Journal of Computers Vol. 30 No. 5, 2019, pp. 220-226
doi:10.3966/199115992019103005017

Hierarchical Online Comment Classification for Internet Word of Mouth Management

Hua Zhao[1], Chunming Xu[1], Zhi Chen[1], Qingtian Zeng[2*]

[1] College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao, 266590, China

[2] College of Electronic Information and Engineering, Shandong University of Science and Technology, Qingdao, 266590, China
{huamolin, 13061301081, binghamton2017, qtzeng}@163.com

Received 13 January 2019; Revised 21 March 2019; Accepted 21 March 2019

Abstract. Online comment analysis is a vitual part of Internet word of mouth management. Online comment is often about several aspects. But only the really related aspect is useful. And at the same time, sentiment analysis is the key part of IWOM. A hierarchical online comment classification for Internet word of mouth management is proposed. Firstly, an aspect classification method based on the automatic recognition of high-frequency aspect words is proposed, with the aim to classify the comments into different aspects. Secondly, a sentiment classification method based on the automatic expansion of the sentiment dictionary is given, judging the sentiment orientations of the comments. Finally, experiments are carried out on the Taobao comments. Experimental results show that the proposed method can find out the different aspects based on the detection of the really related aspects, and at the same time, can judge the sentiment orientation successfully, which is very helpful for IWOM management.

Keywords: aspect classification, Internet word of mouth, online comment, sentiment classification

1 Introduction

In the era of Web2.0, there are various communication methods such as comment website and social network sites. Those methods enable consumers to record or share their experiences and feelings on Internet anytime and anywhere, which forms online commentary big data. These data indicate consumers' comment about a product or a service, which bring IWOM (Internet word of mouth). IWOM is very important for enterprises to obtain development opportunities, to deal with emergencies, to build core competitiveness, and to shape sustainable competitive advantages, but it is also a double-edged sword. Positive word-of-mouth helps the promotion of products and services and brand recognition, while negative ones will lead companies lose their competitive advantage [1]. To this end, online comment analysis for IWOM management has become one of the current research hotspots [2].

Online comments usually contain several different aspects [3], for example quality, service, and so on. In order to deal with IWOM precisely, the first step is to extract content that is truly related to products. At the same time, sentiment orientation analysis is the key part of IWOM management [4]. For these reasons, a hierarchical classification method for IWOM management is proposed, which divides online comments into two levels: the upper level is the aspect classification, which regards the extraction of text as a classification problem, dividing the comments into different aspects based on comment objects; the second level is sentiment classification, that is, through sentiment analysis, dividing comments on each aspect into three categories: positive, negative, and neutral.

In order to evaluate the proposed method, we collect Taobao (http://www.taobao.com) comments automatically using the web crawler, and then verify the performance of the proposed method base on

* Corresponding Author

these comments. Experimental results show that the proposed method can effectively extract the comments that really related to enterprise products, accurately determine the sentiment orientation of the comments, which is very helpful for IWOM management.

2 Related Works

Nowadays, as different kinds of online comments exploding, IWOM management based on online comments analysis becomes a current research hotspot, for these comments contain information that is useful both to consumer and to enterprise. Based on the analysis of the IWOM characteristics and the harm of negative reputation, J.-H. Li et al. proposed an IWOM crisis early warning model based on online comment mining, which includes four steps: warning source monitoring, evaluation, index setting and crisis pre-control [5]. Based on European consumer satisfaction model, Zhang et al. created a fuzzy dictionary of negative IWOM crisis from four aspects: quality, value, reputation and expectations, and then proposed IWOM crisis warning method based on fuzzy-comprehension judgment [6]. Wu and Yan found that factors such as the comments quantity, comment scores, comments quality, and source credibility can significantly affect consumers' adoption of E-WOM (Electronic Word of Mouth) in social networks [7].

