Association for Information Systems
AIS Electronic Library (AISeL)

WHICEB 2018 Proceedings

Wuhan International Conference on e-Business

Summer 6-30-2018

A Text Mining Based Approach for Mining Customer Attribute Data on Undefined Quality Problem

Qing Zhu
International Business School, Shaanxi Normal University, Xi'an, 710119, Chin

Yiqiong Wu
International Business School, Shaanxi Normal University, Xi'an, 710119, China

Yuze Li
Department of Mechanical and Industrial Engineering, University of Toronto, ON, Canada

Renxian Zuo
International Business School, Shaanxi Normal University, Xi'an, 710119, China

A Text Mining Based Approach for Mining Customer Attribute Data on Undefined Quality Problem

Qing Zhu[1*], *Yiqiong Wu*[1], *Yuze Li*[2], *Renxian Zuo*[1]
[1]International Business School, Shaanxi Normal University, Xi'an, 710119, China
[2]Department of Mechanical and Industrial Engineering, University of Toronto, ON, Canada

Abstract: Understanding how the consumer perceives quality is a key issue in supply chain management. However, as the market structure continues to deepen, traditional evaluation methods using SEVRQUAL are unable identify all issues related to customer quality and unable to supply solutions. The maturation of data mining technology, however, has opened the possibilities of mining customer attribute data on quality problems from unstructured data. Based on the consumer perspective, this research uses an unsupervised machine learning text mining approach and the Recursive Neural Tensor Network to resolve the attribution process for undefined quality problems. It was found that the consumer quality perception system has a typical line-of-sight that can assist consumers quickly capture the logical structure of the quality problem. Although attributions related to quality problems are very scattered, a highly unified view was found to exist within each group, and a strategy to solve the undefined quality problem was agreed through group consensus by 61% of the consumers.

Keywords: text mining, supply chain management, quality control

1. INTRODUCTION

Jacoby, Olson and Haddock [1] identified consumer perceived quality as an important criterion for consumer feedback on the quality of goods. Based on Jacoby, Olson and Haddock's initial definition, Grönroos [2] extended the interpretation, as follows: when the quality of the goods are on the same level, due to impact factors such as previous experience and commodity characteristics, consumers develop different perceptions about the quality of the goods. As consumer perceptions of quality are affected by mood and trust, this could affect consumer satisfaction and loyalty [3], [4]. Lee and Lin [5] then extended this concept to marketing strategy, proposing that as consumer satisfaction influences consumer purchase intention, controlling the consumer's perception of good quality is an important part of enterprise marketing strategies [6].

At the same time, Folkes and Kotsos [7] claimed that if managers wanted to increase transaction success, they needed to minimize the differences between the consumer and the seller, and therefore supply chain managers needed to be able to understand product quality from the consumer's perspective [8] and be able to analyze consumer thinking to illuminate the shortcomings in their products and services, reduce potential company losses, and enhance profit margins [9], all of which would result in more effective, targeted supply chain management.

In support of Jacoby's idea of consumer perceived quality, Parasuraman, Zeithaml, and Berry sought to develop an appropriate consumer attribution management tool and proposed the SERVQUAL evaluation method [10], with the understanding that service quality depended on the gap between the consumer's expectations and the perceived quality. They divided service quality into 10 dimensions: tangibles, reliability, responsiveness, reliability, communication, politeness, security, understanding, and accessibility. In subsequent studies, service quality was integrated into 5 dimensions (22 items); materiality, reliability, responsiveness, assurance and empathy, after which this evaluation method was widely used in many fields to measure perceived consumer quality. For example, Marek and Nowacki [11] used these methods to evaluate the tourism quality at the Rogalin

* Corresponding author. Email: zhuqing@snnu.edu.cn

Museum, from which they were able to identify its competitive advantages and disadvantages and offer comprehensive guidance for managing perceived customer quality.

While it has been widely used, the reliability of the SERVQUAL method has been questioned due to its subjectivity, robustness, and variability. A series of revised models on perceived quality evaluations were then developed. Cronin and Taylor [12], for instance, developed the SERVPERF method, in which consumer service expectations and consumer service perceptions were two separate measures. Peter, Churchill, and Brown [13] proposed a non-difference valuation method as they felt that the SERVQUAL method did not account for previous service experience when measuring consumer service expectations, thus weakening the validity of the difference evaluation method.

While the validity of traditional quality perception measurements has been proven and part of the SERVQUAL application can pass the Kaiser-Meyer-Olkin (KMO) and Bartlett tests of sphericity, mainstream measurement methods based on SERVQUAL all tend to formulate project scores from the supply chain management's perspective and fail to directly address consumer concerns or measure consumer expectations. Therefore, the information measured using traditional quality measurement methods is limited both in perspective and effectiveness.

As supply chain managers understand the complete supply chain operation, they consider all quality control links in the supply chain [14], [15]. However, the information asymmetry between consumers and management [16] and the lack of consumer information about the nature of goods during production and sales have resulted in significant differences between consumers and management in terms of the causes of the quality problem attributions [17]. Further, the SERVQUAL tool designs amplify such differences. When quality problems occur, consumers usually attribute the problem to the perceivable end of the supply chain, and do not perceive the overall supply chain. The main cause for the inability of traditional methods such as SERVQUAL to adapt is a difference in perspective: that is, understanding the quality perception of the consumer from the management's perspective can only amplify the differences, which further highlights the invalidity of the tool.

Since 2012, non-structured big data processing technology such as text mining has begun to mature and be applied in areas such as quantitative strategy, market segmentation, prediction and group behavior intervention, and other fields. Therefore, there are new methods now available for measuring the consumer quality perspective. Based on Folkes and Kotsos, this paper uses an unsupervised machine learning algorithm and a recursive neural network method to identify consumer quality perceptions from unstructured data, analyze consumer responses about quality problems, and describe consumer expectations and behavior, while confirming that mainstream measurement methods such as SERVQUAL are not optimal.

