

Online Shopping Behavior Study Based on Multi-granularity Opinion Mining: China Versus America

Qingqing Zhou^{1,2} · Rui Xia³ · Chengzhi Zhang^{1,2,4}

Received: 28 April 2015 / Accepted: 25 January 2016 / Published online: 12 February 2016
© Springer Science+Business Media New York 2016

Abstract With the development of e-commerce, many products are now being sold worldwide, and manufacturers are eager to obtain a better understanding of customer behavior in various regions. To achieve this goal, most previous efforts have focused mainly on questionnaires, which are time-consuming and costly. The tremendous volume of product reviews on e-commerce Web sites has seen a new trend emerge, whereby manufacturers attempt to understand user preferences by analyzing online reviews. Following this trend, this paper addresses the problem of studying customer behavior by exploiting recently developed opinion mining techniques. This work is novel for three reasons. First, questionnaire-based investigation is automatically enabled by employing algorithms for template-based question generation and opinion mining-based answer extraction. Using this system, manufacturers are able to obtain reports of customer behavior featuring a much larger sample size, more direct information, a higher degree of automation, and a lower cost. Second, international customer behavior study is made easier by integrating tools for multilingual opinion mining. Third, this is the first time an automatic questionnaire

investigation has been conducted to compare customer behavior in China and America, where product reviews are written and read in Chinese and English, respectively. Our study on digital cameras, smartphones, and tablet computers yields three findings. First, Chinese customers follow the Doctrine of the Mean and often use euphemistic expressions, while American customers express their opinions more directly. Second, Chinese customers care more about general feelings, while American customers pay more attention to product details. Third, Chinese customers focus on external features, while American customers care more about the internal features of products.

Keywords Sentiment analysis · Opinion mining · Customer behavior study · Automatic questionnaire generation

Introduction

With the rapid development of e-commerce, more and more businesses are operating on a global scale. For manufacturers, customer behavior plays a key role in decision making regarding product design and marketing. It is also important for transactional enterprises to conduct questionnaire-based surveys of customer behavior. Most existing research on customer behavior is conducted via manually administered questionnaires, which is expensive, time-consuming, and prone to error. Nowadays, there is a tremendous volume of product reviews available on e-commerce Web sites, and manufacturers are attempting to understand user preferences by accessing online customer reviews. For example, C&A Marketing¹ analyzes

✉ Chengzhi Zhang
zhangcz@njust.edu.cn

¹ Department of Information Management, Nanjing University of Science and Technology, Nanjing 210094, China

² Alibaba Research Center for Complex Sciences, Hangzhou Normal University, Hangzhou 311121, China

³ School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

⁴ Jiangsu Key Laboratory of Data Engineering and Knowledge Service, Nanjing University, Nanjing 210093, China

¹ <http://www.camarketing.com>.

| | |
|----------------|--|
| English review | <p>★★★★★ Not Happy</p> <p>Battery life is very disappointing. Uses at least 50% charge a day without a lot of use!</p> |
| Chinese review | <p>★★★★★ 正品,质量好,快递迅速 (It's a certified product with high quality, and the express is fast.)</p> <p>外包装规矩, 正品, 质量有保证。亚马逊快递迅速, 服务好</p> <p>(The package is proper, good product, guaranteed quality, Fast express mail by Amazon, good service)</p> |

Fig. 1 Examples of English and Chinese reviews on Apple iPhone 5s

product reviews on Amazon to understand users' requirements, and this information is then used to design new products. To this end, product reviews can reveal the attitudes and preferences of customers in different regions. Two illustrative reviews of the *Apple iPhone 5s* are shown in Fig. 1. As can be seen, the customer in America is unhappy with the battery, while the customer in China is satisfied with the packaging and service. These reviews are typical among American and Chinese consumers, respectively. In general, there are numerous differences between customers in different regions, who usually focus on different aspects of products and display different attitudes in relation to the same aspects. To identify these differences, we exploit recently developed opinion mining techniques to complete questionnaires automatically based on reviews written by Chinese and American customers.

We generate questions automatically based on templates and extract answers by sentiment analysis. (1) *For question generation*, we build question templates for overall sentiment, brand preference, and purchase decision factors, respectively. (2) *For answer extraction*, we use the topic model for aspect extraction, and we conduct multi-granularity sentiment analysis to identify sentiment polarities.

Using three IT product domains as a case study, we aim to identify preferences of customers in different regions. Three interesting results are obtained from the perspective of sociology, based on the empirical results:

- (1) Chinese customers obey the Doctrine of the Mean, while American customers express opinions more directly
- (2) Chinese customers pay more attention to general feelings, while American customers focus on product details
- (3) Chinese customers are more concerned with external features, while American customers pay more attention to internal features.

The remainder of this paper is organized as follows. In “[Related Work](#)” section, we review related work. Data collection is introduced in “[Data Collection](#)” section. In “[Methodology](#)” section, we present our methodology for automatic question generation and answer extraction. “[Empirical Study](#)” section presents our comparisons and

analysis. “[Conclusion and Future Work](#)” section concludes.

Related Work

There are three areas of research related to our study: online shopping behavior, automatic question and answer generation, and sentiment analysis.

Online Shopping Behavior Study

In the field of e-commerce, many studies have focused on online shopping behavior. Chatterjee examined the effect of negative reviews on retailer evaluation and patronage intention, given that the consumer has already made a product/brand decision [9]. The results indicated that the extent of their word-of-mouth search depended on the consumers' reasons for choosing an online retailer. Park et al. [41] proved that the quality of online reviews had a positive effect on customers' purchase intention and that purchase intention increased as the number of reviews increased. They also found that low-involvement consumers were affected by the quantity of reviews rather than review quality, but high-involvement consumers were mainly affected by review quantity when the review quality was high. Lee et al. [21] investigated the effects of negative online consumer reviews on consumers' attitudes to products. The experimental results showed that a high proportion of negative online consumer reviews elicited a conformity effect.

Vermeulen and Seegers [49] applied consideration set theory to model the impact of online hotel reviews on consumer choice. They proposed an experimental study that included review valence, hotel familiarity, and reviewer expertise as independent factors. The results showed that on average, exposure to online reviews enhanced hotel consideration among consumers. Park and Kim [40] showed that the effect of cognitive fit (types of reviews) on purchase intention was stronger for experts than for novices, while the effect of the number of reviews on purchase intention was stronger for novices than for experts. Zhu and Zhan [62] examined how product and

consumer characteristics moderated the influence of online consumer reviews on product sales using data from the video game industry. Their findings indicated that online reviews were more influential for less popular games and games whose players had more Internet experience.

As a result of the emergence of e-commerce and the development of global markets, researchers have begun to compare customers' online shopping preferences and behaviors in different cultures [35]. Mahmood et al. [31] conducted an online shopping behavior study using data from 26 nations and analyzed constructs using a structural equation model. They found that trust and economic conditions, but not educational level and technological savvy, had a significant positive effect on online shopping behavior. Hwang et al. [18] investigated online shopping preferences in three countries, the USA, Korea, and Turkey. They found significant differences in online shopping preferences, especially in relation to information accuracy, security, and price comparison.

From the above, it can be seen that existing studies of customer behavior can be divided into two categories. In the first category, studies have focused on coarse-grained review mining, i.e., only sentiment polarities of reviews are used to conduct studies of customer behavior. In the second category, comparative studies are conducted on customers' online shopping behavior. These are mainly based on traditional survey methods such as questionnaires. These methods have several disadvantages, including small sample sizes, long cycles, and high costs. In summary, existing studies are mainly focused on relatively macroscopic online shopping behavior rather than focusing on specific products or differences in customer attitudes. In contrast, our study compares online shopping behavior by generating questions and answers automatically via multi-granularity review mining.

Automatic Question and Answer Generation

In response to the abovementioned disadvantages of traditional methods, research on automatic question and answer generation has attracted the interest of many researchers. Sumita et al. proposed automatic generation of fill-in-the-blank questions (FBQs) together with testing based on item response theory (IRT) to measure English proficiency [45]. Their method involved three steps. First, an FBQ was generated from a given sentence in English. Second, each of the candidate for incorrect choices was verified using the Web. Finally, the blank sentence, the correct choice, and the incorrect choices surviving verification were arranged to form the FBQ. Brown et al. [6] described an approach to automatically generate questions for vocabulary assessment. Unlike traditional handwritten assessments, they generated six types of vocabulary

questions, including word bank and multiple-choice questions. Their experimental results suggested that these automatically generated questions gave a measure of vocabulary skill that correlated well with subject performance on independently developed manually prepared questions.