As a key part, sentiment classification technology has been adopted widely in the IWOM management. At present, the sentiment classification methods can be divided into two categories: machine learning-based methods [8-9] and sentiment dictionary-based methods [10-11]. Wu et al. surveyed the sentiment classification methods used in the online comment analysis [12]. Vinodhini and Chandrasekaran compared and analyzed the application of various neural network methods on sentiment classification [13], and found that Probabilistic Neural Networks got the best performance based on the experiments on Amazon datasets. Tian et al. combined rules and Support Vector Machine (SVM) for sentiment classification, and the experimental results showed that rules can effectively improve SVM performance [14]. Tang et al. adopted sentiment ontology to help KNN (K-Nearest Neighbors) for sentiment classification [15]. Li and Li constructed a convolutional neural network for sentiment classification of online comments [16]. You and Bai firstly constructed a sentiment classification vocabulary based on certain framework, and then labeled emotional semantic role of each word, which was finally used for sentiment phrase calculation [17]. Different from these above studies focusing on sentiment classification methods, Choi and Lee studied the influence of many other factors on sentiment classification, and found that factors such as dataset size and document length influenced sentiment classification performance as well [18].

3 System Framework

Fig. 1 presents our system framework. Data is the basis of IWOM management. So, the online comments are firstly automatically crawled by the web crawler and stored in the database after pre-processed. Secondly, these comments are classified into different aspects during the aspect classification step. In this step, we adopt this aspect classification to extract the really related comments to the products. So, based on the aspect classification results, the comments classified into product aspect are regarded to be related to the product, which are useful and can be further used during IWOM management. Finally, we carry out sentiment classification for each aspect. During the sentiment classification step, the emotional words with high-frequency would be automatically identified to expand sentiment dictionary, which is then used for sentiment classification.

Taobao is a large e-commerce website in China with huge daily trading volume and considerable quantities of hot commodity comments. Therefore, this paper decided to use the comments about a pair of shoes (which are a Taobao's hot-selling product) as the experiment data. A web crawler is designed to get relevant comment data. By removing invalid data, 2,659 comments are finally retained for experiments. The crawled comments are pre-processed before being used, which include Chinese Word Segmentation (done by jieba tool) and stop words removing. During the segments, a user dictionary collected manually is added into jieba, which improves the segmentation accuracy.

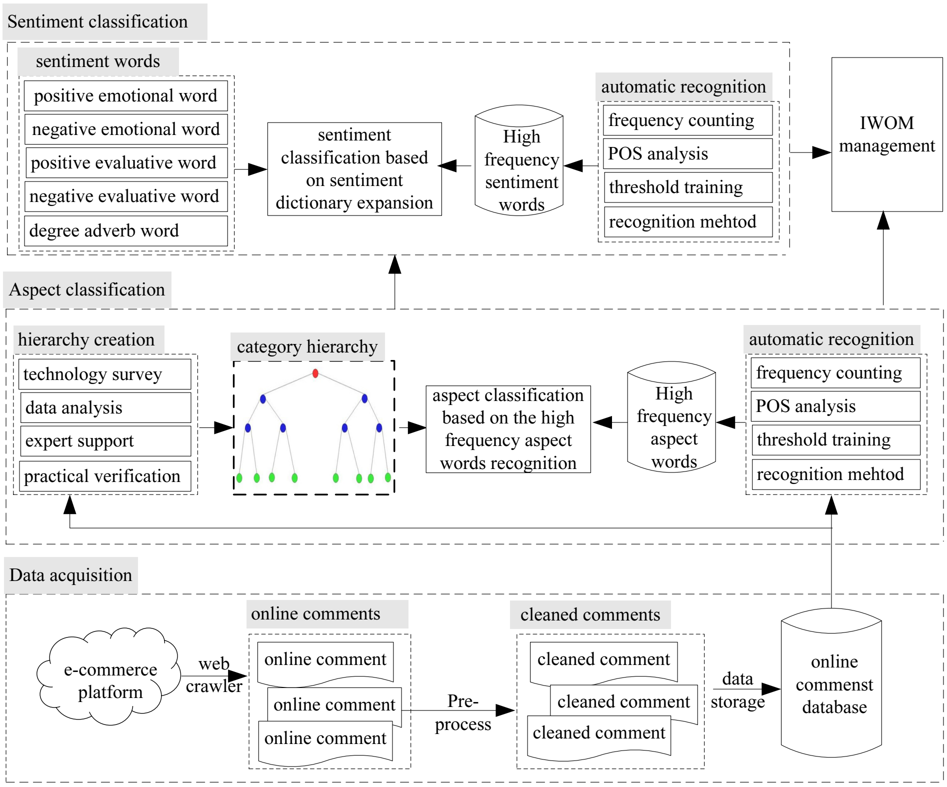

Fig. 1. System framework

4 Hierarchical Classification Method

Based on the analysis of the obtained Taobao comment, we find that most of comments are related to product, seller and delivery. So, the upper level (aspect) is defined to include three classifications, which are product aspect, seller aspect and delivery aspect. And then, in the second level, the comments within an aspect are classified into three classes according to their sentiments, which are positive, negative and neutral.

