for users purchase decisions, but also provide the accurate basis to business for work out direct-sales strategy and improve the product quality. As we all know, it is difficult to get the valuable information from a flood of online review [2]. So how to analyze the review data quickly and find the information we need is a challenging task. The goal of this paper is to automatically acquire and analyze the emotional element in the natural language comments, which contain two subtasks: 1) Emotional element extraction which contains emotional objects and emotional words detection. 2) An analysis of the emotional tendency based on the emotional element.

Although many scholars have conducted some research in sentiment analysis, and made a lot of achievements, but in the fine-grained $\langle emotionalobject, emotionalwords \rangle$ word pair extraction and emotional tendency analysis, the comprehensive effect is still a big gap between the practical distance. In this paper, a fine-grained emotional element detection and emotional tendency judgment method based on conditional random fields (CRFs) and support vector machine (SVM) was proposed. First, syntactic feature, semantic role feature, part-of-speech feature are introduced into this method and emotional word and emotional object are extracted based on multi-granularity CRFs. Then emotional element collocation, syntactic feature, deep semantic of emotional element and emotional word polarity feature are used to construct SVM for emotional tendency judgment. Finally, we get the three-tuple $\langle emotionalobject, emotionalword, emotionaltendency \rangle$. The emotional tendency in the three-tuple represents the emotional object tendency in the sentence environment.

2 Related Work

Recent years, many researchers have made some research on emotional element detection and emotional tendency analysis [3]. In 1993, Agrawal et al. [4] first proposed the concept and model of association rules. Popescu et al. [5] by calculating the value of the mutual information between the noun or noun phrase and the specified identifier, to determine the possibility of belonging to the evaluation object. Jin et al. [6] Proposed rule-based algorithm. They design a variety of collocation rules to match the text by summing up the collocation order of emotional object and emotional words. L. Liu et al.[7] combined conditional random fields with syntax tree pruning for the fine-grained emotion analysis of product reviews. The tri-training method based on MapReduce was used to get the emotional words.

Emotion classification can be divided into four levels according to different granular-ity: word level, sentence level, paragraph level, and chapter level. Different from the level of coarse-grained sentiment analysis, fine-grained sentiment analysis can extract and identify finer emotional information, such as emotional expression range, emotional objects, emotional words, emotional polarity of the emotional elements. Fine grained sentiment orientation analysis technology has attracted more and more attention of researchers. The study of document sentiment classification can be traced back to Hearst [8]. He use the model based on cognitive linguis-tics to judge the whole emotion of the whole document. B. Y. Li et al.[9] proposed a method to analyze the chapter sentiment based on a single annotation cascade model which combined sentence patterns with sentence position as the characteristic. Fu and Wang [10] put forward the

application of fuzzy set theory to Chinese sentence sentiment classification. They define three fuzzy sets to represent the three kinds of emotion polarity (positive, negative and neutral). L. J. Zheng et al.[11] compared the difference between sentence level sentiment classification and chapter level sentiment classification, and find that the granularity has a greater influence on the accuracy of classification.

3 Emotional Element Detection

3.1 Conditional Random Fields

Conditional random fields (CRFs) are undirected graph learning model proposed by Lafferty in 2001[12]. CRFs combined many excellent properties of maximum entropy model (MEM) and hidden markov model (HMM), these excellent properties make the CRFs is very suitable to solve the sequence labeling problem. CRFs is a sequence annotation model, as shown in Figure 1. Let X be a set of input random variables whose values are observed, and Y be a set of output random variables whose values the task requires the model to predict. The random variables are connected by undirected edges indicating dependencies, and let $G = (V, E)$ be the undirected graph such that there is a node $v \varepsilon V$ corresponding to each of the random variables representing an element Y_v of Y.