The lack of automatic methods for evaluating answers is an impediment to progress in the field of question and answer generation. Lin et al. [27] proposed a measure called *Pourpre* for automatically evaluating answers to complex questions based on *n*-gram co-occurrences between machine output and a human-generated answer key. Based on the engagement taxonomy and benefits of question answering during the algorithm visualization, Myller and Niko [36] proposed incorporating automatic question generation into a program visualization tool, *Jeliot 3*. Davidov et al. [13] presented a novel framework for the discovery and representation of general semantic relationships that exist between lexical items. To assess the quality of relationships, they used pattern clusters to automatically generate Scholastic Assessment Test analogy questions. Papasalouros et al. [39] presented an innovative approach to generating multiple-choice questions automatically. Their method was based on domain-specific ontologies and was independent of lexicons or other linguistic resources. With the increasing popularity of community-based question-answering services such as Yahoo! Answers, huge numbers of user-generated questions are now available. Lin used these data along with search engine query logs to create a shared question generation task that aimed to automatically generate questions given a query [26]. Liu et al. [29] presented a novel automatic question generation approach that generated trigger questions as a form of support for students learning through writing. Experimental results showed that the questions generated by their system significantly outscored generic questions on overall quality measures.

From the above, we can see that existing studies are either about information retrieval or about generating tests for skills upgrading. There are few studies focusing on questionnaire generation from online reviews. Our work generates questions and extracts answers automatically to compare online behavior via sentiment analysis of online reviews.

Sentiment Analysis

To identify answers automatically, we conducted multi-granularity sentiment analysis at both the document level and the aspect level. Document-level sentiment analysis aims to determine whether a document expresses a sentiment (i.e., positive or negative) [28, 59]. Aspect-level sentiment analysis performs finer-grained analysis to

identify various aspects and the sentiment polarities that people express in relation to them [19, 60].

Document-Level Sentiment Analysis

Document-level sentiment analysis is used to predict whether the document as a whole expresses a positive sentiment or a negative one [59]. This can be achieved by using one of two learning methods: supervised and unsupervised. In our case, many existing supervised learning methods can be applied [38]. For example, Mullen and Collier [34] used support vector machine (SVM) to conduct sentiment analysis, which brought together diverse sources of potentially pertinent information. To improve the accuracy of sentiment classification, some special techniques for sentiment analysis have been proposed, including those by Xia et al. [53–56], Li et al. [23], Nakagawa et al. [37] and Maas et al. [30]. One of the most classic unsupervised learning approaches to sentiment analysis was proposed by Turney et al. [48]. In their work, they compiled syntactic patterns that people used to express opinions based on parts of speech and then performed sentiment analysis measuring the pointwise mutual information from these patterns, which is efficient. Other unsupervised learning methods such as those of Taboada et al. [46] and Denecke and Kerstin [14] were based on sentiment lexicons. Chikersal et al. described a Twitter sentiment analysis system that classified a tweet as either positive or negative based on its overall tweet-level polarity [10]. The proposed method was evaluated using two publicly available Twitter corpora to illustrate its effectiveness.

Aspect-Level Sentiment Analysis

Rather than providing isolated opinions about a product, customers tend to compare specific features of products. Therefore, it is necessary to conduct aspect-level sentiment analysis [28], including aspect extraction and aspect sentiment identification [15, 60]. Many studies of aspect extraction have been carried out. Popescu and Etzioni [42] improved the mining methods via pointwise mutual information to extract product features and opinions from reviews. Jo and Oh [19] proposed the Sentiment-LDA model and the aspect and sentiment unification model (ASUM) for reviews of electronic devices and restaurants. Their results showed that ASUM outperformed other generative models and performed similarly to supervised classification models. Moghaddam and Este [33] developed some design guidelines for aspect-based opinion mining by discussing a series of increasingly sophisticated LDA models.

For aspect sentiment identification, Ding et al. [15] proposed a simple method based on a lexicon, which was shown

to be quite effective in a large number of domains. Blair-Goldensohn et al. [3] and Yu et al. [60] used similar methods to identify aspect sentiments. Wei and Gulla [51] proposed a hierarchical classification model to determine the scope of each sentiment expression. They also presented an approach for generating multi-unigram features to enhance a negation-aware naïve Bayes classifier for sentiment analysis of sentences in product reviews [52]. Thet et al. [47] performed fine-grained analysis to determine both the orientation and strength of a reviewer's sentiments in relation to various aspects of a movie. Gangemi et al. [16] inferred the conceptual and affective information associated with opinions. To solve contextual polarity ambiguity, Cambria et al. [7] focused on the use of semantics to better explore opinion-level context. Xia et al. [57] explored opinion-level context to solve contextual polarity ambiguity. Das and Bandyopadhyay [12] proposed tagging sentence-level emotions and designed a valence based on the word-level constituents on the SemEval 2007 that affect sensing news corpus. Cambria et al. [8] introduced novel statistical approaches to concept-level sentiment analysis that went beyond a mere syntax-driven analysis of text and provided a semantics-based method. Li et al. [24] implemented a generic stock-price prediction framework by analyzing the impact of news reports on stock prices via sentiment analysis. Rill et al. [43] presented a system called PoliTwi, which was designed to detect emerging political topics earlier on Twitter than on other standard information channels. Agarwal et al. [1] focused on finding good feature sets to conduct concept-level sentiment analysis and achieved excellent performance.

In this study, we conduct multi-granularity sentiment analysis to generate questions and answers automatically to identify review-based differences between customers from different regions. Specifically, we use a supervised machine learning method to conduct document-level sentiment analysis and a lexical affinity method for aspect-level sentiment analysis. We aim to identify differences between opinions expressed by customers in different regions from the perspective of sociology. To the best of our knowledge, this is the first work to compare the online shopping behavior of Chinese and American customers from this perspective via automatic question and answer generation.

Data Collection

This study aims to compare the behavior of online customers in China and America. Therefore, we decided to collect reviews on the same products from Amazon.com and Amazon.cn, respectively. Note that the reviews from Amazon.cn are mostly written in Chinese, while those from Amazon.com are mostly written in English. The sample

Table 1 Statistics of our review corpora

| Products | Brands | #Chinese reviews | #English reviews |
|-----------------|---------|------------------|------------------|
| Digital camera | Cannon | 2245 | 2550 |
| | Nikon | 2172 | 2419 |
| Smart phone | Apple | 1805 | 3578 |
| | Samsung | 2579 | 5525 |
| Tablet computer | Apple | 485 | 3189 |
| | Samsung | 253 | 927 |

Table 2 Statistics on the annotated review corpora

| | #Chinese reviews | #English reviews |
|-----------|------------------|------------------|
| #positive | 1200 | 2400 |
| #negative | 600 | 1200 |
| sum | 1800 | 3600 |

includes reviews of three kinds of IT products: digital cameras (Canon and Nikon), smartphones (Apple iPhone and Samsung Galaxy) and tablet computers (Apple iPad and Samsung Galaxy).² We collected 10,498 Chinese reviews and 18,498 English reviews from October 7, 2009, to August 5, 2014. After deleting low-quality reviews, including reviews with no text, garbled reviews, and advertisements, 9539 Chinese reviews and 18,188 English reviews were included in our sample. Statistics on the sample of reviews are presented in Table 1.

To identify customers' attitudes toward the products, we used the supervised machine learning method to conduct sentiment analysis. Therefore, we constructed a training sample by manually tagging sentiment polarities of 1800 randomly selected Chinese reviews and 3600 randomly selected English reviews. Statistics on these manually annotated reviews are presented in Table 2.

Methodology

Framework

In this study, our aim is to identify the cognitive differences between Chinese and American customers by automatic question and answer extraction, which is based on opinion mining. The framework, which is shown in Fig. 2, includes three major parts.

1. *Automatic question generation* We generate questions automatically based on templates, which includes

overall and product-level, brand-level and aspect-level questions.

2. *Automatic answer extraction* We extract answers for the three levels using multi-granularity sentiment analysis. Specifically, we use overall sentiment analysis to extract answers about overall and product satisfaction and brand satisfaction, and we use aspect-level sentiment analysis to extract answers about frequent aspects and popular aspects.
3. *Comparison and analysis* We use a bilingual dictionary³ to map Chinese and English aspects prior to comparisons. We analyze differences between Chinese and American customers at three levels: overall sentiment, brand preferences, and purchase decision factors.

The methods used for automatic question generation and answer extraction are described in detail in “Automatic Question Generation” and “Automatic Answer Extraction” sections, respectively.

Automatic Question Generation

We automatically generate questions using manually constructed templates. There are three subtasks in this part: (1) overall sentiment, (2) brand preference, and (3) purchase decision factors, which are elaborated as follows.

1. Generating questions of overall sentiment

One of our goals was to compare the differences between Chinese and American customers in relation to overall and product satisfaction. Therefore, we designed the question template shown in Table 3. This template includes overall-level and product-level questions. Regarding overall satisfaction, the blank space is filled with *digital products*. For product-level satisfaction, the blank space is filled with either *digital cameras*, *smartphones*, or *tablet computers*.