4.1 Aspect Classification Based on Automatic Identification of High Frequency Words

The method for automatically identifying high-frequency aspect words in this paper is: count overall frequency of each word within all comments. If the frequency of a word is greater than the high-frequency aspect word threshold and the word is a noun, then this word is regarded as high frequency aspect word. Based on this method, we has extracted 45 high-frequency words, including 32 high-frequency words about product, 11 high-frequency words about sellers, and 2 high-frequency words about delivery, as shown in Table 1. We need to explain here that the number of the words listed in Table 1 is less than 45, because the selected words in our work are Chinese, and here we present their corresponding English for better understanding, some different Chinese words corresponding to same English words, so we combine two same English words, which results in less words in Table 1.

Table 1. High-frequency aspect words

aspect	high-frequency aspect words
product	shoes, shoe welt, shoe type, insole, shoe sole, shoestring, vamp, shoe size, yardage, product, color, quality, shape, style, pattern, size, exterior, appearance, workmanship, thing, price, lining, this shoe, size, real object, cost performance
seller	seller, shop, shop owner, store, this store, attitude towards customers, service, our store, customer service, business
delivery	logistics, delivery

After automatically identifying high frequency aspect words, **Algorithm 1** is used to classify the comments into different aspect class.

```
Algorithm 1. Aspect Classification Algorithm
Input: Pre-processed comments, high-frequency words set of product
       denoted as ProductSet, high-frequency words set of seller
       denoted as SellerSet, high-frequency words set of delivery
       denoted as DeliverySet
Output: Product class denoted as ProductClass, seller class denoted as
        SellerClass, delivery class denoted as DeliveryClass
1. Repeat the process until all online comments have been processed;
2. Read a comment from the database and divide it into several clauses
   with punctuation as a separator: Reveiw={P1, P2,..., Pn}, where
   n=number of punctuation marks +1;
3. for (i=1 to n)
4.     if (Len(Pi)≠0)) //when clause is not empty
5.     {
6.        Pi={w1, w2, ... wim}; // w1, w2, ... wim are the words
          contained by Pi
7.        Class=false;
8.        for (j=1 to im)
9.        {
10.          if (wj∈ProductSet)  {Pi→ProductClass; Class=true;}
11.          if (wj∈SellerSet)   {Pi→SellerClass; Class= true;}
12.          if(wj∈DeliverySet)  {Pi→DeliveryClass; Class= true;}
13.       }
14.       if (Class==false)
15.      {
16.          if (i==1)  Pi→ProductClass;
17.          else  Pi→Class(Pi-1);  // Class(Pi-1) indicates the
   category of Pi-1
18.        }
19.    }
```

4.2 Sentiment Classification Based on Sentiment Dictionary Expansion

The sentiment classification based on sentiment dictionary is one of the important methods in sentiment classification. We use the sentiment dictionary provided by Hownet (http://www.keenage.com/html/c_bulletin_2007.htm) as our basic dictionary, which contains about 17,887 words, dividing into the following categories: (1) positive sentiment words, such as love, happy; (2) negative sentiment words, such as bad, sad; (3) positive evaluative words, such as tuneful, profound; (4) negative evaluative words, such as ugly, painful; (5) degree adverb, which could be subdivided into six categories.

