A Novel Feature-based Method for Sentiment Analysis of Chinese Product Reviews

LIU Lizhen[1], SONG Wei[1], WANG Hanshi[1], LI Chuchu[1], LU Jingli[2]

[1] Information and Engineering College, Capital Normal University, Beijing 100048, P. R. China
[2] Agresearch Ltd, New Zealand

Abstract: Sentiment analysis of online reviews and other user generated content is an important research problem for its wide range of applications. In this paper, we propose a feature-based vector model and a novel weighting algorithm for sentiment analysis of Chinese product reviews. Specifically, an opinionated document is modeled by a set of feature-based vectors and corresponding weights. Different from previous work, our model considers modifying relationships between words and contains rich sentiment strength descriptions which are represented by adverbs of degree and punctuations. Dependency parsing is applied to construct the feature vectors. A novel feature weighting algorithm is proposed for supervised sentiment classification based on rich sentiment strength related information. The experimental results demonstrate the effectiveness of the proposed method compared with a state of the art method using term level weighting algorithms.

Key words: sentiment analysis; sentiment strength; opinion mining; dependency parsing

I. INTRODUCTION

Sentiment Analysis is to identify and extract subjective information in text materials, such as opinions and feelings. With the explosion of Chinese online product reviews, sentiment analysis starts to play an important role in opinion mining and product recommendation. This information can help manufactories to improve their products or services, and also helps potential customers make purchase decisions. Thus, accurate understanding of sentiment expressed in Chinese reviews can bring tremendous business opportunities.

Sentiment (polarity) classification is a main task in sentiment analysis which is usually considered as a binary classification problem, that is to classify a given review's polarity as positive or negative, and extensive study has been conducted on analyzing English and other European languages [1-7]. Many of these methods are directly applied to sentiment analysis of Chinese language without considering the language habits of Chinese users [8], [9] and [10]. Though n-gram based features are still effective for sentiment analysis as for traditional text classification, opinion words (and phrases) , negation words and other syntactic and word dependency related features have a more direct effect for sentiment analysis which may lead to further improvements [11]. Opinions expressed and the corresponding product features should be highlighted for this task. Moreover, few researchers take sentiment strength into consideration for feature weighting in supervised sentiment classification. In fact, sentiment with the same polarity may reflect different degrees of sentiment strength [12]. The degree of sentiment strength also shows the user preferences on products or product features. For instance, "The touch

In this paper, we propose an energy efficient transmission scheme for orthogonal frequency division multiplexing (OFDM) systems in Long Term Evolution (LTE) which is developed based on reduction of the cyclic prefix(CF) energy.

screen is really cool." indicates a very strong positive opinion and a main concern, whereas "The camera is good." indicates a weak positive attitude. In addition, for Chinese sentiment analysis, more difficulties involve. First, Chinese texts have to be segmented. Incorrect segmentation may result in incorrect prediction. Second, comparing with English, the sentiment dictionary resources are incomplete and of difficulty to construct due to that some Chinese words are highly context dependent. Therefore, sophisticated natural language processing technique, such as syntactic parsing, is necessary for Chinese sentiment analysis. The above observations motivate us to conduct research on analyzing both sentiment orientation and sentiment strength of Chinese product reviews with the assistance of dependency parsing.

In this paper, we focus on analyzing opinions in reviews expressed by Chinese customers. We propose a feature-based vector model for in-depth sentiment analysis of review texts. Our method not only identifies modifying relationships between words using dependency parsing, but also incorporates the sentiment strength by considering multiple factors such as the usages of punctuations, over-modifiers and the adverbs of degree in sentiment sentences. The same sentence with different punctuation may express different emotion. For example, "He loves this mobile?" is different from "He loves this mobile!". Generally, "!" express a stronger emotion, while "?" often invert the sentiment polarity of the opinion. Also, over-modifiers such as "too" tends to invert the polarity of the opinion words. For example, "The size of the phone is too big" may indicate a negative polarity. By considering these factors, we build a richer representation of opinions on product features. This representation is more complete compared with the quintuple model proposed by Liu.B [11] which mainly focuses on the polarity of opinions on objects, but ignores the modifying relationship between words and sentiment strength information. Based on the sentiment strength of opinion expressions, we propose a novel weighting algorithm to assign weights to feature-based vectors for classifying the polarity of reviews.

The rest of the paper is organized as follows. In Section 2, related work on sentiment analysis is reviewed. In Section 3, a feature-based vector model and a novel weighting algorithm are proposed. Section 4 describes a novel feature extraction method that is based on dependency parsing. Experimental results and discussions are presented in Section 5, and we conclude in Section 6.