2. Generating questions of brand preference

To compare the differences in brand preference, the question template is *Which brand of _____ do you prefer?* as shown in Table 4. The blank space is filled with either *digital camera*, *smartphones*, or *tablet computers*. Each question has two options.

3. Generating questions of purchase decision factors

To extract frequent and popular aspects, respectively, we designed two question templates: (1) *What aspects of _____ do you care about?* and (2) *What aspects of _____ are you satisfied with?* as shown in Table 5. The blank

² We select the brands with the most two quantity of reviews from Amazon as our data set.

³ <https://catalog.ldc.upenn.edu/LDC2002L27>

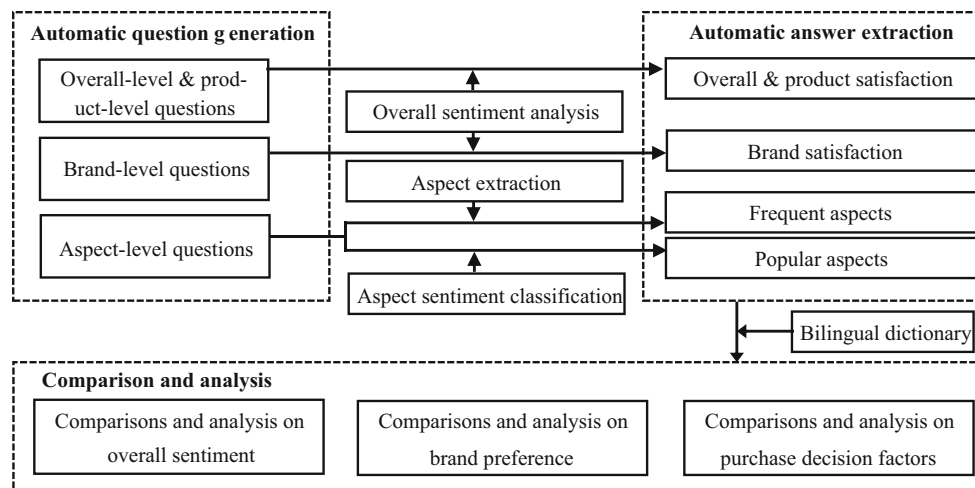


Fig. 2 Framework of opinion comparison between Chinese and American customers

Table 3 Question template for overall sentiment (using digital cameras as an example)

| Do you like ____? | | | |
|-----------------------|-------------|-------------------------|---|
| Level | Question | | |
| Overall satisfactions | Do you like | <i>digital products</i> | ? |
| Product satisfactions | Do you like | <i>digital cameras</i> | ? |

spaces are filled with either *digital cameras*, *smartphones*, or *tablet computers*.

Automatic Answer Extraction

In previous studies, comparisons of online shopping behavior were mainly conducted using traditional questionnaires, which is a costly and time-consuming process. In this study, we conduct the comparison by generating questionnaires and extracting answers automatically. Specifically, we extract answers using opinion mining, including overall sentiment classification, aspect extraction, and aspect sentiment classification. There are three subtasks in this part: (1) extracting answers regarding overall sentiment, (2) extracting answers regarding brand preference, and (3) extracting answers regarding purchase decision factors.

Table 4 Question template for brand preference (using digital cameras as an example)

| | | | |
|------------------------------------|------------------------|----------------|----------------------------------|
| Which brand of ____ do you prefer? | | | |
| Which brand of | <i>digital cameras</i> | do you prefer? | A. <i>Cannon</i> B. <i>Nikon</i> |

1. Extracting answers of overall sentiment

Sentiment classification is a hot topic in natural language processing. It can be used to analyze online public opinion and filter noise. Document-level sentiment analysis is used to predict whether the review as a whole expresses a positive sentiment or a negative one. To extract answers regarding overall and product satisfaction, we conduct overall sentiment classification.⁴ Specifically, we used linear SVM [11] as the classification model with CHI [20] as the feature selection method and term frequency-inverse document frequency (TF-IDF) [44] as the feature weighting method. In addition, we use Jieba⁵ for Chinese word segments and LibSVM as the classifier. For overall satisfaction, we first conduct overall sentiment classification to identify sentiment polarity in reviews. Then, we count the numbers of positive and negative reviews. Finally, we obtain a value for overall satisfaction via the value for overall sentiment (OS), which can be calculated via Eq. 1:

$$OS = \frac{\#pos}{\#pos + \#neg} \quad (1)$$

where $\#pos$ denotes the number of positive reviews, and $\#neg$ denotes the number of negative reviews. The same method is also used in relation to product satisfaction.

For example, to answer the question *Do you like digital products?* we conduct overall sentiment analysis to obtain OS values. The OS values for the Chinese samples and English samples are 0.8873 and 0.8436, respectively. Therefore, it can be seen that most customers are satisfied with digital products.

⁴ Overall sentiment classification here is similar to document-level sentiment classification.

⁵ <https://github.com/fxsjy/jieba>.

Table 5 Question templates for purchase decision factors (using digital cameras as an example)

| | | | |
|-----------------|-----------------|------------------------|-------------------------|
| Frequent aspect | What aspects of | <i>digital cameras</i> | do you care about? |
| Popular aspect | What aspects of | <i>digital cameras</i> | are you satisfied with? |

2. Extracting answers of brand preference

To obtain corresponding results in relation to brand satisfaction, first, we conduct overall sentiment classification to identify the sentiment polarities of reviews about brands A and B. Then, we count the number of reviews for each brand. Finally, we obtain brand preference (BP) values using Eq. 2:

$$BP = \begin{cases} A & \text{if } OS(A) > OS(B), \\ B & \text{otherwise} \end{cases} \quad (2)$$

where $OS(A)$ and $OS(B)$ can be calculated via Eq. 3:

$$OS(X) = \frac{\#pos(X)}{\#pos(X) + \#neg(X)} \quad (3)$$

where $OS(X)$ denotes OS value of brand X, $\#pos(X)$ denotes the number of positive reviews about brand X, and $\#neg(X)$ means the number of negative reviews about brand X.

For example, the answer to the question *Which brand of smartphone do you prefer?* has two options: *Apple* and *Samsung*. We find that $OS(Apple)$ is 0.6401 and $OS(Samsung)$ is 0.7548; hence, the BP is Samsung.

3. Extracting answers of purchase decision factors

To find frequent and popular aspects, we use two processes: (1) aspect extraction and (2) aspect sentiment classification.

Aspect extraction is used to extract frequent aspects. The topic model method is widely used for aspect extraction [25, 61]. In addition, when people comment on different aspects of an entity, the vocabulary that they use generally converges. Thus, nouns that appear frequently are usually genuine and important aspects [38]. In this study, we combine the two methods to extract aspects. Specifically, we first tag the parts of speech for words in each topic generated by LDA using Jieba⁶ for Chinese words and NLTK⁷ for English words and then choose nouns as candidate aspects extracted with LDA [4]. We calculate the sum of probabilities for each noun in every topic and choose the highest ones as candidate aspects.

Then, we use frequent aspect (FA) values to find more candidate aspects, which can be calculated using Eq. 4:

$$FA(Y) = \#pos(Y) + \#neg(Y) \quad (4)$$

where $\#pos(Y)$ denotes the numbers of positive reviews about aspect Y, and $\#neg(Y)$ means the numbers of negative reviews about aspect Y. Finally, we extract the most frequent aspects by sorting the FA values of all candidate aspects.

For example, for the question *What aspects of digital cameras do you care about?*, we first use the combined method to extract frequent aspects about *digital cameras*. Then, we calculate the FA value for each frequent aspect. $FA(package)$ is 709 and $FA(genuine\ product)$ is 684 in the Chinese samples, which are much higher than the FA values of other aspects. Therefore, it can be seen that Chinese customers care more about *package* and whether it is a *genuine product* than about other aspects.

Aspect sentiment classification is used to detect attitudes toward a given aspect. For each review, we first identify sentiment words via sentiment lexicons in Chinese⁸ and English.⁹ It is difficult to determine which aspect a sentiment word is describing when a review expresses sentiments on multiple aspects. In general, if the distance between an aspect word and a sentiment word is short, the sentiment word is more likely to describe the aspect. Therefore, we compute the sentiment polarity of each aspect in a review by measuring the distance between an aspect word and a sentiment word [15]. Sentiment polarity (SP) of aspect Y in review s can be calculated using Eq. 5:

$$SP(Y, s) = \begin{cases} +1 & \text{if } \sum_{w_i: w_i \in s \cap w_i \in V} \frac{w_i.SO}{dis(w_i, Y)} > 0, \\ -1 & \text{otherwise} \end{cases} \quad (5)$$

where $w_i.SO$ is calculated by Eq. 6:

$$w_i.SO = \begin{cases} +1 & \text{if } w_i \text{ is a positive word,} \\ -1 & \text{otherwise} \end{cases} \quad (6)$$

where w_i denotes a sentiment word, V means the set of all sentiment words, s is the review that contains the aspect Y, and $dis(w_i, Y)$ denotes the distance between aspect Y and sentiment word w_i in review s. The sentiment score of the word w_i is denoted by $w_i.SO$.