2. METHODOLOGY

2.1 Text mining

Text mining is a special form of data mining. It can discover and extract implicit valuable information from massive amount of unstructured data, and form knowledge that is easy for users to understand. The implementation of text mining is divided into two steps: text preprocessing and knowledge extraction. Text preprocessing transforms unstructured text into term-document matrix. Knowledge extraction derives facts and knowledge from term-document matrix. Based on different purposes, the task of text miming can be broadly divided into five categories: text classification, text clustering, association rule mining, automatic summarization and topic detection.

2.2 Text preprocessing

In this study, a series of cleaning and feature representation of the text data are carried out by means of NLP and TM packages in R.

2.2.1 Characteristics representation

This study used a Vector Space Model (VSM) to represent the text [18], the fundamental principle for which was to assign different weights to each word, thereby allowing the characteristics vector to be represented as a weighted text, as follows: for text set, a particular text $T = \{t_i\}_{i=1}^{n}$, a particular text $t_i(i \leq i \leq n)$ can be represented as $t_i = w_{i1}, w_{i2}, \dots, w_{im}$. In which, m is the number of characteristics, w_{ij} is the weight of j*th* characteristic in text t_i.

A classic weight assignment method when constructing a vector space model is the TD-IDF method developed by Salton and Buckley [19], which is able to calculate the importance of a particular word in a text and therefore has more accurate representation and clustering results. The formula for the TD-IDF weight assignment method is as follows:

$$W_j = TF_j * IDF_j \tag{1}$$

in which,

$$IDF_j = log\ (\frac{n}{DF_j}) \tag{2}$$

where W_j is the weight of the j*th* characteristic, TF_j is the frequency of the j*th* characteristic in the current text, and DF_j is the frequency of the j*th* characteristic in the overall text set. In a real application, to avoid too broad a variable value range, this study normalized the vector so that the average was 0 and the square difference was 1.

2.2.2 Similarity analysis between texts

Before conducting text clustering analysis, we need to measure the degree of similarity and difference between texts. This document uses hierarchical, k-means and spectral clustering for text clustering.

When conducting text hierarchical clustering, the distance between clusters is optimized by the Ward method. The Ward method requires the degree of separation to be calculated by the Euclidean distance. The Euclidean distance between i*th* text and j*th* text is calculated using the following formula:

$$d(i,j) = \sqrt{\sum_{k=1}^{m}(w_{ik} - w_{jk})^2} \tag{3}$$

Subject to,

$$d(i,i) = 0$$
$$d(i,j) = d(j,i)$$

Dhillon and Modha [20] claimed that the cosine distance was superior to the Euclidean distance for measuring text clustering similarity. Therefore, k-means clustering, spectral clustering, and cosine similarity were used to assess document similarities and obtain the document similarity matrix. The cosine similarity between the i*th* and the j*th* documents was determined using the following formula:

$$sim(d_i, d_j) = \frac{d_i \cdot d_j}{|d_i| \times |d_j|} = \frac{\sum_{k=1}^{m} w_{ik} \cdot w_{jk}}{(\sum_{k=1}^{m} w_{ik}^2)(\sum_{k=1}^{m} w_{jk}^2)} \tag{4}$$

Subject to,

$$sim(d_i, d_j) = 1$$
$$sim(d_i, d_j) = sim(d_j, d_i)$$

2.3 Text clustering

Text clustering is unsupervised learning that involves aggregating massive text data into several classes without prior knowledge or assumptions, thereby ensuring as high a similarity of text data as possible and as low a similarity as possible across the classes. As there is no precise definition for clustering, the clustering algorithm varies with the results.

2.3.1 Hierarchical cluster

Hierarchical clustering has been a common clustering method. In clustering analysis, the basic principle is to select the two classes with the highest similarity aggregation of all the classes. This step is repeated until all data are grouped into a class. Compared with other clustering methods, hierarchical clustering can be applied to arbitrary shapes and attribute data set types [21]; however, the time complexity of the algorithm is relatively high and therefore not suitable for clustering massive amounts of text data [20]. From the notion of basic hierarchical clustering, Rohlf proposed an MST-algorithm based on the minimum spanning tree that was able to optimize hierarchical clustering performances [22]. In this paper, an optimized hierarchical clustering algorithm was adopted, the algorithm for which was as follows:

1. With the known text set $T = \{t_i\}_{i=1}^{n}$ and the differences between documents, *d*;
2. Let every text be a cluster, and then initialize the output table such that $out \leftarrow [\,]$;
3. For any $x \in T$, $size(x) \leftarrow 1$;
4. Suppose $arg\,min\,(d(t_x, t_y))$ is (t_a, t_b), then combine t_a, t_b into a new cluster $(t_a \cup t_b)$:

$$out \leftarrow out + (t_a, t_b, (t_a \cup t_b)) \tag{5}$$

$$T \leftarrow \complement_T t_a \cup \complement_T t_b \cup (t_a \cup t_b) \tag{6}$$

5. Use the Ward method [23] to update the inter-cluster distance:

$$d(t_a \cup t_b, t_k) = \sqrt{\frac{(n_{t_a}+n_{t_k})d(t_a,t_k)+(n_{t_b}+n_{t_k})d(t_b,t_k)-n_{t_k}d(t_a,t_b)}{n_{t_a}+n_{t_b}+n_{t_k}}} \tag{7}$$

6. $size(t_a \cup t_b) \leftarrow size(t_a) + size(t_b)$
7. Repeat steps 4 to 6 until a cluster of size *n* is obtained