II. Related Work

The approaches for sentiment analysis can be roughly categorized into two groups: approaches based on machine learning techniques and approaches based on lexicons. Most studies have been focusing on training sentiment classifiers based on machine learning algorithms. Pang et al. [7] use three traditional machine learning methods to classify movie reviews of Internet Movie Database (IMDB), i.e., Naive Bayes, Maximum Entropy and Support Vector Machines (SVM). Their experimental results demonstrate that all these three methods outperform human-produced baselines, and SVM achieves the best result. Mullen and Collier [6] employ a hybrid SVM approach by using potentially pertinent information, including several favorability measures of terms and knowledge of the topics. Blitzer et al. [9] investigate domain adaptation for sentiment classifiers, focusing on online reviews of different types of products. Prabowo and Thelwall [13] conduct experiments on movie reviews, product reviews and Myspace comments by combining a rule-based classifier and a supervised learning algorithm, and find that the hybrid classification shows an improvement in accuracy. Machine learning approaches work well when large labeled corpora are available for training and validation [14].

There are also some lexicon-based methods for sentiment analysis, which measure the polarity of text based on lexicons derived from resources such as WordNet [15], MPQA lexicon [16] and SentiWordNet [17]. Turney [18] predicts the sentiment orientation of phrases by Point-wise Mutual Information (PMI) based on pre-defined seed words, and classifies a document as positive or negative by the average semantic orientation of the phrases. Kim and Hovy [19] build three models to assign a sentiment category to a given sentence by combining the individual sentiments of sentiment-bearing words. Kennedy and Inkpen [20] classify reviews based on the number of positive and negative terms they contain. Devitt and Ahmad [21] explore a computable metric of positive or negative polarity in financial news texts.

Lexicon-based approaches outperform machine learning approaches when a complete sentiment dictionary is available. However, the difficulty of generating a complete sentiment dictionary hinders the application of lexicon-based approaches in sentiment analysis. Therefore, our proposed method is based on machine learning techniques.

In recent years, Chinese sentiment analysis has attracted researchers' interests. Ye, Shi and Li [22] proposed an improved semantic approach for Chinese review classification by applying the PMI-IR (Information Retrieval) method to measure the similarity between pairs of words. Li and Sun [23] compare performances of four machine learning methods (i.e., SVM, NB, ME and ANN) for sentiment classification of Chinese reviews using N-grams features. Wan [9] utilizes reliable English resources to improve Chinese sentiment analysis by employing machine translation and ensemble techniques. Tan and Zhang [8] have demonstrated that SVM is the best classifier for Chinese text sentiment classification. Wan [10] proposed a bilingual co-training approach to make use of both the English view and the Chinese view based on additional unlabeled Chinese data. Wen et al. proposed a semantic based sentiment analysis method [24] where negation words and adverbs of degree were exploited. However, they didn't consider the features in sentences. In this study, we focus on developing dependency parsing and considering sentiment strength to improve the machine learning approaches for sentiment classification of Chinese product reviews.

Due to the complexity of semantic expression, simple words and phrases cannot express the semantic orientation accurately and traditional text vector models don't show good enough performances in sentiment classification. Therefore, in recent years, feature based sentiment analysis has drawn attention. This kind of methods first identifies the targets on which opinions have been expressed in a sentence, and then determines the opinion polarity as positive or negative. The targets are products and their features (aspects, attributes). Some research has been conducted to focus on the definition of feature based opinion model. Kim and Hovy [19] represent an opinion as a quadruple (topic, holder, Claim, Sentiment). Liu [12] proposed a feature-based sentiment analysis model (object, feature, opinion, opinion holder, time) to express an opinionated document. Thus, the task for a feature based sentiment analysis is to find all the opinion tuples and then determine the sentiment of a review.

However, these models ignore the language habits of Chinese users. For example, users tend to use many adverbs of degree to express opinions. Ignoring such characteristics cannot express the sentiment strength accurately. Also, these models lack concern on the over-modifier of opinion words and the usage of punctuations which may invert the sentiment polarity. In this paper, we propose a feature-based vector model for Chinese reviews which extends existing opinion models. Firstly, we use dependency parsing to identify modifying relationships between words in order to find the features opinions are expressed on, and the syntactic constituents like adverbs and punctuations which modify the opinion words. Then we use the over-modifiers and punctuations to express the sentiment strength of opinions more completely in order to enrich the opinion model. Finally, we propose a novel weighting algorithm based on sentiment strength to assign proper weights to features for effectively classifying reviews' sentiment polarity.

III. Feature-Based Sentiment Analysis for Chinese Reviews

The feature-based vector model represents the essential information for classifying Chinese reviews. In this section, we introduce the general form of the model and the weighting algorithm for assigning weights to model components. In next section, we will introduce how to construct the model based on dependency parsing in detail.