To find popular aspects, we first conduct aspect sentiment classification to calculate SP values of aspects in each review. Then, we use popular aspect (PA) values to extract popular aspects, which can be calculated using Eq. 7:

⁶ <https://github.com/fxsjy/jieba>.

⁷ <http://www.nltk.org/>.

⁸ http://ir.dlut.edu.cn/EmotionOntologyDownload.aspx?utm_source=weibolife.

⁹ <https://hlt.fbk.eu/technologies/sentiwords>.

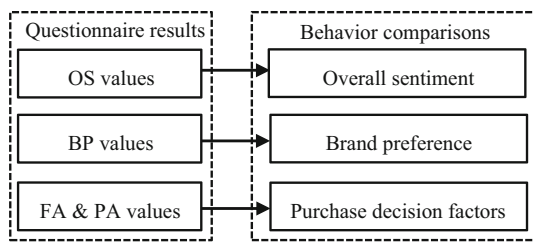


Fig. 3 The flowchart of empirical study for online shopping behavior

$$PA(Y) = \frac{\sum_{i=1}^m (SP(Y, s_i) + 1)}{2 \sum_{i=1}^m |SP(Y, s_i)|} \quad (7)$$

where m denotes the number of reviews about aspect Y .

For example, for the question *What aspects of digital cameras are you satisfied with?*, $PA(\text{quality})$ is 0.9900 and $PA(\text{flashlight})$ is 0.9611 in English samples, which are much higher than PA values of other aspects. Therefore, it can be seen that English customers are more satisfied with the overall quality and the flashlight than with other aspects.

Empirical Study

We conducted our analysis using digital cameras, smartphones, and tablet computers. The process is shown in Fig. 3. The results of our investigation are presented in “[Questionnaire Investigation Results](#)” section. Based on the analysis in “[Questionnaire Investigation Results](#)” section, we present the results of our analysis of online consumer behavior in “[Customer Behavior Comparison Results](#)” section. Finally, in order to illustrate the value of our proposed method, we present an example of a smartphone promotion in “[A Case for Smartphone Promotion](#)” section.

Questionnaire Investigation Results

We obtained our results using multi-granularity sentiment analysis. There are three types of results: (1) overall sentiment, (2) brand preference, and (3) purchase decision factors, which are elaborated as follows.

Overall Sentiment

The results of the analysis of overall sentiment include overall satisfaction and product satisfaction. Regarding

overall satisfaction, the question we asked was *Do you like digital products?* Our sample consisted of 9539 Chinese reviews and 18,188 English reviews. Table 6 shows the results in terms of overall satisfaction. It can be seen that the OS value in Chinese reviews is higher than that in English reviews, which indicates that Chinese customers’ satisfaction regarding digital products is higher than that of American customers.

For product satisfaction, the question we asked was *Do you like digital camera/smartphone/tablet computer?* We collected 4417 Chinese reviews and 4969 English reviews about digital cameras, 4384 Chinese reviews and 9103 English reviews about smartphones, and 738 Chinese reviews and 4116 English reviews about tablet computers. Table 7 shows the results in relation to product satisfaction. For digital cameras, the Chinese OS value is 91.78 %, while the English OS value is 94.52 %. For smartphones, the Chinese OS value is 82.27 %, while the English OS value is 70.96 %. For tablet computers, the Chinese OS value is 92.14 %, while the English OS value is 87.61 %. The results indicate that (1) for digital cameras, both Chinese and American customers are satisfied with the cameras, although the American customers’ satisfaction tends to be slightly higher, and (2) for smartphones and tablet computers, the OS values of American customers are significantly lower than those of Chinese customers.

Brand Preference

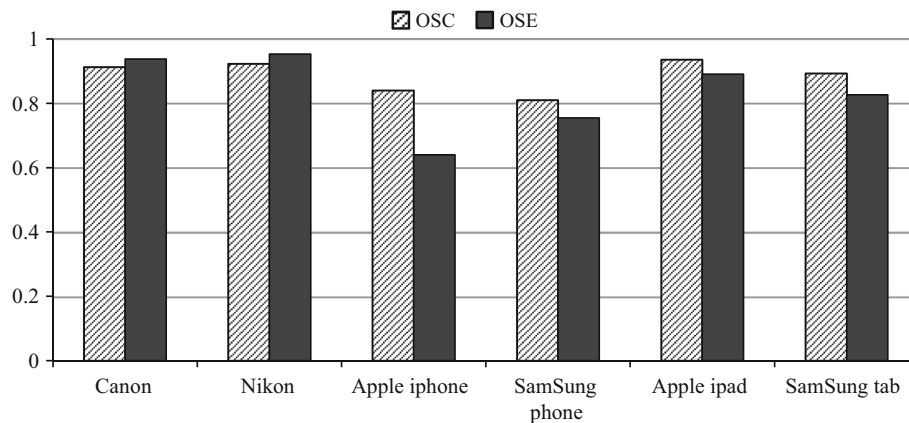
The results in relation to brand preference are shown in Fig. 4. OSC denotes the Chinese OS value for each brand, while OSE denotes the English OS value for each brand. For digital cameras, we collected 4795 reviews of Canon products (2245 Chinese reviews and 2550 English reviews) and 4591 reviews of Nikon products (2172 Chinese reviews and 2419 English reviews). For smartphones, we collected 5383 reviews of Apple products (1805 Chinese reviews and 3578 English reviews) and 8104 reviews of Samsung products (2579 Chinese reviews and 5525 English reviews). For tablet computers, we collected 3674 reviews of Apple products (485 Chinese reviews and 3189 English reviews) and 1180 reviews of Samsung products (253 Chinese reviews and 927 English reviews). It can be seen from Fig. 4 that the differences between the OSCs are smaller than those between the OSEs in all three domains, especially in the smartphone domain, which indicates that

Table 6 Results of overall sentiment

| | #Review | #Pos. | #Neg. | OS |
|-----------------|---------|--------|-------|--------|
| Chinese reviews | 9539 | 8341 | 1198 | 0.8873 |
| English reviews | 18,188 | 14,763 | 3425 | 0.8436 |

Table 7 Results of product sentiment

| Products | #Review | #Pos. | #Neg. | OS |
|-----------------|---------|-------|-------|--------|
| Chinese reviews | | | | |
| Camera | 4417 | 4054 | 363 | 0.9178 |
| Smartphone | 4384 | 3,607 | 777 | 0.8227 |
| Tablet computer | 738 | 680 | 58 | 0.9214 |
| English reviews | | | | |
| Camera | 4969 | 4697 | 272 | 0.9452 |
| Smartphone | 9103 | 6460 | 2,643 | 0.7096 |
| Tablet computer | 4116 | 3606 | 510 | 0.8761 |

**Fig. 4** OS values of different brands

American customers have more obvious brand preferences than Chinese customers.

Purchase Decision Factors

To extract purchase decision factors, we performed three subtasks: (1) extracting the most frequent aspects, (2) extracting the most popular aspects, and (3) extracting frequent aspects that are a concern of both Chinese and American customers.

1. Most frequent aspect extraction and comparison

In this study, we use gensim¹⁰ to identify topics with both aspects and sentiments. The LDA algorithm in gensim is based on Hoffman et al. [17]. They developed an online variational Bayes algorithm for LDA. We use the POS taggers Jieba and NLTK to perform Chinese and English POS tagging, respectively. We choose nouns extracted by LDA as candidate aspects and then choose the top 10 aspects in terms of aspect frequency as real aspects via FA values. For the question *Which aspects of digital camera / smartphone / tablet computer do you care about / are you*

satisfied with?, we collected 3280 Chinese reviews and 7098 English reviews about digital cameras, 4040 Chinese reviews and 21,798 English reviews about smartphones, and 1474 Chinese reviews and 8780 English reviews about tablet computers. The 10 most frequent aspects of digital cameras, smartphones, and tablet computers are given in Tables 8a, 9a, and 10a respectively. We use the entropy value to measure the differences between Chinese and American customers regarding the 10 most frequent aspects, which can be calculated using Eq. 8:

$$E_{\text{fre}} = - \sum_{i=1}^{N_{\text{fre}}} \left(\frac{FA_i}{\sum_{i=1}^{N_{\text{fre}}} FA_i} * \log \frac{FA_i}{\sum_{i=1}^{N_{\text{fre}}} FA_i} \right) \quad (8)$$

where FA_i means frequent aspect (FA) value of the i th aspect, which can be calculated by Eq. 4, and N_{fre} denotes the number of frequent aspects. Table 11 shows the entropy values of Chinese and American customers regarding frequent aspects. From Table 11, it can be seen that Chinese customers' concerns about frequent aspects are less homogeneous than those of American customers in two of the three domains. Hence, we can conclude that concerns about frequent aspects are obviously different in all three domains between Chinese and American

¹⁰ <http://radimrehurek.com/gensim/>.