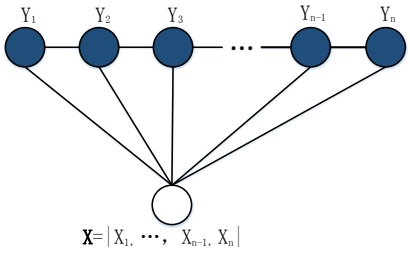

Figure 1: Graphical structure of a chain-structured CRFs for sequences

In the task of Sequence labeling, random variables $X = |X_1 \cdots X_{n-1}, X_n|$ and $Y = |Y_1 \cdots Y_{n-1}, Y_n|$ are the observe sequence and target sequence respectively, then linear-chain CRFs thus define the conditional probability of a state sequence Y given an input sequence X to be

$$P(Y \mid X) = \frac{1}{Z(X)} exp\left[\sum_{i,k} \lambda_k f_k (y_{i-1}, y_i, x) + \sum_{i,k} \mu_k g_k (y_i, x)\right] \quad (1)$$

where f_k is the transition feature function of the entire observation sequence and the labels at position i and $i-1$ in the label sequence; g_k is a state feature function of the label at position i and the observation sequence; and λ_k and μ_k are parameters to be estimated from training data. $Z(X)$ is a normalization factor over all state sequences.

3.2 Select multi-granularity features

Most traditional methods extracted emotional objects by using CRFs[13], while emotional words were ignored. Or found the emotional words first and then extracted the emotional objects by association rules. Different from the traditional element detection, emotional objects and emotional words will be extracted synchronously in this article. This article focused on the sentence structure and semantic, so the dependency syntax features and syntax tree features were introduced. The features used in the detection are as follows:

- Word Feature (WF): Words are the smallest grammatical unit that can express the meaning of words. Emotional words and emotional objects are both consisted of word.

- Part of Speech Feature (POS): POS is an implicit feature, one of the characteristics of natural language processing frequently used. Some POS used in this article are as follows: Structure word, Preposition, Entity name, Adjective, Personal Pronouns, Place name, Omit the word, Demonstrative Pronoun, Interjection, Modal Particle, Tense word,

- Semantic Role Dependency Parsing (DP): Semantic role dependency parsing is not only a more advanced and more in-depth implicit feature, but also is a manifestation of sentence semantic. The semantic roles used are as follows: Subject, Object, Preposition, Relevance, Complement, Questions linked, Punctuation, Sigh, Tense, Parallel, Attribute

- Parent of the word in syntax parse tree (PW): In the syntax parse tree, each word has a parent node and has some kind of relationship with its parent node.

3.3 Sequence labeling with CRFs

Emotional word and emotional object detection can be considered as marking out some specific words from the word sequence. Therefore, this problem can be handled by CRFs. The location mark set is symbolized as: BA, BB, EA, EB, SA, SB, O, P. The specific meaning of BA is object in front of emotional word. BB means emotional word in front of emotional object. EA means emotional object behind emotional word. EB means emotional word behind emotional object. SA means emotional object without emotional word. SB means emotional word without emotional object. P means punctuation and O means others. This paper only considered the situation of the emotional objects and the emotional words in pairs. The output format of CRFs is emotional object and emotional word pair.

In the label list, BA and EA both represent emotional object, BB and EB both represent emotional word. But the difference between them is the location relationship. Strictly speaking, SA is not emotional object, it is just object. Although SB is an emotional word, but cannot find the emotional object corresponding to it. But SA and SB approximate emotional objects and emotional words from the feature level, so separate them into two new categories. The number of labeled O is much more than others. To balance the data and improve precision of the tag, punctuations are separated from O, and marked as P.