3.1 Feature-based six-tuple vector model

We define a general opinionated document D as $\{(t_1, \omega_1), (t_2, \omega_2), \ldots, (t_n, \omega_n)\}$, where ω_i is the weight of the opinion i, i=1…n. t_i is a feature-based vector that represents a reviewer's opinion on a feature of the target object or the object itself. To incorporate sentiment strength information, we define three variables to express the modifiers of an opinion, i.e., adverbs of general degree, over-modifier and negation words. The adverbs of degree express strength of sentiment polarity. Over-modifiers are the adverbs of over degree like "too" etc., which may invert the polarity of opinion words and have an influence on the sentiment orientation of reviews. Negation words can invert the polarity of opinion words. Thus, accurate analysis of modifiers of opinion words is the basis for effective sentiment classification. Each feature-based vector can be represented as a Six-Tuple vector:

$$t_i=(f, o, omc, gms, nmc, p) \qquad (1)$$

where, f is a feature (or attribute) which a reviewer comments on, o is an opinion word modifying feature f, ***omc*** is the number of over-modifier of opinion word o. There are 30 frequently used over-modifiers in Chinese language, and Chinese like to use them to express their feelings. For this reason, we extract the over-modifiers around the opinion word. The ***gms*** is the average score of general modifiers of opinion word o, ***nmc*** is the number of negation words and ***p*** is the punctuation of the opinion sentence, which reflects the tone of reviewers. Generally, the number of over-modifiers and negation words of an opinion word is no more than 2. ***P*** has three different states, i.e., statement, exclamation and interrogation. For example, "这件裙子真的非常漂亮！但是尺寸有点太大了，质量也不是很好。" ("The dress is really very beautiful! However, the size is a little too big, and the quality is not so good."). This review can be represented by the feature-opinion vector as follows:

(dress, beautiful, 0, 4, 0, exclamation),
(size, big, 1, 2, 0, statement),
(quality, good, 0, 3, 1, statement)

We employ the Chinese sentiment lexicons derived from HowNet, which is a widely used online common-sense knowledge database unveiling inter-conceptual relations and inter-attribute relations of concepts as connoting in lexicons of the Chinese and their English equivalents. HowNet classifies adverbs of degree into six levels as shown in Table 1.

In the feature-based vector model, levels 1-5 can be seen as adverbs of general degree, which can impact on the intensity of polarity, but cannot invert the sentiment polarity. We assign scores (1, 2, 3, 4, 5) to level 1-5, respectively. The value of *gms* is the average score of the adverbs of general degree for each opinion word. For each feature, we compute the average *gms*. This score is used to express the intensity of polarity. Words in level 6 can be seen as over-modifiers, which may invert the semantic orientation. The quantity of over-modifiers and negation words indicate whether the semantic orientation is inverted. Therefore we define *omc* as the number of over-modifiers associated with the opinion word and *nmc* as the number of negation words.

3.2 Feature weighting algorithm

Feature weighting is important for learning algorithms. Previous studies usually focused on the significant words for classification. Weighting algorithms such as TF and TFIDF are widely applied [25]. However, sentiment is not usually highlighted by repeating the same terms [26]. Here we propose a novel feature weighting algorithm, called High Adverb of Degree Count (HADC) based on the feature-based vector model to measure the importance of features. For each feature in a review, HADC can be described as follows:

$$hadc_f = \frac{1 + degadv_i}{n + \sum_{i=1}^{n} degadv_i} \quad (2)$$

Where n is the number of opinion word occurrences in the review and $degadv_i$ represents the number of adverbs of degree that are used to modify the ith opinion word which is expressed on feature f. In a nutshell, the main idea behind the algorithm is that sentiment words that an reviewer uses more adverbs of degree to describe is more likely to be the important elements so that the corresponding features should be more important to represent the sentiment of Chinese reviews. The reason is that Chinese people usually use more adverbs of degree to emphasize emotion. For an example, 这是一个好用的手机 , 我真的很喜欢它的触摸屏 , 声音也非常好 ("It is a nice mobile phone, I really love the touch screen very much and voice is very clear, too"), the reviewer uses two adverbs of degree to emphasize the love for screen. This shows a user preference over the feature touch screen. Thus combining the polarity of opinion words expressed on the features, it is expected to more accurately express the weights of features.

IV. Feature-Based Six-Tuple Vectors Extraction

In this section, we introduce the extraction of Six-Tuple vectors from freestyle texts for the feature-based vector model. The main challenge of extraction is to correctly identify the modifying relationship between words which helps to find features the opinions expressed on and sentence constituents indicating the sentiment strength of the opinions. In this paper, we apply dependency parsing to achieve this goal.