2.3.2 K-means clustering

K-means clustering is a common clustering method based on centroids [21] that has a lower time complexity and a higher computational efficiency; however, the algorithm is not suitable for non-convex data, does not have robustness, is more sensitive to outliers, and can easily fall into a local optima. Therefore, the clustering results are more susceptible to the influence of the number of predefined clusters [21]. Pelleg and Moore [24] proposed an X-Means algorithm that could automatically determine the number of K clusters using optimization. Therefore, this study used the split level algorithm, the steps for which were as follows:

1. With a known text set $T = \{t_i\}_{i=1}^{n}$;
2. Initialize the number of K clusters;
3. Randomly select the clustering centroid $C = \{c_k\}_{i=1}^{K}$;
4. Cluster the text objects into the nearest-located cluster and obtain K classes $\{W_k\}_{i=1}^{K}$, as defined in the equation (8):

$$W_k = \{t_i \in T | k = arg\,min_{j=1,\dots,K} \|t_i - c_j\|\} \tag{8}$$

Subject to, $W_k \subset T, W_l \cap W_q = \emptyset, \cup_{k=1}^{K} W_k = T$;

5. Use equation (9) to update each centroid

$$c_k = \frac{1}{|W_k|} \sum_{t_i \in W_k} t_i \tag{9}$$

6. Repeat step 4 and 5 until a stable clustering result is obtained.

The algorithm optimizes the clustering result using the iterative method in the following equation (10) to minimize the sum of square errors $E_T(C)$:

$$E_T(C) = \sum_{k=1}^{K} \sum_{t_i \in W_k} \|t_i - c_k\|^2 \tag{10}$$

2.3.3 Spectral clustering

The essence of the spectral clustering algorithm is to use the eigenvector of the Laplace matrix. The relationships between the texts are used to build graph $G = (V, E, W)$ with n nodes, with the vertex $V = \{1, \dots, n\}$ representing each text. The edge $E \subseteq V \times V$ in the graph illustrates the relationships between texts, and the weight of the edges $W = (w_{ij})_{n\times n}$ shows the strength of the relationships between the texts. The goal of spectral clustering is to divide the graph model into a number of subgraphs and minimize segmentation losses [25]. The spectral clustering algorithm has the ability to converge the clustering results to a global optimum and is not sensitive to outliers. However, the spectral clustering time complexity and the number of clusters k needs to define in advance are high [21]. The algorithm is as follows:

1. Obtain a similarity matrix F_{ij} between the texts;
2. The Laplace matrix $L = D^{-1/2}FD^{-1/2}$ is constructed, in which D is the diagonal matrix of the diagonal elements $D_{ii} = \sum_{j=1}^{n} F_{ij}$;
3. The eigenvectors $s_1, s_2, \dots, s_k$ that correspond to the minimum eigenvalues of the first k of L are calculated, and the matrix $S = [s_1, s_2, \dots, s_k] \in R^{n\times k}$ is obtained;
4. Consider each line in the S as a point in space R^k, and use the k-means clustering algorithm to obtain k text clusters.

2.4 Sentiment analysis - recursive neural tensor network

As grammar rules are recursive, Socher, Perelygin, and Wu [26] combined them with a corresponding algorithm to fully analyze a text. Then, based on the existing recursive neural network (RNN) and matrix-vector recursive neural network (MV-RNN) models, they proposed a recursive neural tensor network (RNTN) for fine grained sentiment text classification. The fine grained sentiment classifications had 5 emotional levels; very negative, negative, neutral, positive, and very positive. For fine grained sentiment analysis, the algorithm increased the accuracy from 44.4% to 45.7%.

2.4.1 Neural network calculation process

In recursive neural models, the compositional vector representations for phrases of different length and syntactic type can be computed in a bottom up recursive fashion using different compositionality functions.

Based on the current RNN model, Socher, Perelygin, and Wu proposed a new model called the RNTN, which was able to compute a sentence tree with detailed emotional information using the recursive combination between words and phrases

Figure 1 gives an example of a three tensor layer neural network. The computation process for a single tensor layer is as follows:

1. Each word in the sentence is represented as a d-dimensional vector. All word vectors are initialized by random sampling each value from a uniform distribution *U(-0.0001,0.0001)*;
2. The output of a tensor product $h \in R^d$ is defined as:

$$h = \begin{bmatrix} b \\ c \end{bmatrix}^T V^{[1:d]} \begin{bmatrix} b \\ c \end{bmatrix} \tag{11}$$

where $V^{[1:d]}$ is the tensor that defines the multiple bilinear forms;

3. The first parent vector *p1* is computed:

$$p1 = f\left(\begin{bmatrix} b \\ c \end{bmatrix}^T V^{[1:d]} \begin{bmatrix} b \\ c \end{bmatrix} + W \begin{bmatrix} b \\ c \end{bmatrix}\right) \tag{12}$$

where W is the sentiment classification matrix;

4. The next parent vector *p2* in the tri-gram is computed

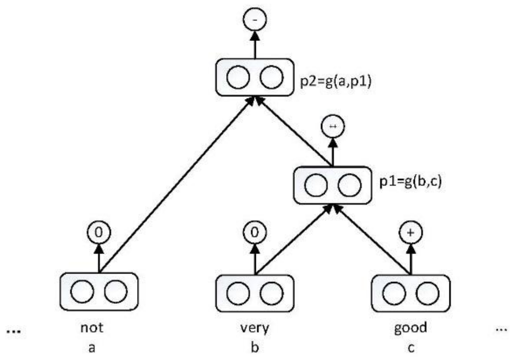


Figure 1. An example of a recursive neural tensor network

using the same weights:

$$p2 = f\left(\begin{bmatrix} a \\ p1 \end{bmatrix}^T V^{[1:d]} \begin{bmatrix} a \\ p1 \end{bmatrix} + W \begin{bmatrix} a \\ p1 \end{bmatrix}\right) \tag{13}$$

5. Steps 3 and 4 are repeated and each parent vector is computed in a bottom up fashion until the top parent vector is reached and the final sentiment orientation determined.