3.4 Emotional object and word detection

The complexity of annotation symbols affects the accuracy of emotional objects and emotional word detection. BA and EB, BB and EA are in pairs in most cases, but some special structure of sentence and CRFs mislabeled could cause BA, BB, EA, EB appear alone or in the wrong order. It would reduce the detection accuracy if these errors cannot be deal with. Some rules as follows were used for hierarchical filtering errors:

- yesterday/O received/O phone/SA. This is a non-opinion sentence, just have an object phone, no emotional word, so just simply ignore it.
- Clothes/BA bought/O yesterday/O,/P very/O beautiful/EB. Clothes is emotional object, beautiful is emotion word. But these two words are distributed in two different clauses. Pair-words in different sub- sentence may cause errors, so the situation is ignored.
- Product/BA quality/BA is/O very/O good/EB. The product and quality are labeled as BA. But they are not parallel relationship, so merge it into an emotional object.

4 Emotional Tendency Judgment

4.1 Support Vector Machine

SVM was proposed by the Cones and Vapnik in 1995 [14]. It is a statistical machine learning method and is mainly used to solve classification problem. The mechanism of SVM is looking for a hyperplane meet for the requirement of classification, which is a best support vector to distinguish two different classes, as shown in Figure 2.

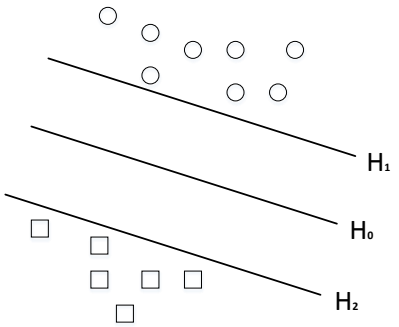

Figure 2: Separating hyperplane

H_1 and H_2 are parallel to H_0 and they are close to the two kinds of samples. Such distance between H_1 and H_2 is called the interval. The optimal classification line is to find the classification line which makes the interval maximum. The samples on H_1 and H_2 called support vector. Use the following formula to express the classification line:

$$w \cdot x + b = 0 \quad w \in R, b \in R \tag{2}$$

Only by minimizing the $\|w\|^2/2$ then we can get the optimal classification line. $\|w\|$ is the two norm of w. Because the fine-grained analysis of the emotional tendency in this article can be regarded as a classification problem, so SVM can be used to solve this problem.

4.2 Features description

The same emotional word would possibly exhibit different emotions when it modified different emotional objects. Therefore, to analyze the emotional tendency more accurately, only consider the emotional word is not enough. The sentence structure and the dependent relationship of emotional words and emotional objects must be combined. Following are some features and descriptions adopted in emotional tendency judgment: (1)Emotional object (SS): The same word when modified the different emotional objects, tendency may be different. So the impact on the final analysis result by emotional object need be taken into consideration. (2) Emotional word (SW): Emotional words indicate emotional tendency, each word has a basic emotional feeling. We use a dictionary to do the vectorization of $\langle emotionalobject, emotionalwords \rangle$, each word has a corresponding numerical representation in the dictionary. Finally, we can convert a two-tuple into a N dimensional feature vector $(N_1 N_2 \cdots N_n)$. (3) Emotional object semantic code (SSC): Assuming the emotional tendency of "processor frequency is very high" is known, but "CPU" specific meaning is unknown. And also the relationship between CPU and processor is unknown. It is difficult to know the emotional tendency of "CPU frequency is very high". If the fact that "processor" and "CPU" have the same meaning is known, then the latter's emotional tendency can be judged correctly and easily. In order to make computer know two words whether have similar meaning, the code of emotional object semantic (SSC) was introduced. In Sec 4.3, a method would be introduced in detail to get the semantic code. (4) Emotional word semantic code (SWC): Like SSC, SWC is used to show two emotional words whether have the same or similar meaning. (5) Emotional tendency inversion (ETI): Whether there is a word to reverse emotional tendencies, Just go through from the syntax parsing tree to find the sentence whether contains a negative word as adverbial to modify emotional word. (6) Basic emotional tendency of emotional word (BPW):The basic emotional polarity can get from HowNet. If the word is not in HowNet, the basic emotional polarity can get by SO-PMI algorithm [15]. Pointwise Mutual Information (PMI) is calculated by the following formula:

$$PMI(w_1, w_2) = log \frac{P(w_1, w_2)}{P(w_1)P(w_2)} \tag{3}$$

$p(w_1, w_2)$ is the co-occurrence frequency of w_1 and w_2. SO-PMI is calculated by the following formula:

$$SO\_PMI(w) = \sum_{pw \in P_{set}} PMI(w, pw) - \sum_{nw \in N_{set}} PMI(w, nw) \tag{4}$$

Here, Pset is positive emotion word collection and Nset is negative emotion word collection.