At first, we combines “positive sentiment words” and “positive evaluative words” to create a positive word set, then combines “negative sentiment words” and “negative evaluative words” to form a negative word set. Secondly, uses the high frequency word identification (in Section 3.1.1) to identify some adjective high frequency words and added to the corresponding collection as sentiment words. Finally, use **Algorithm 2** to classify sentiment.

```
Algorithm 2. Sentiment Classification Algorithm
Input: Comments belonging to a certain aspect class, positive words
       set denoted as PositiveSet, negative words set denoted as
       NegativeSet, degree adverbs set denoted as DegreeSet
Output: The sentiment category of each input comment
1. Repeat the process until all comments in this aspect have been
   processed;
```

```
2.  Read a comment, remove the high frequency words listed in Table 1,
    result denoted as rComment, assuming rComment={w_1, w_2,..., w_m},
    where w_i is the word included in rComment;
3.  sum=0;
4.  have=false;
5.  for (i=1 to m)
6.  {
7.      if (w_i∈PositiveSet OR w_i∈NegativeSet)
8.      {
9.         have=true;
10.        Denote the weight of the w_i as Weight(w_i);
11.         if (i>1)
12.            if (w_i-1∈DegreeSet)
13.            {
14.               Denoted the weight of w_i-1 as Weight(w_i-1);
15.               sum=sum+ Weight(w_i)* Weight(w_i-1);
16.            }
17.          else
18.             sum=sum+ Weight(w_i);
19.      }
20. }
21. if (have==false  OR  sum==0)   return  Neutral evaluation / no
    evaluation;
22. if (sum>0)   return  Positive evaluation
23. else        return  Negative evaluation;
```

5 Experimental Results and Analysis

In order to evaluate the proposed hierarchical classification algorithm, 2,000 comments are randomly selected from 2,659 Taobao comments as training corpus, and the remaining 659 comments are used as test data. Meanwhile, precision and recall are used as evaluation metric. Through multiple trainings, the high-frequency word threshold is set to 5, the positive word weight is set to 1, and the negative word weight is set to -1. The degree adverb weight is set to be one of six levels according to their degrees, which are 2, 1.5, 1, 0.7, 0.5, and 0.2, respectively. In the case of these parameters values, the experiments were carried out on the test corpus, the experiment results are shown in Table 2 and Table 3, respectively.

Table 2. Experimental results of aspect classification

category	original data	test results	correct classification	precision	recall
product	2659	2601	2502	96.2%	94.1%
seller	245	296	198	66.9%	80.8%
delivery	183	190	178	93.7%	97.3%
average	--	--	--	86.6%	90.7%

Table 3. Experimental results of sentiment classification

category	sentiment	original data	classification results	correct classification	precision	recall
product	positive	2599	2523	2455	97.3%	94.5%
	neutral	45	114	32	28.1%	71.1%
	negative	15	22	5	22.7%	33.3%
seller	positive	243	241	240	99.6%	98.8%
	negative	2	4	1	25%	50%
delivery	positive	176	165	164	99.3%	93.1%
	negative	7	17	5	29.4%	71.4%
	average	--	--	--	66.9%	85.4%

The following conclusions can be obtained from the above experimental results:

(1) When consumers publish online comments, they pay more attention to the products themselves. At the same time, among several product attributes, consumers are most concerned about "quality" and "price". So, we can conclude that, in the era of Web2.0, improving the cost-effectiveness of products is the most fundamental guarantee for good IWOM.

(2) Besides products themselves, consumers will also pay attention to sellers and delivery. So, when selling goods on the e-commerce platform, it is also necessary for manufacturer to carefully select the cooperative sellers and delivery, for avoiding bad impact on IWOM.

(3) The data set also has some influence on the performance of the online comments analysis. For example, in above experiments, most comments in online comments related to product, and most of them are positive comments, so the performance of aspect classification for product is the best. As for the other aspects which have less data, the precision is poor, but the recall is relatively high. The reason for this phenomenon is that if there are more positive comments, it is easier to identify the nouns or emotional words which are useful for classification when counting high-frequency words, thus the dictionary coverage rate would be improved, and the precision will be improved correspondingly.

6 Conclusions

After price competition, quality competition and service competition, the competition between enterprises has entered the stage of word-of-mouth competition. IWOM is the core component of current word-of-mouth management, and online comment analysis is a key part of IWOM management. This paper proposes and implements a hierarchical online comment classification method for IWOM management. Firstly, the aspect classification is implemented based on the automatic recognition of high-frequency aspect words, and the comments are divided into different aspects according to their evaluation object. Secondly, the sentiment classification is implemented based on the expansion of the sentiment dictionary, and the sentiment orientation of each comment is determined. Finally, experiments were conducted on Taobao product comments which are automatically obtained. The experimental results show that the method can effectively find the perspective of consumers when they publish online comments, and accurately determine their sentiment tendencies, providing technical support for enterprise to management IWOM.

