), 843-867.

[14] Page, L. , Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank citation ranking: bringing order to the web. Stanford InfoLab Technical Report, 1-17.

[15] Park, Y., Park, S., Jung, W., & Lee, S. G. (2015). Reversed CF: a fast collaborative filtering algorithm using a k-nearest neighbor graph. Expert Systems with Applications, 42(8), 4022-4028.

[16] Su, G. C., & Kim, S. B. (2017). Feature network-driven quadrant mapping for summarizing customer reviews. Journal of Systems Science & Systems Engineering, 646-663.

[17] Tian, F., Wu, F., Chao, K. M., Zheng, Q., Shah, N., & Lan, T., et al. (2016). A topic sentence-based instance transfer method for imbalanced sentiment classification of chinese product reviews. Electronic Commerce Research & Applications, 16, 66-76.

[18] Wang, C., & Tang, X. (2016). The Online Debate Networks Analysis: A Case Study of Debates at Tianya Forum. International Symposium on Knowledge and Systems Sciences (pp.140-150). Springer Singapore.

[19] Xu L H, Lin H F, Zhao J. (2008). The construction of emotional lexical ontology. Journal of Chinese Information Processing, 22(1), 116-122.

[20] Yang, X., Liang, C., Zhao, M., Wang, H., Ding, H., & Liu, Y., et al. (2017). Collaborative filtering-based recommendation of online social voting. IEEE Transactions on Computational Social Systems, 4(1), 1-13.

[21]Yang, X., Yang, G., & Wu, J. (2016). Integrating rich and heterogeneous information to design a ranking system for multiple products. Decision Support Systems, 84, 117-133.
[22]Yu, J., Zha, Z. J., Wang, M., & Chua, T. S. (2011). Aspect ranking: identifying important product aspects from online consumer reviews. Meeting of the Association for Computational Linguistics: Human Language Technologies. Association for Computational Linguistics, 1496-1505.
[23]Zhang, K., Cheng, Y., Liao, W. K., & Choudhary, A. (2011). Mining millions of reviews: a technique to rank products based on importance of reviews. International Conference on Electronic Commerce, 1-8.
[24]Zhang, X., Qiao, Z., Tang, L., Fan, W., Fox, E., & Wang, G. (2016). Identifying Product Defects from User Complaints: A Probabilistic Defect Model. In Proceedings of 22nd Americas Conference on Information Systems (AMCIS).
[25]Zhang, Z., Guo, C., & Goes, P. (2013). Product comparison networks for competitive analysis of online word-of-mouth. ACM Transactions on Management Information Systems, 3(4), 1-22.

Chonghui Guo is a professor of the Institute of Systems Engineering, Dalian University of Technology, Dalian, China. He received the B.S. degree in Mathematics from Liaoning University in 1995, M.S. degree in Operational Research and Control Theory in 1999 and Ph.D. degree in Management Science and Engineering from Dalian University of Technology in 2002. He was a postdoctoral research fellow in the Department of Computer Science in Tsinghua University, Beijing, China. His studies concentrate on data mining and knowledge discovery. He has published over 100 peer-reviewed papers in academic journals and conferences, besides 5 text-books and 2 monographs. He has been the Principal Investigator on over 10 research projects from the Government and the Industry.

Zhonglian Du received her B.S. degree in 2015 from School of Business, Jiangnan University. She is currently a master student in the Institute of Systems Engineering, Dalian University of Technology. Her general research interests include text mining and business intelligence. Her papers have been published and presented on journals and conferences such as the System Engineering and the 18th International Symposium on Knowledge and Systems Sciences.

Xinyue Kou received her B.S. degree in 2016 from School of Economics and Management, China University of Petroleum. She is currently a master student in the Institute of Systems Engineering, Dalian University of Technology. Her general research interests include two-sided matching and group decision. Her papers have been published and presented in journals and conferences such as the Expert Systems with Applications and Knowledge-Based Systems.