4.3 Semantic code acquisition

Semantic code refers to the number of words with the same or similar meaning. That is to say if two words have a similar meaning, they belong to a same collection, also have a same semantic code. Semantic code can be constructed by a synonym dictionary. We proposed an algorithm to convert the word into SSC as follows.

To cluster words, the first step is to map the word to the N-dimensional vector space according to its context. Semantic code for each word can calculate by automatic coding neural network like Feedforward Neural Network Language Model (NNLM)[16] , it divided neural network into the input layer, hidden layer and output layer. The computational complexity per each training example is:

$$Q = N * D + N * D * H + H * V \tag{5}$$

Where N is the n of the n-gram, D is the dimension of each word, H is the number of nodes in the hidden layer, and V is the number of nodes in the output layer. Through multilayer neural network, each word can be mapped to a N-dimensional space, and can easily calculate the Euclidean distance between any two words w_i and w_j.

$$S(i,j) = \sqrt{\sum_{k=1}^{n} \left(x_{ik} - x_{jk}\right)^2} \tag{6}$$

Then using the K-means algorithm to cluster all the words according to their meaning vector, center coordinates of each category is calculated by the following formula.

$$x_i = \frac{\sum_{j=1}^{m} x_{ji}}{m} \tag{7}$$

According to the formula , the words belong to which category can be determined. By the above method to learn from a large number of data, thus the words can convert into vectors, then using the K-means algorithm to cluster all the words according to their meaning vector, give each category a fixed unique number, this number is the semantic code of this category words.

5 Experiment and Result Analysis

5.1 Emotional element detection

In order to verify the effectiveness of the proposed method, we have carried out a series of comparison experiments. First we used the product reviews from www.tmall.com to construct review corpus. It contains 4146 review data. We used 2500 as the training set, and the rest as a test set, marked as Corpus 1. We used the method based on CRFs model and the method based on association rules to detect the emotional objects and emotional words. The result is shown in Figure 3.

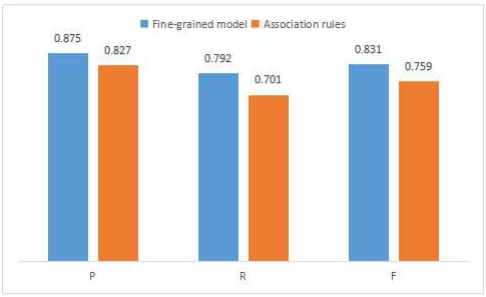

Figure 3: Emotional element detection result comparison on product reviews

5-fold cross-validation was used to optimize parameters on the training set. We can see that the precision is high and the recall rate is relatively low. The high accuracy is because the features we selected make the detection rules stricter. As long as the condition gets satisfied, almost can guarantee it corrects. Recall rate is relatively low is not only because of the great arbitrariness of colloquial language in comments expressed. Another important reason is that there are a lot of typos and punctuation missing in the reviews. It reduced the accuracy of word segmentation and POS tagging. We can also see that the method based on CRFs get a better result than the method based on association rules because the limitation of the method based on association rules is relatively bigger. It can get good effect on regular sentences, but it performed badly for the complex sentences of greater freedom degree.