2.4.2 Model training

When the syntax tree is generated, the model trains a softmax classifier from top to bottom through the vector labels of each node. This semantic distribution relationship can be expressed as:

$$y^a = softmax(W_s a) \tag{14}$$

in which, $W_s \in R^{5\times d}$ is the emotional classification matrix, and a is the operation node for the classification.

2.5 Price decomposition model

From the sentiment classification prediction model of the text, the sentimental trends in each sentence are integrated into the sentiment fluctuation in which *t* represents the number of sentences in the current comment; however, the *t* values of each text are not equal. To identify similar trends in the sentiment fluctuations, a price decomposition model was applied to divide each comment into positive sentiment fluctuations E_{inc} and negative sentiment fluctuations E_{dec}. The price decomposition model, which was first proposed by Oscar in 1972[27], decomposes price into a rise and a fall and allows for the asymmetric effect of demand to be studied. For a comment containing *t* sentences, the specific decomposition formulas are as follows:

$$\begin{cases} E_{inc} = \sum_{i=1}^{t-1} max\,\{0, (E_{i+1} - E_i)\} \\ E_{dec} = \sum_{i=1}^{t-1} min\,\{0, (E_{i+1} - E_i)\} \end{cases} \tag{15}$$

3. DATA COLLECTION

As the object of this study was the general end consumer, it does not include the “industrial market”, “raw material market” or “intermediate manufactured goods market” as these could result in consumer ambiguity. Further, to ensure scientific questionnaire validity, a real online shopping situation was simulated that applied real evaluation rules. Study objects were required to provide comments on three evaluation categories; the quality of the goods, the logistics service, and consumer service attitudes (total simulation of an Alibaba shopping scenario) and used Likert scales ranging from -2 to 2 to represent their level of satisfaction with -2 being very unsatisfactory and 2 being very satisfactory.

Hoffman [28] claimed that education, income, gender, occupation, and other factors affected online consumer shopping behavior. However, with the increased popularity of network technology, demographic characteristics are expected to gradually decline. For example, Zellner [29] found that gender, income, and education levels did not contribute to online shopping differences, and Doolin, Dillon, and Thompson [30] also found no significant correlations between a consumer’s age and online shopping behavior. Therefore, as it has been repeatedly shown that demographic characteristics were less related to consumer online shopping behavior, it is reasonable to surmise that the experimental results were not biased.

This study conducted a questionnaire survey posted electronic questionnaire online. A total of 788 questionnaires were collected and 508 valid samples obtained. After data collection was completed, the score items and text comments were separated and stored, and were then read separately into the software for analysis.

4. DATA ANALYSIS

The data analysis was divided as follows: (1) an analysis of the perceived quality measurements based on the SDERQUAL structure and the scores for the quality of the goods, the logistics service, and the customer service attitude; (2) a description of the consumer perception system and the problem detection expressions; (3) presentation and measurement of consumer attributions; (4) analyses of consumer strategies and expected actions; and (5) an analysis of the relevant consumer logic and emotional consumer attributions.

4.1 Score item analysis

In this study, the three dimensions expressed using the Likert scale were turned inside out, so that the degree of dissatisfaction and traditional attribution expressions could be displayed, as shown in Figure 2, with the distance between the midpoint and the center indicating the degree of satisfaction; that is, the shorter the distance, the higher the product satisfaction. In this paper, the three scores (the quality of the goods, the logistics service, and consumer service attitudes) from the 508 individual subjects were connected in a closed triangle, in which the degree of overlap is shown in the color transparency.

Table 1 shows the correlation coefficients for the score items, with all relevant significant relationships showing horizontal dominance. The correlation coefficient between quality of goods and logistics was shown to be negative, indicating that most participants perceived the poor quality to be attributable to either the goods or the logistics. Therefore, as a generally negative evaluation was given for the goods and logistics services under this endogenous evaluation system, the accuracy of the attributions could be questioned. However, in real management activities, as the usual practice is to review the system from management to the operating level in response to negative feedback, management cannot generally respond quickly.

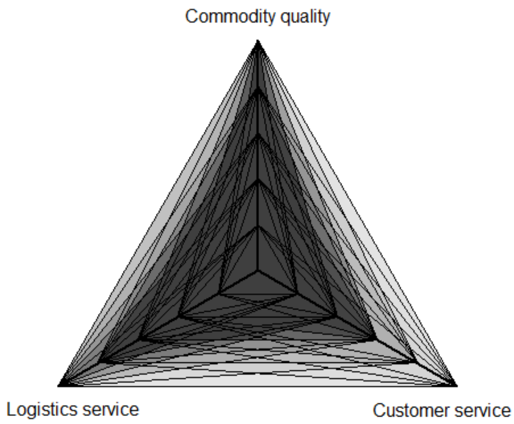


Figure 2. Score item analysis

Table 1. Correlation coefficients between score items

Correlation coefficients	Commodity quality	Logistics service	Customer service
Commodity quality	1	-0.198**	0.132**
Logistics service	-0.198**	1	0.329**
Customer service	0.132**	0.329**	1

Note: **: significantly correlation on 0.01 level.

4.2 Text mining

4.2.1 Cause attribution measurements

After separately preprocessing the text comments, the word frequency statistics were counted. Excluding entries that had less than 15 words, the resulting word cloud is shown in Figure 3.