Acknowledgements

This work is supported by the social science and humanity on Young Fund of the ministry of Education (16YJCZH154, 18YJAZH017); CAS Key Lab of Network Data Science and Technology Open Fund Project (No. CASNDST201706); the SDUST Research Fund (No. 2015TDJH102); the SDUST Cultural Construction Project (WHJSZ-1705); International cooperation and training program for outstanding young teachers in Shandong Colleges and Universities; Sci & Tech. Development Fund of Shandong Province of China (ZR2017MF027), and the Taishan Scholar Climbing Program of Standing Province.

References

[1] H. Yu, Z.-P. Wang, Research on the effects of online word-of-mouth communication, Journal of Intelligence 32(6)(2013) 100-106.

[2] H. Kim, D.-M. Hanssens, Advertising and word-of-mouth effects on pre-launch consumer interest and initial sales of experience products, Journal of Interactive Marketing 37(2)(2017) 57-74.

[3] S. Dong, X.-H. Wang, Z.-H. Ge, Analysis on the content characteristics of the online comment about b2c shopping websites based on text mining, Library Theory and Practice 6(2017) 54-58.

[4] C. Hung, H.-K. Lin, Using objective words in Senti-WordNet to improve word-of-mouth sentiment classification, IEEE Intelligent Systems 28(2)(2013) 47-54.

[5] J.-H. Li, Y.-S. He, Y.-Q. Ma, S.-P. Li, Z.-W. Li, Research of corporate IWOM crisis early warning based on online comments mining in the era of big data, Journal of Intelligence 34(2)(2015) 53-58.

[6] Y.-F. Zhang, H. Li, L.-H. Peng, The fuzzy inference of internet word of mouth crisis early warning of negative comment mining, Library and Information Service 60(9)(2016) 75-82.

[7] P.-F. Wu, Q. Yan, Impact of online comments on adoption of electronic word-of-mouth in social networks, Journal of Beijing University of Posts and Telecommunications (Social Sciences Edition) 17(1)(2015) 52-61.

[8] F. Kokkinos, A. Potamianos, Structural attention neural networks for improved sentiment analysis, in: Proc. Conference of the European Chapter of the Association for Computational Linguistics, 2017.

[9] X.-W. Liao, Y.-Y. Xie, J.-J. Wei, L. Gui, X.-Q. Cheng, G.-L. Chen, Perspective level Microblog sentiment classification based on convolutional memory network, Pattern Recognition and Artificial Intelligence 31(3)(2018) 219-229.

[10] D Bollegala, D Weir, J Carroll, Cross-domain sentiment classification using a sentiment sensitive thesaurus, IEEE Transactions on Knowledge & Data Engineering 25(8)(2013) 1719-1731.

[11] X.-D. Yan, T. Huang, Tibetan sentence sentiment classification based on emotion dictionary, Journal of Chinese Information Processing 32(2)(2018) 75-80.

[12] Y.-L. Wu, Y. Huang, X.-F. Wang, Research on online users comments in Chinese: basing on the perspective of affective computing, Information Science 35(6)(2017) 159-163.

[13] G. Vinodhini, R.-M. Chandrasekaran, A comparative performance evaluation of neural network based approach for sentiment classification of online comments, Journal of King Saud University (Computer and Information Sciences) 28(1)(2016) 2-12.

[14] F. Tian, F. Wu, K.M. Chao, Q.H. Zheng, N. Shah, T. Lan, J. Yue, A topic sentence-based instance transfer method for imbalanced sentiment classification of Chinese product comments, Electronic Commerce Research & Applications 16(3)(2016) 66-76.

[15] X.-B. Tang, J. Zhu, F.-H. Yang, Online comment sentiment classification based on ontology and KNN, Information Studies: Theory & Application 39(6)(2016) 110-114.

[16] J. Li, H. Li, Research on product feature extraction and sentiment classification of short online comment based on deep learning, Information Studies: Theory & Application 41(2)(2018) 143-148.

[17] L.-P. You, X.-Y. Bai, Hotel service quality measurement based on online comments sentiment analysis and TOPSIS, Information Science 35(10)(2017) 13-17.

[18] Y. Choi, H. Lee, Data properties and the performance of sentiment classification for electronic commerce applications, Information Systems Frontiers 19(3)(2017) 993-1012.