Figure 4: Emotional element detection result comparison on micro-blog

In order to improve the robustness of the method, we collected 2000 micro-blog review data and made artificial marks on them. Then 1000 of them were used as the training data and the other as the test data to construct micro-blog corpus which marked as Corpus 2. The result is shown in Figure 4. We can see that though the result is not good as the previous result on review data, but it is still a good result. The decline of precision is not obvious which verified the effectiveness of the proposed detection method. The recall rate fell more, mainly because the content of micro-blog is more diverse than the reviews. It means more complex sentences

and more new words. So emotional object and emotional word detection is more difficult. This is also the reason why the research on micro-blog corpus generally got low recall rate.

5.2 Emotional tendency judgment

Emotional tendency labeling was performed by artificial for the emotional object and emotional word pair which extracted in the previous section. In order to verify the validity of the fine-grained sentiment classification method, we used three methods to judge the emotional tendency on the same corpus for comparing.

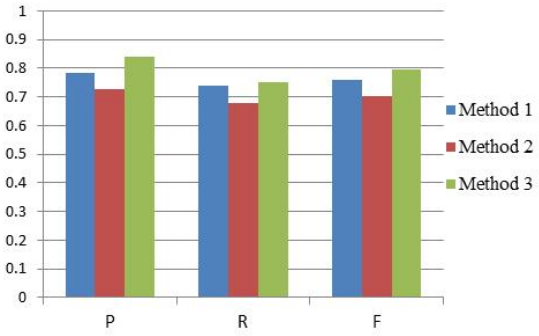

Figure 5: Emotional tendency judgment result comparison on product reviews

Method 1: the method based on sentence; Method 2: the proposed fine-grained classification method; Method 3: the proposed fine-grained classification method with semantic code. The unlabeled review and micro-blog mixed data containing 500 thousand product reviews and 500 MB micro-blog were used to get the semantic code. The comprehensive result is shown in Figure 5. It should be noted that the fine-grained model is based on the emotional element detection which is only recognized the objects and words that extracted in section 5.1. So the results of the previous task have a direct impact on the recall rate of current experiments. We did the same comparison experiment on the corpus 2. The result is as shown in Figure 6.

It can be observed from Fig. 5 and Fig. 6 that the fine-grained classification model with semantic code is significantly better than the sentence level emotional judgment. The introduction of semantic code greatly improved the accuracy of emotion classification. The reason why semantic code can get a better result is that fine grained itself is targeted to emotional objects and emotional words. The introduction of semantic code further made up the loss of the sentence and semantic. And the method based on sentence classification is so general that erroneous judgement was happened easily for the sentence of more than one emotional object.

To have an intensive study on the deep learning based semantic code, we used the two different sources to obtain the SC. One is the mixed unlabeled corpus used in section 5.2; the

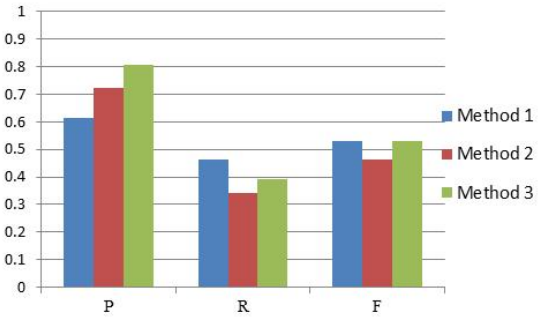

Figure 6: Emotional tendency judgment result on micro-blog

other is a text source of 1GB micro-blog data. Then we did the emotional tendency judgment on both corpus 1 and corpus 2. The experimental results showed that SC1 gets by product reviews have a bigger influence on product review corpus than micro-blog corpus and the SC2 gets by micro-blog as well has a bigger influence on micro-blog corpus. This means the consistent performance of the corpus can effectively improve the final emotion judgment. This phenomenon is due to the expression and word usage habit of the different types of corpus will result in large differences. So the characterization ability of semantic codes iterated from different corpus will have differences in different areas. The experimental results above further demonstrated the effectiveness of the proposed method in this paper.