Table I *Levels of adverbs of degree in HowNet*

Level	Degree	Instance
1	欠 | slightly	slightly, just, merely
2	稍 |-ish	quite, some, passably
3	较 |more	such, more, so, further
4	很 |very	really, much, greatly, especially
5	极其 |extremely/ 最 |most	absolute, deeply, fully
6	超 |over	excessive, too, over

4.1 Dependency parsing

Dependency parsing reveals syntax structures by analyzing dependency relationship between language units. The main elements of dependency structures are dependency arcs. Each arc is directed pointing from core word to a subordinate word. Its label indicates the dependency relationship type. Each sentence can be represented as a parsing tree by a set of dependent arcs. For example, the dependency parsing tree of the sentence "衣服的质量非常好." (The quality of the clothes is very good.) is shown in Fig.1. (good, quality, SBV) represents a dependency arc pointing from "good" to "quality". "good" is the core word, and "quality" is the subordinate word, and "SBV" is the dependency relationship. Dependency relationships cover a set of grammatical relations including local dependencies, simple-sentence dependencies and complex-sentence dependencies.

Dependency parsing has been widely used in sentiment analysis, due to its high performance of syntactic parsing at sentence level [27] and [28]. The research result in Zhou M [29] shows that dependency parsing is more suitable for analyzing Chinese sentence than Phrase Structure Grammar. Although phrase structure grammar has achieved great success in English, it cannot effectively describe the relationship between words in Chinese sentences. However, dependency parsing can

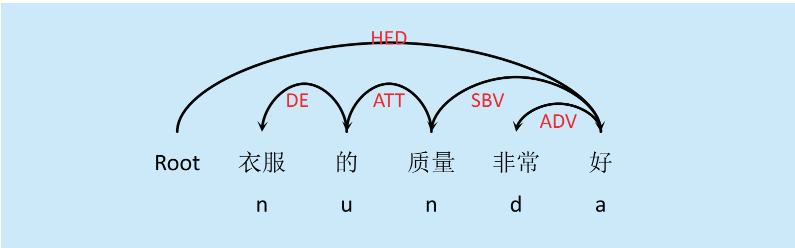

Fig.1 *Dependency parsing tree of the sentence "衣服的质量非常好"*

generate a clear structure of dominance and subordination relation between components. Furthermore, dependency parsing does not emphasis on the word order in sentences and can adapt to the flexibility of Chinese word order. However, dependency parsing is usually used for extract opinion targets but dependencies between other constitutes are neglected.

4.2 Feature-based six-tuple vectors extraction based on dependency parsing

We introduce the algorithm to extract feature-based six-tuple vectors based on dependency parsing. Before extract the six-tuple vectors, we preprocess review sentences for further feature extraction including Chinese word segmentation, parts-of-speech tagging, and dependency parsing. We use the Language Technology Platform developed by Harbin Institute of Technology (HIT-LTP) [30] to process reviews.

Our feature-based six-tuple vectors extraction algorithm has three inputs, i.e., a list of sentiment words, a list of modifiers, and review data, which are described as follows.

– A list of **sentiment words** that is used to express opinions. We use four subsets of HowNet sentiment vocabulary, i.e., positive evaluation words subset (e.g. 有效的 (effective), 安全的 (safe), negative evaluation words subset (e.g. 假的 (false), 差的 (inferior)), positive sentiment subset (e.g. 高兴 (happy), 爱 (love))and negative sentiment words subset(e.g. 失望 (disappointed), 后悔 (regret)). These words could express strong appraisal emotions, which are the most widely used sentiment words.
– A list of **modifiers** that includes negation words and adverbs of degree. Negation words usually invert the semantic orientation of reviews, such as "no" and "not". Adverbs of degree are words that have impact on the intensity of polarity, such as "more", "especially", "merely". HowNet sentiment word set provides a list of adverbs of degree containing 219 words. Furthermore, these words are divided into six levels according to their different intensity, which are shown in Table 1.
– **Review data**. Review data is transformed into an appropriate format by language processing module HIT-LTP. Then the review data is operated in DOM (Document Object Model) form, and saved as a XML file. The appropriate XML format is shown in Fig.2.