Through an effective combining of the words and word clouds, a series of elements surrounding the quality problems; “broken hole”, “customer service”, “logistics”, “quality” and “seller”; were clearly exposed. In contrast to traditional methods for measuring perceived quality (Figure 2), managers are able to quickly locate the immediate causes for quality problems, and take measures to prevent the problem from further deteriorating. This direct effect was demonstrated in a preliminary analysis of the output structure of the consumer perception system, from which it was observed that to indirectly express the consumer cause attribution component, the system could quickly and accurately capture the main quality issues without the assistance of a non-difference variable structure. The highly abstract main components have a certain strategic significant for management; however, in recent years market segmentation and customer-orientation has become more important, making information such as “broken hole” more significant than highly abstract variables such as a score of 0, especially when these highly abstract principal components are only substitutes for a real influencing factor. Although these abstract measures are useful from a management perspective, convincing calculations are not possible because of the failure of entropy in the abstraction process.

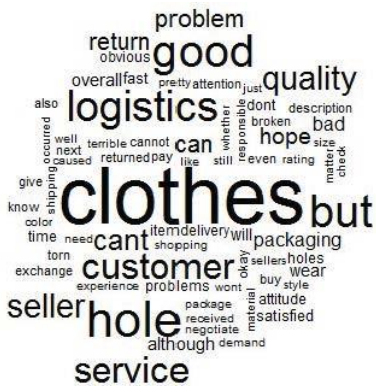
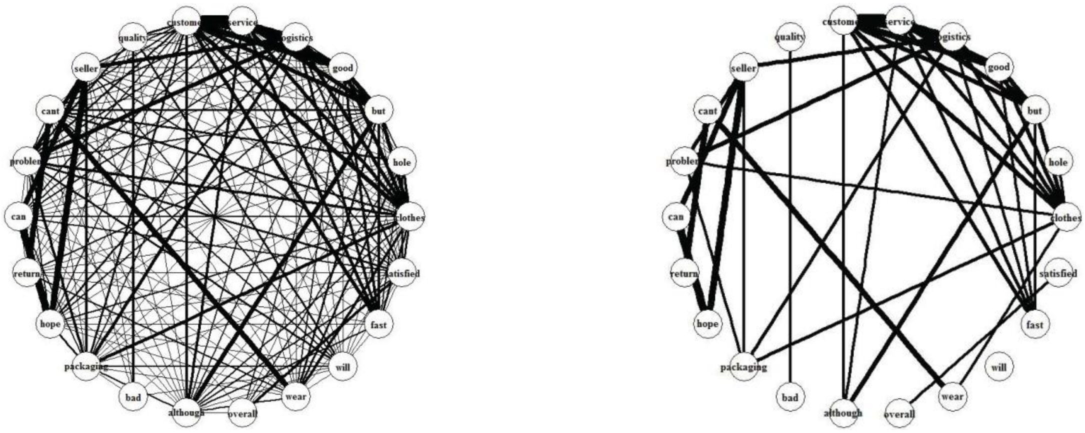


Figure 3. Word cloud.

4.2.2 Word association analysis

To maintain a maximum level of information entropy in the consumer perception system, based on the cause attribution estimates, a consumer perceived quality logic and uses *findAssocs( )* function was constructed in the tm Package in R to determine the associations between 22 entries.

Based on characteristics such as the word set sparsity and information redundancy, word entries with frequencies greater than 40 were selected, and a relationship graph drawn that showed the combined logical relationships between the high frequency words (Figure 4 (a)). To highlight the entry logics that had strong associations, the associations less than 0.2 were removed between the 22 entries, resulting in Figure 4 (b).



(a) The logical structure. **(b) The simplified logical structure.**

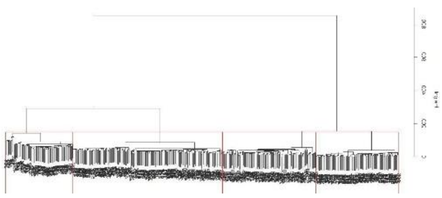
Figure 4. Word association analysis.

Figures 4 (a) and (b) appear to indicate that consumer perceived quality cannot be measured using SERVQUAL, its algorithms, or using abstract management theories. The basic cause attribution principle is to construct a unique consumer preference structure by emphasizing certain aspects in the logical structure shown in Figure 4. In addition, as the cause and results are consistent, this means that the cause is also the result. Nikhashemi and Tarodfer [31] found a high degree of similarity between consumer preferences and consumer perceived quality and predicted that the high latitude and endogeneity in the consumer cause attribution structure could be responsible. In this paper, the ordinary least square (OLS) estimation result was deemed unacceptable, which also indirectly reflected the two elements integrated structure. However, as language structure is highly logical and self-consistent, the mixing of causal factors may be normal.

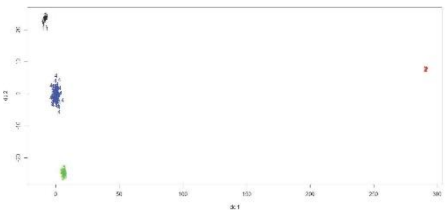
Using Figure 4 and the grammar rules, the associated entries were combined into the phrases that the consumers paid more attention to: “hole”, “can’t wear”, “can’t return”, “poor quality” and “seller does not return”. The relevance was further differentiated based on Figure 4 (b), and the logically isolated components removed so as to obtain a two logic structure that had certain associations; “nice clothes”, “but”, “clothing hole”, “can’t wear” and “seller does not return” and “no return”. As “poor quality” was an isolated structure, it was not seen as affecting the consumer tendencies towards cause attribution. Compared with the results of the measurement analysis in Section 2.1, consumers do not pay much attention to summarizing and criticizing their reasons for poor quality and are also unable to deduce strong attribution factors from “poor quality”. However, phrases such as “the seller does not return”, “no return” or “demand return” are at the center of the review comment logic, which revealed that consumers tend to have a certain strategy when assessing quality perception. Traditional perception methods have failed to identify consumer strategies and expected behaviors, and generally, consumer enthusiasm for quality management activities has also been misunderstood.