Table 8 Top aspects of digital cameras

| Chinese reviews | | English reviews | | Chinese reviews | | English reviews | |
|---------------------------------|-----|-----------------|-------|--------------------------------|--------|-----------------|--------|
| Aspects | FA | Aspects | FA | Aspects | PA | Aspects | PA |
| <i>(a) Top frequent aspects</i> | | | | <i>(b) Top popular aspects</i> | | | |
| Package | 709 | Lens | 2,118 | Genuine product | 0.9985 | Quality | 0.9900 |
| Genuine product | 684 | Quality | 1,604 | Logistics | 0.9295 | Flashlight | 0.9611 |
| Price | 488 | Price | 994 | Price | 0.9262 | Battery | 0.9402 |
| Lens | 467 | Lighter | 950 | Package | 0.9027 | Price | 0.9054 |
| Logistics | 312 | Screen | 406 | Quality | 0.8901 | Lens | 0.8820 |
| Quality | 273 | Sensor | 272 | Appearance | 0.8772 | Appearance | 0.8720 |
| Battery | 149 | Button | 270 | Lens | 0.8287 | Screen | 0.8670 |
| Service | 82 | Appearance | 250 | Battery | 0.7987 | Button | 0.8593 |
| Screen | 59 | Battery | 234 | Screen | 0.7288 | Sensor | 0.8382 |
| Appearance | 57 | Shutter | 188 | Service | 0.6829 | Shutter | 0.8298 |

Table 9 Top aspects of smartphones

| Chinese reviews | | English reviews | | Chinese reviews | | English reviews | |
|---------------------------------|-----|-----------------|------|--------------------------------|--------|-----------------|---------|
| Aspects | FA | Aspects | FA | Aspects | PA | Aspects | PA |
| <i>(a) Top frequent aspects</i> | | | | <i>(b) Top popular aspects</i> | | | |
| Logistics | 760 | App | 6910 | Genuine product | 0.9848 | Quality | 0.97198 |
| Genuine product | 726 | Screen | 3200 | Price | 0.8697 | Appearance | 0.9154 |
| Screen | 494 | Battery | 2458 | Logistics | 0.8592 | Camera | 0.8747 |
| Service | 481 | Apperance | 2034 | Quality | 0.813 | Battery | 0.8373 |
| Appearance | 391 | Price | 1538 | Appearance | 0.7928 | Charger | 0.8363 |
| Price | 376 | Camera | 1516 | Screen | 0.7874 | Screen | 0.8363 |
| Battery | 274 | Service | 1164 | Service | 0.7775 | Price | 0.8349 |
| Quality | 246 | Quality | 1142 | App | 0.6918 | Service | 0.8196 |
| App | 159 | Charger | 1014 | Memory | 0.6917 | Button | 0.7981 |
| Memory | 133 | Button | 822 | Battery | 0.6496 | App | 0.737 |

customers. Specifically, from Table 8a, we can see that for digital camera reviews, the most frequent aspect in Chinese reviews is *package*, followed by *genuine product* and *price*, while American customers are most concerned about the *lens*, followed by *quality* and *price*. From Table 9a, we can see that for smartphone reviews, the most frequent aspect in Chinese reviews is *logistics*, followed by *genuine product* and *screen*, while American customers are most concerned about the *apps*, followed by *screen* and *battery*. From Table 10a, we can see that for tablet computer reviews, the most frequent aspect in Chinese reviews is *screen*, followed by *logistics* and *price*, while American customers are most concerned about *apps*, followed by *screen* and *appearance*.

2. Most popular aspect extraction and comparison

The ten most popular aspects for digital cameras, smartphones, and tablet computers are given in Tables 8b,

9b, and 10b respectively. We also use the entropy value to measure the differences between Chinese and American reviews regarding the ten most popular aspects, which can be calculated using Eq. 9:

$$E_{\text{pop}} = - \sum_{i=1}^{N_{\text{pop}}} (PA_i * \log PA_i) \quad (9)$$

where PA_i means PA value of the i th aspect, which can be calculated using Eq. 7, and N_{pop} denotes the number of popular aspects. Table 12 shows the entropy values of Chinese and American customers regarding popular aspects. From Table 12, it can be seen that Chinese customers' concerns about popular aspects are less homogeneous than those of American customers in all domains. Hence, we can conclude that concerns about popular aspects obviously differ between Chinese and American customers in all three domains, and the differences are much larger than those relating to frequent aspects. From

Table 10 Top aspects of tablet computers

| Chinese reviews | | English reviews | | Chinese reviews | | English reviews | |
|---------------------------------|-----|-----------------|------|--------------------------------|---------|-----------------|--------|
| Aspects | FA | Aspects | FA | Aspects | PA | Aspects | PA |
| <i>(a) Top frequent aspects</i> | | | | <i>(b) Top popular aspects</i> | | | |
| Screen | 512 | App | 2385 | Service | 0.95556 | Quality | 0.9969 |
| Logistics | 305 | Screen | 1584 | Genuine product | 0.8385 | Size | 0.9654 |
| Price | 145 | Size | 1473 | Battery | 0.8356 | Service | 0.9466 |
| Genuine product | 130 | Price | 832 | Appearance | 0.8305 | Resolution | 0.945 |
| Resolution | 85 | Battery | 648 | Button | 0.7714 | Price | 0.9447 |
| Quality | 85 | Quality | 636 | Resolution | 0.7647 | Battery | 0.9414 |
| Battery | 73 | Resolution | 400 | Quality | 0.7177 | Button | 0.9317 |
| Appearance | 59 | Processor | 282 | Price | 0.7103 | App | 0.8973 |
| Service | 45 | Button | 278 | Logistics | 0.7049 | Screen | 0.8813 |
| Button | 35 | Service | 262 | Screen | 0.6191 | Processor | 0.8511 |

Table 11 Entropy values about frequent aspects

| | Digital cameras | Smart phones | Tablet computers |
|---------|-----------------|--------------|------------------|
| Chinese | 0.8795 | 0.9436 | 0.8415 |
| English | 0.8495 | 0.8959 | 0.8832 |

Table 12 Entropy values about popular aspects

| | Digital cameras | Smart phones | Tablet computers |
|---------|-----------------|--------------|------------------|
| Chinese | 0.5549 | 0.7788 | 0.8361 |
| English | 0.4269 | 0.6045 | 0.2887 |

Table 8b, we can see that for digital camera reviews, the most popular aspect among Chinese customers is *genuine product*, followed by *logistics* and *price*, while American customers' most popular aspects are *quality*, *flashlight*, and *battery*. Chinese customers are most unsatisfied with the *service* aspect, while American customers are most dissatisfied with *shutter*. From Table 9b, we can see that for smartphone reviews, Chinese customers' most popular aspect is *genuine product*, followed by *price* and *logistics*, while American customers' most popular aspects are *quality*, *appearance*, and *camera*. Both Chinese and American customers are dissatisfied with *apps*, while Chinese customers are most unsatisfied with *battery* and American customers are most dissatisfied with *apps*. From Table 10b, we can see that for tablet computer reviews, Chinese customers' most popular aspect is *service*, followed by *genuine product* and *battery*, while American customers' most popular aspects are *quality*, *appearance* and *service*. Both Chinese and American customers are dissatisfied with *screen*, while Chinese customers are most dissatisfied with *screen* and American customers are most dissatisfied with *processor*.

The responses to the questionnaire survey show that the aspects of greatest concern to Chinese and American customers differ; thus, the aspects with which they are most satisfied or most dissatisfied are different. In addition, we can see that aspects that are of concern to customers in the two regions constitute a large proportion of the most frequent aspects, although their attitudes are not the same. For example, American customers are quite satisfied with *quality*, while Chinese customers are not. This suggests that manufacturers should strive to improve the quality of products in the Chinese market, especially in the tablet computer domain. For aspects with low satisfaction ratings in reviews from the two regions, enterprises need to pay more attention to customers' opinions in relation to their products. It is worth noting that Chinese customers show a high level of satisfaction with *price*, except in the tablet computer domain. This suggests that in the process of product promotion and marketing in China, price is not an important point; hence, inputs for promoting product price can be reduced appropriately. However, in the tablet computer domain, Chinese customers are not satisfied with *price*. This may be because the popularity of the tablet computer is not sufficiently widespread. Smartphones have almost become a necessity of life, while digital cameras are required by photography enthusiasts. Quality information about smartphones has become quite comprehensive; therefore, the price sensitivity is low [2], and most consumption of digital cameras is conspicuous consumption [5], which occurs when customers buy unnecessary or luxury products to display their financial resources and social status. Hence, consumers are willing to pay high prices for digital cameras. These two products are almost at two consumption extremes. The degree of quality information and level of conspicuous consumption for tablet computers lie somewhere in between these extremes, resulting in a low level of satisfaction regarding price.