In each six-tuple, opinion feature f can represent attributes of reviewed objects, or objects themselves. Opinion features can be identified and marked by analyzing the result of dependency parsing. Considering the time complexity and computational complexity, we only focus on the dependencies that are useful for feature extraction, such as SBV, VOB, ADV, ATT, COO, CMP and HED. Descriptions of these dependencies are shown in table 2. The input is the preprocessed review data. Firstly, if the subordinate word is in SBV or VOB, and the core word in ATT is noun or pronoun, then they are more likely to be a feature, because the features are usually nouns and pronouns. Secondly, the components, which have coordinative relation (COO) with the words identified in previous step, are iden-

Table II *Dependency relations*

symbol	relation
SBV	subject-verb
VOB	verb-object
ADV	adverbial
ATT	attribute
COO	coordinate
CMP	complement
HED	head

tified as the features as well. Finally, pronouns are replaced by the previous noun that is identified as another feature occurrence. The algorithm to process one review is shown in Fig 3.

The next step is to extract the other five elements of a six-tuple. We search the dependencies of the opinion feature and identify the opinion words associated with a feature in the sentence. If the opinion feature is a subordinate word in structure SBV, and the core word can be found in the sentiment word dictionary, then the core word is viewed as an opinion word. Otherwise, we search for ADV, VOB and CMP relation around opinion feature and if the subordinate word in above relations can be found in the sentiment word list, then the subordinate word is viewed as an opinion word. If the reviewed object is a subordinate word in structure SBV and the core word can be found in sentiment word dictionary, then the subordinate word is an opinion word. If the feature is the core word in structure ATT and the subordinate word can be found in the sentiment word list, then the subordinate word is an opinion word. For example, “ 漂亮的裙子 ”(It is a beautiful skirt), “ 漂亮 ”(beautiful) is the opinion word, while “ 裙 子 ”(skirt) is the corresponding feature. Finally, if opinion words are in coordinative relation (COO), then the words in COO are opinion words too.

Furthermore, we identify the modifiers of each opinion word by analyzing the ADV relation before an opinion word. We identify modifiers and modifying types depending on a list of modifier. The scores of *omc*, *gms*, *nmc* are initialized to zero. Each adverb and inversion word is associated with the opinion word that the modifier is the most likely to describe. For each over-modifier, the score of the *omc* is increased by one. Similarly, *nmc* is increased by the number of inversion words. For each *gms* we compute the average score per adverb of general degree associated with the opinion word. Finally, we also need to identify the tone of the review sentence and assign it to *p*. If the punctuation is “!”, *p* is set to “exclamation”, else if the punctuation is “?”, *p* is set to “interrogation” and *p* is set to “statement” in other cases.

```
<?xml version="1.0" encoding="gbk"?>
<xml4nlp>
  <note sent="y" word="y" pos="y" ne="y" parser="y"
  wsd="y" srl="y"/>
  <doc>
    <para id="0">
      <sent id="0"cont=" 衣服的质量非常好。">
        <word id="0"cont=" 衣服 "pos="n" ne="O"
        parent="1" relate="DE"/>
        <word id="1"cont=" 的 "pos="u" ne="O"
        parent="2" relate="ATT"/>
        <word id="2"cont=" 质量 "pos="n"ne="O"
        parent="4" relate="SBV"/>
        <word id="3"cont=" 非常 "pos="d" ne="O"
        parent="4" relate="ADV"/>
        <word id="4"cont=" 好 "pos="a" ne="O"
        parent="-1" relate="HED"/>
        <word id="5"cont="。" pos="wp" ne="O"
        parent="-2" relate="WP"/>
      </sent>
    </para>
  </doc>
</xml4nlp>
```

Fig.2 *The format of a processed review*

V. EXPERIMENTS

5.1 Datasets and evaluation metrics

```
Input: a review preprocessed by HIT-LTP
Output: a set of opinion features
1.    Retrieval core words in ATT relation or the
subordinate word in SBV, VOB which form the candidates
W
2.    For each word w in W
3.      If the part of speech of w is noun or pronoun:
4.        Assign feature tag to w
5.        If w has coordinative relation COO:
6.          Expand features from COO
7.        Endif
8.      Endif
9.    Replace all tagged pronouns with the nearest
previous noun
10.   Endfor
```

Fig.3 *Algorithm for opinion feature extraction*

In order to evaluate the effectiveness of the proposed method, we conducted experiments on ChnSentiCorp[1] which is a publicly available corpus for Chinese sentiment analysis. We extracted 3500 documents from the corpus, which consists of customer reviews in three domains, i.e., digital product, book, and hotel. The digital product review data is crawled from 360buy.com which is a popular E-commerce website in China. The book review data is collected from dangdang.com, and the hotel review data is crawled from Ctrip which is one of the most well-known hotel booking websites in China. The size and the distribution of these three datasets is shown in Table 3.The datasets contain balanced positive and negative documents so that we can focus on evaluating the effectiveness of the features rather than other factors (such as imbalanced training data) which may affect the performance of classifiers.