4.2.3 Text clustering

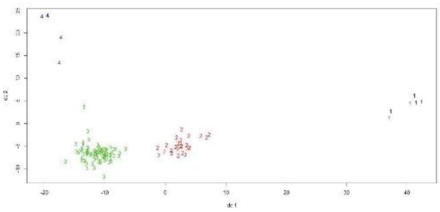
To determine which solution was satisfactory for most consumers, a text clustering method was used to classify consumer comments. As there were individual deviations in the clustering results because of the randomness of the clustering algorithm, hierarchical clustering, k-means clustering and spectral clustering were combined to derive a general proportion for the total number of consumer comments.



(a) Hierarchical clustering



(b) K-means clustering



(c) Spectral clustering

Figure 5. Visual clustering results

Figure 5 shows the results of these three clustering methods. Figure 5 (a) is the visual result for the hierarchical text clustering, from which it can be seen that the consumer reviews were clustered into four distinct categories. Table 2 displays the typical comments in each category; the 1st type simply objectively describes the quality problem and suggests a desired solution (return and exchange); the 2nd type believes the seller is responsible for the quality problem; the 3rd type believes the logistics are the cause of the problem; and the 4th type points out the quality problem directly and demands a return.

As k-means clustering and spectral clustering require the number of clusters to be customized, the number of clusters in the hierarchical clustering were referred to and initialized as 4 to facilitate comparisons across the categories. Figure 5 (b) shows the visual results for the k-means clustering and Table 3 gives a representative evaluation of the 4 categories. The 1st consumer type attributed the commodity problem to the common responsibility of both the seller and the logistics; the 2nd consumer type simply pointed out the commodity quality problems; the 3rd consumer type simply described the problem and demanded a return; and the 4th consumer type described the problem and also gave positive evaluations for the logistics and customer service.

Table 2. Typical comments from each category (hierarchical clustering)

Category	Comment
1	There are holes in the packaging and clothing. Whether it is a seller problem or a logistics problem, the customer should be able to return the item. Because I paid for the clothes, I should receive it in good condition. I am satisfied with the clothes except for the hole. The clothes feel comfortable when I wear it and it is the size recommended by customer service. I hope the seller can negotiate with the logistics and make me satisfied. I will also accept a replacement if it can’t be returned. I

	still trust the clothing quality.
2	I hope it won't happen again. The seller needs to improve greatly and correct the problem. It needs to know that the customer is good.
3	It's good overall except for the fact that the logistics caused the hole in the clothes.
4	The quality of this cloth is terrible, I need a sales return.

As k-means clustering and spectral clustering require the number of clusters to be customized, the number of clusters in the hierarchical clustering were referred to and initialized as 4 to facilitate comparisons across the categories. Figure 5 (b) shows the visual results for the k-means clustering and Table 3 gives a representative evaluation of the 4 categories. The 1st consumer type attributed the commodity problem to the common responsibility of both the seller and the logistics; the 2nd consumer type simply pointed out the commodity quality problems; the 3rd consumer type simply described the problem and demanded a return; and the 4th consumer type described the problem and also gave positive evaluations for the logistics and customer service.

Table 3. Typical comments from each category (k-means clustering)

Category	Comment
1	There is a hole in the packaging as well as the clothing. So I think the logistics should pay more attention. This hole is very obvious and I can't wear the clothes at all. I need to give a bad rating for the logistics. Also I hope the seller can pay more attention and package it better so it won't break entirely. Also I hope the seller can label the fragile packages to eliminate these kinds of problems. I can only give an okay review to the seller and a bad review for the logistics.
2	The clothing quality is bad.
3	There is a hole and I can't return the clothes. I am very disappointed.
4	Overall is okay but there is a hole which requires careful consideration, customer service is very good .

Figure 5 (c) gives the visual result for the spectral clustering. Table 4 gives the typical comments from each category classified using the spectral clustering; the 1st consumer type attributed the problem to the seller and logistics; the 2nd consumer type gave a positive review for the logistics, seller, and customer service despite the quality problems; the 3rd consumer type pointed out the hole in the clothes and suggested several solutions (return and replacement); however, as the holes seriously affected the consumer's overall impression, the 4th consumer type attributed the problem to either the seller or to logistics.

Table 4. Typical comments from each category (spectral clustering)

Category	Comment
1	There is a hole in the packaging as well as the clothing. So I think the logistics should pay more attention. This hole is very obvious and I can't wear the clothes at all. I need to give a bad rating for the logistics. Also I hope the seller can pay more attention and package it better so it won't break entirely. Also I hope the seller can label the fragile packages to eliminate these kinds of problems. I can only give an okay review to the seller and a bad review for the logistics.
2	Customer service is very good and the clothes match the descriptions. There are no color differences and the size is appropriate The packaging and clothes are torn.
3	There was a hole in the clothes when it was sent and I can't return it I can't wear it at all so bad rating.
4	The clothes are good and the customer service is very good. But the package is broken and the clothes are torn as well. It is torn, it is torn. I'm applying for compensation with logistics.

Table 5 shows the numbers from each clustering method.

Table 5. Numbers for each clustering method

Category	Hierarchical clustering	K-means clustering	Spectral clustering
1	193	115	84
2	87	25	120
3	121	62	232
4	107	306	72

A comparison of the results from the three clustering methods indicated that the internal group structure and logical structure were relatively stable. Although the three clustering results were not highly consistent, the three typical cluster structures effectively subdivided the consumer comments into typical sub classes.