3. Extracting frequent aspect that is concerned by both Chinese and American customers

Frequent aspects that are subject to concern by both Chinese and American customers are shown in the following figures. Figure 5 shows frequent aspects of digital camera reviews, and it can be seen that there are six aspects that both Chinese and American customers are concerned about. *Screen* is the aspect with the lowest satisfaction level in both Chinese and English reviews. Chinese customers are not satisfied with *battery*, although American customers show high satisfaction levels. Attitudes toward *price* are reversed, with Chinese customers being more satisfied than American customers.

Figure 6 shows frequent aspects of smartphone reviews. As can be seen, there are seven aspects that both Chinese and American customers are concerned about. Satisfaction with *apps* is low in both Chinese and English reviews. Chinese customers are not satisfied with *battery*, while

American customers show a high level of satisfaction. Again, their attitudes toward *price* are reversed, with Chinese customers being more satisfied than American customers.

Figure 7 shows frequent aspects of tablet computer reviews. As can be seen, there are eight aspects that both Chinese and American customers are concerned about. *Screen* is the aspect with the lowest satisfaction levels in both Chinese and English reviews. Chinese customers are not satisfied with *quality*, while American customers show a high level of satisfaction. Their attitudes toward *service* are reversed, with Chinese customers being more satisfied than American customers.

Customer Behavior Comparison Results

By analyzing the responses to the questionnaire survey, we are able to compare customer behavior in the two regions.

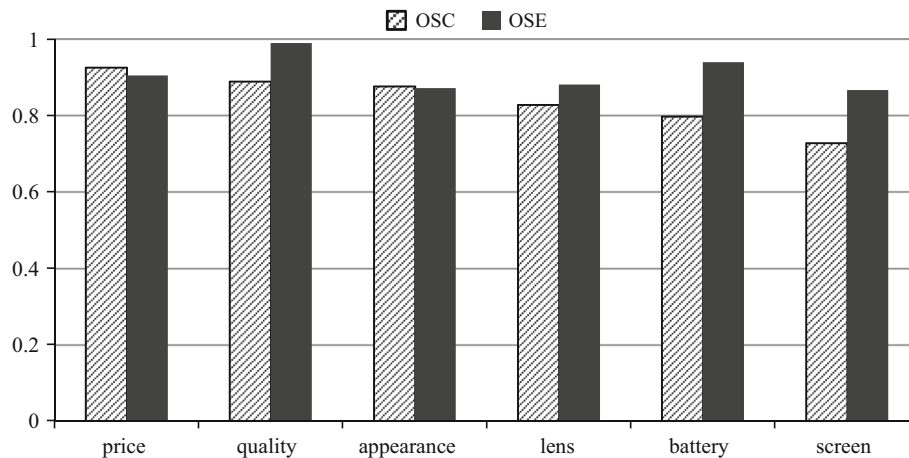


Fig. 5 PA values of frequent aspects about digital cameras

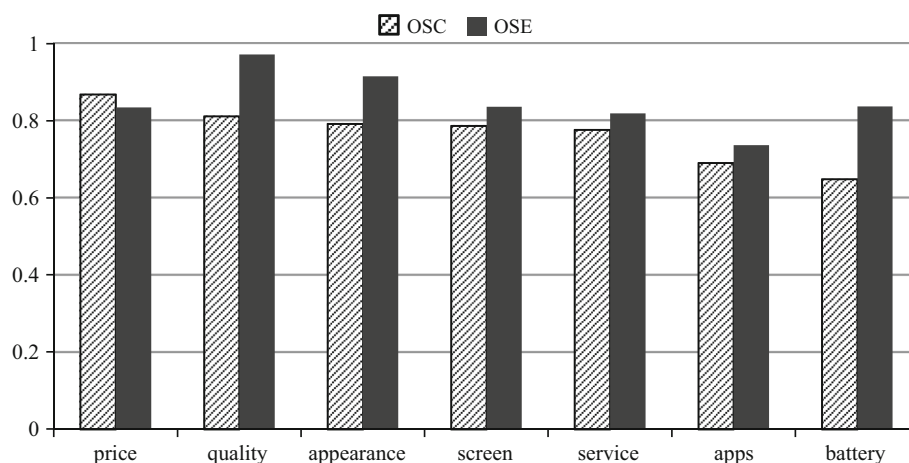


Fig. 6 PA values of frequent aspects about smartphones

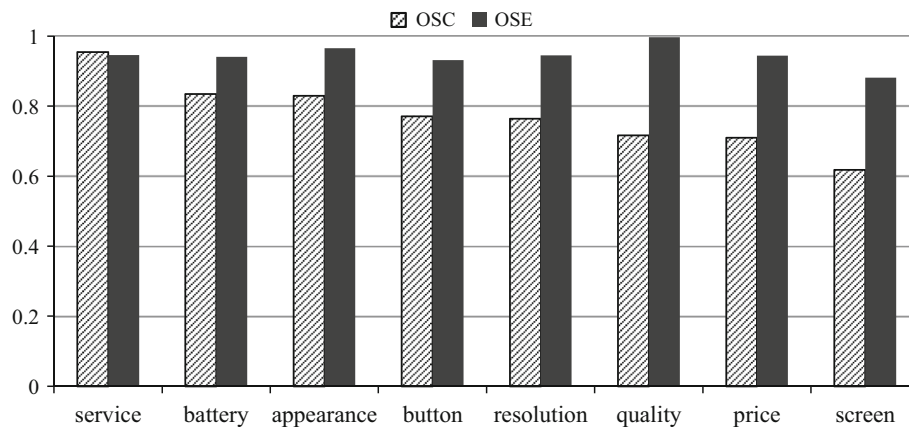


Fig. 7 PA values of frequent aspects about tablet computers

Corresponding with the questionnaire structure, there are three areas of customer behavior comparison: overall sentiment, brand preference, and purchase decision factors.

1. Customer behavior comparison on sentiment expression

The results presented in “[Overall Sentiment](#)” section show that in relation to both overall satisfaction and product-level satisfaction, the OS values of Chinese customers are higher than those of American customers in relation to most conditions; hence, we demonstrated that Chinese customers are more likely to express positive feelings. From the sociological point of view, Chinese people may be influenced by traditional culture, i.e., the *Doctrine of the Mean* [22]. This is an important part of Confucian philosophy and involves an attitude of compromise, rather than leaning toward either side. The Doctrine of the Mean has been internalized by Chinese people for thousands of years and is considered part of the national character. Hence, Chinese reviews include a wide variety of euphemisms, and they are accustomed to providing positive evaluations [58]. In contrast, Western culture encourages a more direct form of expression by American customers.

2. Customer behavior comparison on brand preference

The results presented in “[Brand Preference](#)” section show that American customers have more obvious BPs than Chinese customers. For example, because Samsung and Apple are both high-end smartphone brands, Chinese customers are not particularly concerned about specific aspects and have no obvious BP; however, American customers prefer some aspects of Samsung smartphones. This may be related to differences in cultures or customers’ experiences in different countries or regions [32, 50]. Based on our results, we can deduce that American customers pay more attention to details than Chinese customers in relation to IT products.

3. Customer behavior comparison on purchase decision factors

The results presented in “[Purchase Decision Factors](#)” section show that Chinese customers are concerned about the external features of products, such as the *logistics* and *package*, while American customers are concerned about product performance, including aspects such as the *lens* and *quality*. From the sociological point of view, as a group, Chinese customers are more concerned about the external features of a product, while American customers pay more attention to the internal features. It is worth noting that *genuine product* is a frequently mentioned aspect in Chinese reviews. This suggests that the Chinese market should be subject to tight supervision regarding product quality to enhance customers’ purchasing confidence. This supervision should come from both government and the manufacturing enterprises themselves. Additionally, multinational marketing needs to be guided in the light of its general trend, and different market conditions and customer behaviors should be taken into consideration.