Table III *The size of three sentiment corpora*

Domains	Sentiment	Documents
Digital product	Positive	800
	Negative	800
Hotel	Positive	450
	Negative	450
Book	Positive	500
	Negative	500

We adopt commonly used metrics in sentiment analysis to evaluate the effectiveness of our proposed method, Precision, Recall, F1 and the accuracy. The Precision, Recall and F1 are defined as follows respectively:

$$Precision = \frac{number of correct positive (or negative) predictions}{number of positive (or negative) predictions} \quad (3)$$

$$Recall = \frac{number of correct positive (or negative) predictions}{number of positive (or negative) examples} \quad (4)$$

$$F1 = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (5)$$

The Accuracy is calculated as follows:

$$Accurary = \frac{number of correct positive and negative predictions}{number of positive and negative examples} \quad (6)$$

5.2 Comparisons and analysis

5.2.1 Evaluation of different dependencies for six-tuple extraction

We analyzed the importance of different dependency relations for opinion feature extraction. We randomly selected 209 reviews and annotated the opinions in these reviews. Table 4 shows the ratio that Six-tuples appear in different dependency relations. As shown in Table 4, dependencies, such as COO, SBV, ADV, HED and CMP, often appear in review sentences and the elements of Six-tuples are more likely to appear in these dependency relations. Thus, we focus on these dependencies to extract Six-tuples from reviews.

5.2.2 Evaluation of feature weighting algorithms

We combine our feature-based vector model with the HADC weighting algorithm for sentiment polarity classification of Chinese product reviews. In this section, we compare the pro-

Table IV *Contrast of different dependency relations that contains elements of six-tuples*

Dependency relationship	Occurrences	Elements of Six-tuple	Ratio
SBV	344	196	53.43%
ADV	576	261	45.31%
VOB	288	92	31.94%
CMP	73	35	47.95%
ATT	411	99	24.09%
COO	49	30	61.22%
HED	203	114	56.16%
others	643	109	16.95%

[1]http://www.searchforum.org.cn/tansongbo/corpus-senti.htm

posed method with 2 baselines:

- TF: Proposed feature based vector model is adopted and the frequency is used to weight each feature.
- TFIDF: Similar to TF, TFIDF score is used to weight each feature. Though TFIDF is a classic strategy for IR [25], but it has been adapted for sentiment classification and proved to be effective on a wide selection of datasets [31].

We want to examine that whether the proposed novel weighting algorithm is superior to the traditional TF and TFIDF weighting algorithms. Since it is a supervised classification problem, any classification algorithm can be applied here. As the high performance for sentiment classification, we adopted SVM as the classifier in our experiments. We used 5-folds across variations for evaluation. And for each training datasets, across variations were also used for tuning the parameters of the SVM classifier with the RBF (radial basis function) kernel.

Table 5 presents the performances of different algorithms. It shows that our proposed novel weighting algorithm achieves better results than the TF algorithm and TF-IDF algorithm in most cases. The algorithm's performance in these three datasets verifies our initial hypothesis that opinion words, which reviewers use more adverbs of degree to describe indicate stronger sentiment strength. And the corresponding features are more likely to be the most important elements to represent the sentiment in Chinese reviews.

Fig 4 shows the accuracies of the classified reviews for different datasets with different weighting algorithms. It demonstrates that the HADC algorithm, which weights each six-tuple based on the number of adverbs of degree, outperforms TF and TFIDF weighting algorithms. The average accuracy of HADV is 0.891, which is 3% greater than TF(0.861) weighting algorithm, and 2% greater than TFIDF(0.872) weighting algorithm.

The experimental results show that our proposed six-tuple model and HADC weighting algorithm succeed to classify more than 87.5% of the reviews correctly in the worst case (Digital product dataset), and has an accuracy close to 91% in the best case.

VI. Conclusions

In this paper, we presented a feature-based vector model and a novel weighting algorithm for in-depth sentiment analysis of Chinese reviews. Specifically, adverbs of degree and punctuations were used to express the intensity of polarity more accurately. We evaluated our proposed method on three review data sets

Table V *Performance of the proposed method and the baselines*

	Method	Positive Precision	Positive Recall	Positive F1	Negative Precision	Negative Recall	Negative F1
Digital product	TF	0.8421	0.8595	0.8507	0.8540	0.8361	0.8450
	TFIDF	0.8301	0.8958	0.8617	0.8868	0.8166	0.8503
	HADC	0.8501	0.9458	0.8954	0.9389	0.8333	0.8830
Hotel	TF	0.8482	0.9111	0.8785	0.904	0.8370	0.8692
	TFIDF	0.8671	0.9185	0.8920	0.9133	0.8592	0.8854
	HADC	0.8985	0.9185	0.9084	0.9166	0.8962	0.9063
Book	TF	0.8034	0.9266	0.8606	0.9133	0.7733	0.8375
	TFIDF	0.9285	0.7852	0.8509	0.8160	0.9403	0.8738
	HADC	0.8452	0.9281	0.8847	0.9166	0.8231	0.8673