The similarities between the groups indicated that overall, most consumers felt that they needed to solve the undefined quality problems, with 60% of total consumer comments agreeing with this perspective. A further 25% of consumers believed the problem was caused by logistics, 21% believed the problem was caused by product quality, and 54% did not attribute the problem to any causes. However, no consumers attributed the quality problems to warehousing or any point prior to warehousing on the supply chain, and no other strategies besides "return" and "exchange" were mentioned; therefore, when managers are seeking to deal with undefined quality problems, controlling outcomes could be better than controlling the process [32].

Therefore, this study cautiously concluded the following. First, there were obvious disparities and frames of reference in the consumer perception system. In this study, of the samples (55 cases), all of which were from consumers with a higher education in business, management or economics, none were found that extended the quality problem to the supply chain before warehousing. Second, the attribution expressions were deemed to be insufficient for the consumer perception system as only 47% of the sample gave any cause attribution, with nearly half only stating the facts and asking for a return. Third, the inter-group consumer analyses suggested that strategy and behavior expectations were unified, a return was a clear proposition, and there was a general group consensus; however, there were no strategies other than "return" and "exchange". In comparison to conclusions obtained from traditional perception quality analysis methods, this study of consumer reviews revealed the consumer's requirements for after-sales service, which in turn revealed the problems with traditional perception quality analysis methods. From the comment information analyses, it was determined that the consumer was dissatisfied with the customer service. The contradiction between traditional perception evaluations and the text mining method proved that consumers have cognitive dissonance when using traditional methods to evaluate perceived quality.

4.2.4 Sentiment analysis

Unstructured text data contains logical information and sentiment information (tendencies and fluctuations). This paper explored the correlations between logic and sentiment information in text reviews to analyze the relationships between consumer sentiment and the consumer perception system. From the database and equations (15), the sentiment tendency E_{inc} and E_{dec} were calculated for each sentence in 508 reviews.

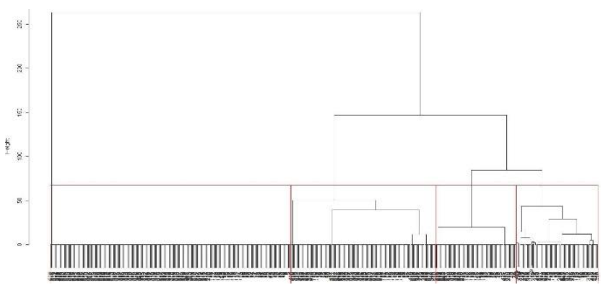


Figure 6. Sentiment clustering results

E_{inc} and E_{dec} were then used as cluster variables to cluster the comments and derive the cluster visualization results (Figure 6).

To compare the results of the hierarchical text clustering, it was divided into 4 categories. Figure 7 shows the comparisons for the numbers of comments in each category for the hierarchical text clustering and the sentiment analysis clustering, with the number of connections indicating the number of comments in the two classes. Figure 7 indicated that there was no statistical significance in the correlation between the logical hierarchical clustering results and the sentiment analysis clustering results. Jang and Numkang [33] found that consumer sentiment had a significant impact on consumer preferences and perception. While this study used different analytical angles and methods, it was not possible to obtain evidence to support these conclusions.

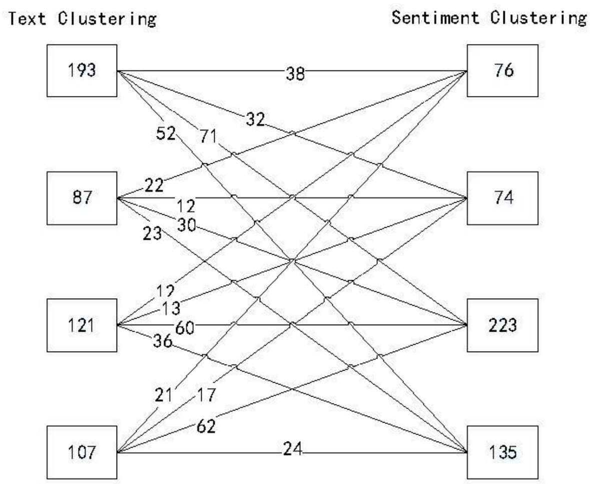


Figure 7. Text clustering and sentiment clustering comparison

There are several possibilities for these results.

1. Although the virtual experimental environment maximized the extent to which the consumer perception process and consumer behavior were reproduced, because the consumer results were simulated, the psychological defense mechanism possibly weakened the emotional volatility [34] and was therefore not able to accurately measure the dynamic wave data.

2. Related demand urgency research has highlighted that demand driven consumption could affect consumption mood fluctuations, which can reduce psychological defense mechanisms and cause consumer mood swings. The virtual consumer environment and objects designed in this study did not take account of the urgent demand characteristics, so there were insufficient strong emotional volatilities.

3. As the consumer perception system has separate logical attribution and emotional volatility forms under certain constraints, it is surmised that there is no correlation between the logical structure and emotional fluctuation when there is a general constraint condition. When the constraint condition changes, the violent fluctuations in emotions may replace or change the logical structure.

4. The algorithmic tool results in a loss of the time series characteristics, resulting in structural deviations.

5. CONCLUSIONS

This research explored the consumer quality perception system using text mining technology. Different from traditional perception evaluation methods, this research captured consumers' highly unified solutions to undefined quality problems, and proved that the consumer quality perception system had a typical line-of-sight, with consumers usually attributing undefined quality problems to only those aspects of the supply chain they can perceive. It was found that customer service control was more efficient than process control. However, determining more accurate ways to assess the sentiment tendencies in texts to explore the relationship between consumer sentiments and their perceptions requires further study.

ACKNOWLEDGEMENT

This research was supported by the National Natural Science Foundation of China Committee 71350007, the Science and Technology Department of Shaanxi Province Government 2013KCX(2)008, the Shaanxi Land Construction Group SXDC201208.