A Case for Smartphone Promotion

We now present an imaginary case using online reviews from China and America. Suppose that Peter is a staff member of a smartphone marketing company and wants to promote smartphones in China and America. As customers from different regions may have different preferences and concerns about different aspects regarding smartphones, it is necessary for product marketing staff to understand customer behavior in different regions. As a senior marketing staff member, Peter has received a brief report about the differences between Chinese and American customers, such as that shown in Table 13. Based on this report, Peter

Table 13 Example report of differences between Chinese and American customers

| (a) Overall comparison | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|------------------------|------------------|--|--------------------|------------|-----------|--------|-----------------|--------|---------|--------|------------|--------|-------|-------|---------|-------|---------|-------|------|-------|--------|-------|--------|-------|--------|------------|------|--------|--------|--------|---------|--------|------------|-------|-------|-------|--------|-------|---------|-------|---------|-------|---------|-------|--------|-------|
| Comparisons | | Chinese customers | American customers | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Sentiment expression | Euphemistic | Purchase decision influencing factors | Direct | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Brand preference | | Obvious | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | | <div><div><div>Chinese top aspects</div><table><thead><tr><th>Aspect</th><th>Percentage</th></tr></thead><tbody><tr><td>logistics</td><td>18.81%</td></tr><tr><td>genuine product</td><td>17.97%</td></tr><tr><td>service</td><td>12.23%</td></tr><tr><td>appearance</td><td>11.91%</td></tr><tr><td>price</td><td>9.68%</td></tr><tr><td>battery</td><td>9.31%</td></tr><tr><td>quality</td><td>6.78%</td></tr><tr><td>apps</td><td>6.09%</td></tr><tr><td>screen</td><td>3.94%</td></tr><tr><td>memory</td><td>3.29%</td></tr></tbody></table></div><div><div>American top aspects</div><table><thead><tr><th>Aspect</th><th>Percentage</th></tr></thead><tbody><tr><td>apps</td><td>31.70%</td></tr><tr><td>screen</td><td>14.68%</td></tr><tr><td>battery</td><td>11.28%</td></tr><tr><td>appearance</td><td>9.33%</td></tr><tr><td>price</td><td>7.96%</td></tr><tr><td>camera</td><td>6.95%</td></tr><tr><td>service</td><td>5.34%</td></tr><tr><td>quality</td><td>5.24%</td></tr><tr><td>charger</td><td>4.65%</td></tr><tr><td>button</td><td>3.77%</td></tr></tbody></table></div></div> | Aspect | Percentage | logistics | 18.81% | genuine product | 17.97% | service | 12.23% | appearance | 11.91% | price | 9.68% | battery | 9.31% | quality | 6.78% | apps | 6.09% | screen | 3.94% | memory | 3.29% | Aspect | Percentage | apps | 31.70% | screen | 14.68% | battery | 11.28% | appearance | 9.33% | price | 7.96% | camera | 6.95% | service | 5.34% | quality | 5.24% | charger | 4.65% | button | 3.77% |
| Aspect | Percentage | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| logistics | 18.81% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| genuine product | 17.97% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| service | 12.23% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| appearance | 11.91% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| price | 9.68% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| battery | 9.31% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| quality | 6.78% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| apps | 6.09% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| screen | 3.94% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| memory | 3.29% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Aspect | Percentage | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| apps | 31.70% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| screen | 14.68% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| battery | 11.28% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| appearance | 9.33% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| price | 7.96% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| camera | 6.95% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| service | 5.34% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| quality | 5.24% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| charger | 4.65% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| button | 3.77% | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

should create different marketing programs for customers in different regions.

First, regarding the matter of sentiment expression, Peter has taken the euphemistic expressions of Chinese customers into consideration. Therefore, he has raised the threshold for Chinese satisfaction that he obtained from test marketing. Second, Peter is promoting an unpopular smartphone brand. When he promotes this smartphone to American customers, he focuses on contrasting this brand with popular brands and emphasizing the advantages of this brand. Third, Chinese and American customers have different views in terms of the aspects that are of most concern, so Peter emphasizes different aspects to different customers. When Peter promotes the smartphone to Chinese customers, he allocates most resources, including people, finance, and time, toward promoting the smartphone's external features. He chooses good logistics companies to ensure reliable, speedy delivery, and good e-commerce Web sites to ensure that the most important qualities of the smartphone are emphasized. When Peter promotes the smartphone to American customers, most of the resources are allocated to ensuring that the smartphone is of high quality. He emphasizes the need for software compatibility, high-definition screens, and long-lasting batteries.

Marketing based on the above analyses may improve customer satisfaction, which can in turn improve sales volume. Moreover, because of the targeted marketing programs, marketing costs are likely to be greatly reduced.

Conclusion and Future Work

In this paper, we present a framework for developing an automatic questionnaire related to online customer behavior to explore differences between the online shopping behavior of Chinese and American customers. We generate questions and extract answers automatically via multi-granularity opinion mining. Three important conclusions can be drawn from our analysis:

1. Chinese customers obey the Doctrine of the Mean and often use euphemistic expressions, while American customers use more direct expressions
2. American customers pay more attention to product details than Chinese customers
3. Chinese customers are more concerned about external features, while American customers are more concerned about internal features.

We investigate differences between Chinese and American online shopping behavior using reviews posted on e-commerce Web sites. This method can be extended to investigate differences between Chinese and American

consumers on other Web sites and differences between users from other countries or regions. In a word, this is a study about exploring differences in customers behavior by mining online reviews, which provide some reference values for research on features including user cognition and sentiment analysis.

The following future work is planned. First, this was our first attempt to generate questionnaires automatically, and only simple questions based on templates were generated. Therefore, we will conduct further research into areas such as question and answer technologies. Second, reviews from greater numbers of Web sites and regions will be compared and analyzed to determine whether differences in online shopping behaviors exist among customers who speak different languages and/or are from different continents. Third, we will conduct further research into sentiment analysis technologies to obtain more comprehensive sentiment results, and additional measurement techniques will be used to quantify differences in online shopping behaviors. Finally, these research topics can be expanded into other domains, such as catering, tourism, and entertainment.

Acknowledgments This work is supported by the Major Projects of National Social Science Fund (No. 13&ZD174), the National Social Science Fund Project (No. 14BTQ033), the Natural Science Foundation of China (Nos. 61305090, 61272233), and the Opening Foundation of Alibaba Research Center for Complex Sciences, Hangzhou Normal University (No. PD12001003002003). We thank the reviewers for the insightful comments.

Compliance with Ethical Standards

Ethical Standard All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Conflicts of Interest Qingqing Zhou, Rui Xia and Chengzhi Zhang declare that they have no conflict of interest.

Informed Consent All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2008 (5). Additional informed consent was obtained from all patients for which identifying information is included in this article.

Human and Animal Rights This article does not contain any studies with animals performed by any of the authors.

References

- Agarwal B, Poria S, Mittal N, Gelbukh A, Hussain A. Concept-level sentiment analysis with dependency-based semantic parsing: a novel approach. *Cogn Comput*. 2015;7(4):487–99.
- Bagwell LS, Bernheim BD. Veblen effects in a theory of conspicuous consumption. *Am Econ Rev*. 1996;86(3):349–73.
- Blair-Goldensohn S, Hannan K, McDonald R, Neylon T, Reis GA, Reynar J. Building a sentiment summarizer for local service reviews. In: WWW workshop on NLP in the information explosion era, 2008;14.
- Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. *J Mach Learn Res*. 2003;3:993–1022.
- Boulding W, Lee E, Staelin R. Mastering the mix: Do advertising, promotion, and sales force activities lead to differentiation? *J Mark Res*. 1994;159–72.
- Brown JC, Frishkoff GA, Eskenazi M. Automatic question generation for vocabulary assessment. In: Proceedings of the conference on human language technology and empirical methods in natural language processing. Association for Computational Linguistics; 2005. p. 819–26.
- Cambria E, Hussain A, Havasi C, Eckl C. SenticSpace: visualizing opinions and sentiments in a multi-dimensional vector space. In: Knowledge-based and intelligent information and engineering systems. Berlin: Springer; 2010. p. 385–93.
- Cambria E, Schuller B, Liu B, Wang H, Havasi C. Statistical approaches to concept-level sentiment analysis. *IEEE Intell Syst*. 2013;28(3):6–9.
- Chatterjee P. Online reviews: do consumers use them? *Adv Consum Res*. 2001;28:129–33.
- Chikersal P, Poria S, Cambria E, Gelbukh A, Siong CE. Modelling public sentiment in twitter: Using linguistic patterns to enhance supervised learning. In: Computational linguistics and intelligent text processing. Berlin: Springer; 2015. p. 49–65.
- Cortes C, Vapnik V. Support-vector networks. *Mach Learn*. 1995;20(3):273–97.
- Das D, Bandyopadhyay S. Sentence-level emotion and valence tagging. *Cogn Comput*. 2012;4(4):420–35.
- Davidov D, Rappoport A. Unsupervised discovery of generic relationships using pattern clusters and its evaluation by automatically generated sat analogy questions. In: ACL. Citeseer; 2008. p. 692–700.
- Denecke K. Using sentiwordnet for multilingual sentiment analysis. In: Data engineering workshop, 2008. ICDEW 2008. IEEE 24th international conference on. IEEE; 2008. p. 507–12.
- Ding X, Liu B, Yu PS. A holistic lexicon-based approach to opinion mining. In: Proceedings of the 2008 international conference on web search and data mining. ACM; 2008. p. 231–40.
- Gangemi A, Presutti V, Reforgiato Recupero D. Frame-based detection of opinion holders and topics: a model and a tool. *Comput Intell Mag IEEE*. 2014;9(1):20–30.
- Hoffman M, Bach FR, Blei DM. Online learning for latent dirichlet allocation. In: Advances in neural information processing systems. 2010; p. 856–64.
- Hwang W, Jung HS, Salvendy G. Internationalisation of e-commerce: a comparison of online shopping preferences among Korean, Turkish and US populations. *Behav Inf Technol*. 2006;25(1):3–18.
- Jo Y, Oh AH. Aspect and sentiment unification model for online review analysis. In: Proceedings of the fourth ACM international conference on Web search and data mining. ACM; 2011. p. 815–24.
- Lancaster HO, Seneta E. Chi-square distribution. New York: Wiley Online Library; 2005.
- Lee J, Park DH, Han I. The effect of negative online consumer reviews on product attitude: an information processing view. *Electron Commer Res Appl*. 2008;7(3):341–52.
- Legge J, et al. The confucian analects, the great learning & the doctrine of the mean. Cosimo, Inc. 2009.
- Li S, Lee SYM, Chen Y, Huang CR, Zhou G. Sentiment classification and polarity shifting. In: Proceedings of the 23rd international conference on computational linguistics. Association for Computational Linguistics 2010. p. 635–43.