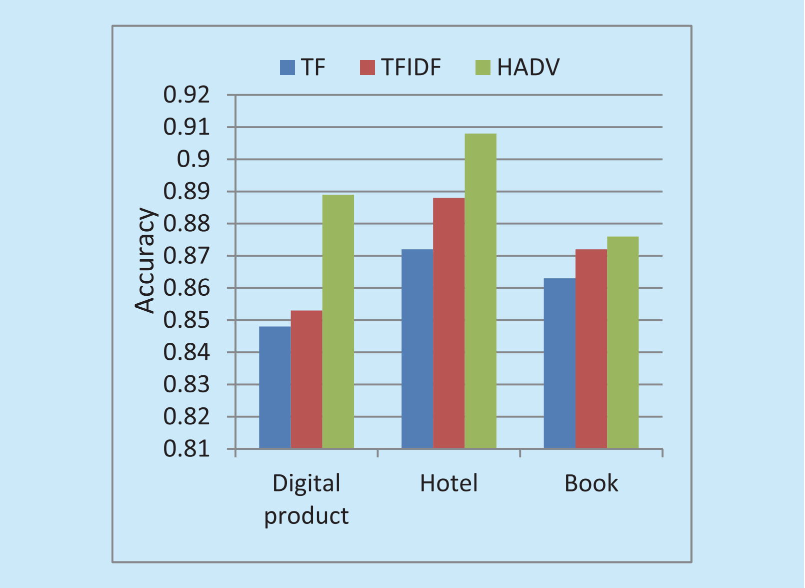

Fig.4 *Accuracy of three weighting methods for sentiment classification*

from different domains. The experimental results demonstrated that the sentiment strength information is useful for weighting features and further improves the performance of sentiment classification.

ACKNOWLEDGMENT

This work was supported in part by National Natural Science Foundation of China under Grants No. 60970052; the Beijing Natural Science Foundation under Grants No.4133084; the Beijing Educational Committee Science and Technology Development Planned under Grants No. KM201410028017 and the Beijing Key Disciplines of Computer Application Technology.

References

[1] Li S, Zhou G, Wang Z, Lee SYM, Wang R (2011) Imbalanced sentiment classification. In: Proceedings of CIKM-2011, pp 2469–2472.

[2] Tang H, Tan S, Cheng X (2009). A survey on sentiment detection of reviews. Expert Systems with Application 36(7):10760–10773.

[3] Blitzer J, Dredze M, Pereira F (2007) Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In: Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, pp 440–447.

[4] Hu M, Liu B (2004) Mining and summarizing customer reviews. In: Proceedings of the ACM SIGKDD conference on Knowledge discovery and data mining (KDD), pp 168–177.

[5] Narayanan R, Liu B, Choudhary A (2009) Sentiment analysis of conditional sentences. In: Proceedings of the conference on empirical methods in natural language processing (EMNLP), pp 180–189.

[6] Mullen T, Collier N (2004) Sentiment analysis using support vector machines with diverse information sources. In: Proceedings of EMNLP-2004, pp 412–418.

[7] Pang B, Lee L, Vaithyanathan S (2002) Thumbs up? Sentiment classification using machine learning techniques. In: Proceedings of the ACL-02 conference on Empirical methods in natural language processing, pp 79-86.

[8] Tan S, Zhang J (2008) An empirical study of sentiment analysis for chinese documents. Expert Systems with Applications 34(4):2622-2629.

[9] Wan X (2008) Using bilingual knowledge and ensemble techniques for unsupervised Chinese sentiment analysis. In: Proceedings of EMNLP-2008, pp 553-561.

[10] Wan X (2011) Bilingual co-training for sentiment classification of Chinese product reviews. Computational Linguistics 37 (3):587-616.

[11] Thelwall, Mike, Kevan Buckley, and Georgios Paltoglou. "Sentiment strength detection for the social web." Journal of the American Society for Information Science and Technology 63.1 (2012): 163-173.

[12] Liu B (2010) Sentiment Analysis and Subjectivity. In: Indurkhya N and Damerau FJ (ed) Handbook of Natural Language Processing, 2nd edn., pp 627–667.

[13] Prabowo R, Thelwall M (2009) Sentiment analysis: A combined approach. Journal of Informetrics, 3(2): 143–157.