REFERENCES

[1] Jacoby J, Olson J, Haddock R. (1971). Price, brand name, and product composition characteristics as determinants of perceived quality. Journal of Applied Psychology, 55(6): 570-579.

[2] Grönroos C. (1984). A Service Quality Model and its Marketing Implications. European Journal of Marketing, 18(4): 36-44.

[3] Smith A, Bolton R. (2002). The effect of customers' emotional responses to service failures on their recovery effort evaluations and satisfaction judgments. Journal of the Academy of Marketing Science, 30(1): 5-23.

[4] Marakanon L, Panjakajornsak V. (2017). Perceived quality, perceived risk and customer trust affecting customer loyalty of environmentally friendly electronics products. Kasetsart Journal of Social Sciences, 38(1): 24-30.

[5] Lee G, Lin H. (2005). Customer perceptions of e-service quality in online shopping. International Journal of Retail & Distribution Management, 33(2): 161-176.

[6] Steenkamp J. (1990). Conceptual model of the quality perception process. Journal of Business Research, 21(4): 309-333.

[7] Folkes V, Kotsos B. (1986). Buyers' and Sellers' Explanations for Product Failure: Who Done It? Journal of Marketing, 50(2): 74-80.

[8] Brochado A, Pereira C. (2017). Comfortable experiences in nature accommodation: Perceived service quality in Glamping. Journal of Outdoor Recreation & Tourism, 17: 77-83.

[9] Buzzell R, Gale B. (1989). The PIMS principles : linking strategy to performance. Journal of Marketing, 53(2): 126.

[10] Parasuraman A, Zeithaml V, Berry L. (1985). A conceptual model of service quality and its implications for future research. Journal of Marketing, 49: 41-50.

[11] Marek M. (2005). Evaluating a museum as a tourist product using the servqual method. Museum Management & Curatorship, 20(3): 235-250.

[12] Cronin J, Taylor S. (1992). Measuring service quality: A reexamination and extension. Journal of Marketing, 56(3): 55-68.

[13] Peter J, Churchill G, Brown T. (1993). Caution in the Use of Difference Scores in Consumer Research. Journal of Consumer Research, 19(4): 655-662.

[14] Stylidis K, Madrid J, Wickman C, & Söderberg R. (2017). Towards Overcoming the Boundaries between Manufacturing and Perceived Quality: An Example of Automotive Industry. Procedia CIRP, 63: 733-738.

[15] Choi T, Eboch K. (1998). The tqm paradox: relations among tqm practices, plant performance, and customer satisfaction. Journal of Operations Management, 17(1): 59-75.

[16] Özer Ö, Raz G. (2011). Supply chain sourcing under asymmetric information. Production & Operations Management, 20(1): 92-115.

[17] Davis D, Allen J, Crosenza R. (1988). Segmenting local residents by their attitudes, interests, and opinions toward tourism. Journal of Travel Research, 27(2): 2-8.

[18] Gerard S, Wong A, Yang C. (1974). A vector space model for automatic indexing. Communications of the Acm, 18(11): 613-620.

[19] Gerard S, Christopher B. (1988). Term-weighting approaches in automatic text retrieval. Information Processing and Management, 24(5): 513-523.

[20] Dhillon, Inderjit S, Dharmendra S. (2001). Concept decompositions for large sparse text data using clustering. Machine Learning, 42(2): 143-175.

[21] Xu D, Tian Y. (2015). A comprehensive survey of clustering algorithms. Annals of Data Science, 2(2), 165-193.

[22] Rohlf F. (1982). 12 single-link clustering algorithms. Handbook of Statistics, 2(82): 267-284.

[23] Ward J. (1963). Hierarchical grouping to optimize an objective function. Publications of the American Statistical Association, 58(301): 236-244.

[24] Dan P, Moore A. (2000). X-means: Extending K-means with Efficient Estimation of the Number of Clusters. Seventeenth International Conference on Machine Learning. Morgan Kaufmann Publishers Inc, 727-734.

[25] Luxburg, Ulrike. (2007). A tutorial on spectral clustering. Statistics & Computing, 17(4): 395-416.

[26] Socher R, Perelygin A, Wu J, Chuang J, Manning C, & Ng A, et al. (2013). Recursive deep models for semantic compositionality over a sentiment treebank, EMNLP, 1631: 1631-1642.

[27] Oscar R. (1972). Positivistic measures of aggregate supply elasticities: some new approaches: comment. American Journal of Agricultural Economics, 54(3): 530-531.

[28] Hoffman D, Kalsbeek W, Novak T. (2000). Internet and web use in the united states: baselines for commercial development. Communications of the Acm, 39: 36-46.

[29] Zellner, Wendy. (2000) Wooing the Newbies.(efforts of online retailers to attract new customers). Business Week, 3681: 116.

[30] Bill D, Stuart D, Fiona T, Jim C. (2005). Perceived risk and the internet shopping experience in online purchasing behaviour. Journal of Global Information Management, 13(2): 66.

[31] Nikhashemi S, Tarofder A, Gaur S, Haque A. (2016). The effect of customers' perceived value of retail store on relationship between store attribute and customer brand loyalty: some insights from malaysia . Procedia Economics & Finance, 37: 432-438.

[32] Liu S, Wang L, & Huang W. (2017). Effects of process and outcome controls on business process outsourcing performance: moderating roles of vendor and client capability risks. European Journal of Operational Research, 260(3): 1115-1128.

[33] Jang S C, & Namkung Y. (2009). Perceived quality, emotions, and behavioral intentions: application of an extended mehrabian–russell model to restaurants. Journal of Business Research, 62(4): 451-460.

[34] Kuftyak E. (2014). Psychological defense and coping in the prognosis of children mental adaptation. Procedia - Social and Behavioral Sciences, 159: 180-183.