24. Li X, Xie H, Chen L, Wang J, Deng X. News impact on stock price return via sentiment analysis. *Knowl Based Syst.* 2014;69:14–23.
25. Lin C, He Y. Joint sentiment/topic model for sentiment analysis. In: *Proceedings of the 18th ACM conference on information and knowledge management.* ACM; 2009. p. 375–384.
26. Lin CY. Automatic question generation from queries. In: *Workshop on the question generation shared task*, 2008; p. 156–64.
27. Lin J, Demner-Fushman D. Methods for automatically evaluating answers to complex questions. *Inf Retr.* 2006;9(5):565–87.
28. Liu B. *Sentiment analysis and opinion mining: synthesis lectures on human language technologies.* San Rafael: Morgan & Claypool Publishers; 2012.
29. Liu M, Calvo RA, Rus V. Automatic question generation for literature review writing support. In: *Intelligent tutoring systems.* Berlin: Springer; 2010. p. 45–54.
30. Maas AL, Daly RE, Pham PT, Huang D, Ng AY, Potts C. Learning word vectors for sentiment analysis. In: *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies-volume 1.* Association for Computational Linguistics; 2011. p. 142–150.
31. Mahmood MA, Bagchi K, Ford TC. On-line shopping behavior: cross-country empirical research. *Int J Electron Commer.* 2004;9(1):9–30.
32. McWilliam G. Low involvement brands: is the brand manager to blame? *Mark Intell Plann.* 1997;15(2):60–70.
33. Moghaddam S, Ester M. On the design of lda models for aspect-based opinion mining. In: *Proceedings of the 21st ACM international conference on information and knowledge management.* ACM; 2012. p. 803–12.
34. Mullen T, Collier N. Sentiment analysis using support vector machines with diverse information sources. In: *EMNLP*, vol 4. 2004; p. 412–8.
35. MY W. Comparing consumers' on-line shopping behaviors in taiwan and the united states. In: *Proceedings of the annual meeting of the international communication association* 2010.
36. Myller N. Automatic generation of prediction questions during program visualization. *Electron Notes Theor Comput Sci.* 2007;178:43–9.
37. Nakagawa T, Inui K, Kurohashi S. Dependency tree-based sentiment classification using crfs with hidden variables. In: *Human language technologies: the 2010 annual conference of the North American Chapter of the Association for Computational Linguistics.* Association for Computational Linguistics 2010. p. 786–94.
38. Pang B, Lee L, Vaithyanathan S. Thumbs up?: sentiment classification using machine learning techniques. In: *Proceedings of the ACL-02 conference on empirical methods in natural language processing-volume 10.* Association for Computational Linguistics 2002. p. 79–86.
39. Papasalouros A, Kanaris K, Kotis K. Automatic generation of multiple choice questions from domain ontologies. In: *e-Learning.* Citeseer 2008. p. 427–34.
40. Park DH, Kim S. The effects of consumer knowledge on message processing of electronic word-of-mouth via online consumer reviews. *Electron Commer Res Appl.* 2009;7(4):399–410.
41. Park DH, Lee J, Han I. The effect of on-line consumer reviews on consumer purchasing intention: the moderating role of involvement. *Int J Electron Commer.* 2007;11(4):125–48.
42. Popescu AM, Etzioni O. Extracting product features and opinions from reviews. In: *Natural language processing and text mining.* Berlin: Springer; 2007. p. 9–28.
43. Rill S, Reinelt D, Scheidt J, Zicari RV. Politwi: early detection of emerging political topics on twitter and the impact on concept-level sentiment analysis. *Knowl Based Syst.* 2014;69:24–33.
44. Salton G, Yu CT. On the construction of effective vocabularies for information retrieval. In: *ACM SIGPLAN notices*, vol 10. ACM; 1973. p. 48–60.
45. Sumita E, Sugaya F, Yamamoto S. Measuring non-native speakers' proficiency of english by using a test with automatically-generated fill-in-the-blank questions. In: *Proceedings of the second workshop on building educational applications using NLP.* Association for Computational Linguistics 2005. p. 61–8.
46. Taboada M, Brooke J, Tofiloski M, Voll K, Stede M. Lexicon-based methods for sentiment analysis. *Comput Linguist.* 2011;37(2):267–307.
47. Thet TT, Na JC, Khoo CS. Aspect-based sentiment analysis of movie reviews on discussion boards. *J Inf Sci.* 2010;0165551510388123.
48. Turney PD. 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.* Association for Computational Linguistics 2002. p. 417–24.
49. Vermeulen IE, Seegers D. Tried and tested: the impact of online hotel reviews on consumer consideration. *Tour Manag.* 2009;30(1):123–7.
50. Warrington P, Shim S. An empirical investigation of the relationship between product involvement and brand commitment. *Psychol Mark.* 2000;17(9):761–82.
51. Wei W, Gulla JA. Sentiment learning on product reviews via sentiment ontology tree. In: *Proceedings of the 48th annual meeting of the association for computational linguistics.* Association for Computational Linguistics 2010. p. 404–13.
52. Wei W, Gulla JA, Fu Z. Enhancing negation-aware sentiment classification on product reviews via multi-unigram feature generation. In: *Advanced intelligent computing theories and applications.* Berlin: Springer; 2010. p. 380–91.
53. Xia R, Wang T, Hu X, Li S, Zong C. Dual training and dual prediction for polarity classification. In: *ACL*, vol 2. 2013; p. 521–5.
54. Xia R, Zong C. Exploring the use of word relation features for sentiment classification. In: *Proceedings of the 23rd international conference on computational linguistics: posters.* Association for Computational Linguistics 2010. p. 1336–44.
55. Xia R, Zong C, Hu X, Cambria E. Feature ensemble plus sample selection: domain adaptation for sentiment classification. *Intell Syst IEEE.* 2013;28(3):10–8.
56. Xia R, Zong C, Li S. Ensemble of feature sets and classification algorithms for sentiment classification. *Inf Sci.* 2011;181(6):1138–52.
57. Xia Y, Cambria E, Hussain A, Zhao H. Word polarity disambiguation using bayesian model and opinion-level features. *Cogn Comput.* 2015;7(3):369–80.
58. Yau OH. Chinese cultural values: their dimensions and marketing implications. *Eur J Mark.* 1988;22(5):44–57.
59. Yessenalina A, Yue Y, Cardie C. Multi-level structured models for document-level sentiment classification. In: *Proceedings of the 2010 conference on empirical methods in natural language processing.* Association for Computational Linguistics 2010. p. 1046–56.
60. Yu J, Zha ZJ, Wang M, Chua TS. Aspect ranking: identifying important product aspects from online consumer reviews. In: *Proceedings of the 49th annual meeting of the association for computational linguistics: human language technologies, vol 1, HLT '11.* Association for Computational Linguistics, Stroudsburg, PA, USA 2011. p. 1496–505.
61. Zhao WX, Jiang J, Yan H, Li X. Jointly modeling aspects and opinions with a maxent-lda hybrid. In: *Proceedings of the 2010 conference on empirical methods in natural language processing.* Association for Computational Linguistics 2010. p. 56–65.
62. Zhu F, Zhang X. Impact of online consumer reviews on sales: the moderating role of product and consumer characteristics. *J Mark.* 2010;74(2):133–48.