[14] Andreevskaia A, Bergler S (2008) When specialists and generalists work together: Overcoming domain dependence in sentiment tagging. In: Proceedings of ACL-08: HLT, pp 290-298.

[15] Miller G, Beckwith R, Fellbaum C, Gross D, Miller K. (1990). Introduction to wordnet: An on-line lexical database. International Journal of Lexicography, 3(4): 235–312.

[16] Wilson T, Wiebe J, Hoffmann P (2005). Recognizing contextual polarity in phrase-level sentiment analysis. In: Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLT/EMNLP), pp 347–354.

[17] Esuli A, Sebastiani F (2006) Sentiwordnet: A publicly available lexical resource for opinion mining. In: Proceedings of the 5th conference on language resources and evaluation (LREC), Genova, Italy.

[18] Turney PD (2002) Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In Proceedings of ACL-02 ,pp 417–424.

[19] Kim SM, Hovy E (2004) Determining the sentiment of opinions. In: Proceedings of COLING-04, 20th International Conference on Computational Linguistics, pp 1367-1373.

[20] Kennedy A, Inkpen D (2006) Sentiment classification of movie reviews using contextual valence shifters. Computational Intelligence 22(2):110-125.

[21] Devitt A, Ahmad K (2007) Sentiment polarity identification in financial news: A cohesion based approach. In: Proceedings of ACL-07, pp 984-991.

[22] Ye Q, Shi W, Li Y (2006) Sentiment classification for movie reviews in Chinese by improved semantic oriented approach. In: Proceedings of 39th Hawaii International Conference on System Sciences.

[23] Li J, Sun M (2007) Experimental study on sentiment classification of Chinese review using machine learning techniques. In Proceedings of

Natural Language Processing and Knowledge Engineering.2007, pp 393–400.
[24] Wen B. et al. (2010) Text Sentiment Classification Research based on Semantic Comprehension. Computer Science 37(6):261-264.
[25] Salton G, Buckley C (1988) Term-weighting approaches in automatic text retrieval. Information Processing and Management 24 (5):513–523.
[26] Eirinaki M, Pisal S, Singh J (2012) Feature-based opinion mining and ranking. Journal of Computer and System Sciences 78(4): 1175-1184.
[27] Joshi M, Penstein-Rosé C (2009) Generalizing dependency features for opinion mining. In: Proceedings of the ACL-IJCNLP 2009 Conference Short Papers, pp 313–316.
[28] Thet TT, Na JC, Khoo CSG (2010) Aspect-based sentiment analysis of movie reviews on discussion boards. Journal of Information Science, 36(6): 823-848.
[29] Zhou M (2000) A Block-based Robust Dependency Parser for Unrestricted Chinese Text.In: Proceeding of the second workshop on Chinese language processing: held in conjunction with the 38th Annual Meeting of the Association for Computational Linguistics.
[30] Che W, Li Z, Liu T (2010) LTP: A Chinese Language Technology Platform. In: Proceedings of 23rd International Conference on Computational Linguistics: Demonstrations, pp 13-16.
[31] Paltoglou, Georgios, and Mike Thelwall. "A study of information retrieval weighting schemes for sentiment analysis." Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, 2010.

Biographies

LIU Lizhen, received a PhD in Computer Application from the Beijing Institute of Technology, China. She is currently a Professor at the Capital Normal University. Her research interests include text mining, sentiment analysis and knowledge acquisition.She has published in journals and conferences like Knowledge and Information Systems, International Journal of Information & Computational Science, Journal of Software, Journal of Computers, IEEE World Congress on Intelligent Control and Automation, CSCWD.

SONG Wei, received a PhD in Computer Application from the Harbin Institute of Technology, China. He is currently a lecturer at the Capital Normal University. His research interests focus onSentiment Analysis, Personalized Search and Social Computing. He has published his important work on the Proceedings of the ACM SIGIR Conference, Coling and so on. Email:msongwei@gmail.com

WANG Hanshi, received a PhD in Computer Application from the Beijing Institute of Technology, China. He is currently a lecturer at the Capital Normal University. His research interests focus on computational linguistics, sentiment analysis, and natural language processing, especially unsupervised methods in the area. He has published his important work on the famous journal of Computational Linguistics, and other international conferences. Email: necrostone@sina.com.

LI Chuchu, is a master's candidate of Technology of Computer Application at Capital Normal University. Now she is a student in the third year of graduate study. Her research interests include sentiment analysis and machine learning.

LU Jingli, received a doctoral degree in Computer Science from Massey University, New Zealand. She is currently a Software Engineer at Agresearch Ltd, New Zealand. Her research interests include Data Mining and Machine Learning, Fuzzy Set Theory, and Speech Signal Processing. She has published several papers in highly-rankly international journals and conferences.