6 Conclusion and Future Work

By analyzing the experimental results, it can be concluded that the proposed method can ensure the correct rate and the recall rate of product review corpus is high. Although the recall rate of micro-blog corpus is low, but compared with other similar methods, also achieve a higher performance because of the introduction of semantic information with deep features. Although this article has achieved good experimental results, but for cross-cutting and complex sentences, accuracy and recall rate can still be further improved. Further work will be carried out on the selection of optimal features and unknown word processing.

Acknowledgment

The work is supported by the Natural Science Foundation of Anhui Province(1508085QF119) and State Key Program of National Natural Science of China(61432004). This work was partially supported by the China Postdoctoral Science Foundation funded project(2015M580532). This research has been partially supported by National Natural Science Foundation of China under Grant No.61472117.

References

[1] Hu, Weishu, Z. Gong, and J. Guo. "Mining Product Features from Online Reviews." IEEE, International Conference on E-Business Engineering IEEE, 2010:24-29.

[2] Mackiewicz, J, and D. Yeats. "Product Review Users' Perceptions of Review Quality: The Role of Credibility, Informativeness, and Readability." Professional Communication IEEE Transactions on 57.4(2014):309-324.

[3] Hong, Yili, P. Y. Chen, and L. M. Hitt. "Measuring Product Type With Dynamics of Online Product Review Variance." Social Science Electronic Publishing (2012).

[4] Agrawal, Rakesh, et al. "Mining association rules between sets of items in large databases." ACM SIGMOD International Conference on Management of Data ACM, 1993:207-216.

[5] Popescu, Ana Maria. "Extracting product features and opinions from reviews." Hlt/emnlp on Interactive Demonstrations Association for Computational Linguistics, 2005:32-33.

[6] Jin, Wei, and H. H. Ho. "A novel lexicalized HMM-based learning framework for web opinion mining NOTE FROM ACM: A Joint ACM Conference Committee has determined that the authors of this article violated ACM's publication policy on simultaneous submissions. Therefore ACM has shut of." International Conference on Machine Learning ACM, 2009:465-472.

[7] Q Liu, B Ma. Product features and emotion tendency mining[J].Information Technology and Informatization. 2015(12)

[8] Hearst, Marti A. "Direction-Based Text Interpretation as an Information Access Refinement." Text-based intelligent systems L. Erlbaum Associates Inc. 1999.

[9] Benyang, L. I., et al. "Single-label Cascaded Model for Document Sentiment Analysis." Journal of Chinese Information Processing 26.4(2012):3-158.

[10] Fu, Guohong, and X. Wang. "Chinese sentence-level sentiment classification based on fuzzy sets." International Conference on Computational Linguistics: Posters Association for Computational Linguistics, 2010:312-319.

[11] Zheng, Li Juan, H. W. Wang and K. Q. Guo. "Sentiment Classification of Chinese Online Reviews: A Comparison between Sentences and Paragraphs." Journal of the China Society for Scientific and Technical Information, 32(4)(2013) 376384

[12] Lafferty, John D., A. Mccallum, and F. C. N. Pereira. "Conditional Random Fields: Probabilistic Models For Segmenting And Labeling Sequence Data." 3.2(2001):282–289.

[13] Xu, B., et al. "Extraction of Opinion Targets Based on Shallow Parsing Features." Zidonghua Xuebao/acta Automatica Sinica 37.10(2011):1241-1247.

[14] Joachims, Thorsten. Text categorization with Support Vector Machines: Learning with many relevant features. Machine Learning: ECML-98. Springer Berlin Heidelberg, 1998:137-142.

[15] Lou, De Cheng, and T. F. Yao. ”Semantic polarity analysis and opinion mining on Chinese review sentences.” Journal of Computer Applications 26.11(2006):2622-2625.

[16] Mikolov T, Sutskever I, Chen K, et al. Distributed Representations of Words and Phrases and their Compositionality[J]. Advances in Neural Information Processing Systems, 2013, 26:3111-3119.