ISSN2349-4506
Impact Factor: 2.265

Global Journal of Engineering Science and Research Management

TARGET-DEPENDENT SENTIMENT ANALYSIS FOR PRODUCT COMMENTS

LI LIU, YONGHENG WANG*, SHIJUN ZHANG
College of Information Science and Engineering, Hunan University, Changsha 410082, China
E-mail:270818745@qq.com, wyh@hnu.edu.cn, 1187977620@qq.com

KEYWORDS: Product Comments; Sentiment Analysis; MapReduce; Tri-training; Conditional Random Fields(CRF); Syntax Tree Pruning

ABSTRACT
Traditional sentiment analysis method analyzes the whole sentiment polarity of comments without concerning about the relevant targets. Existing target-dependent sentiment analysis usually ignores the multi-target and multi-opinion sentence, which causes wrong target identification. In this paper, we propose a novel target-dependent method based on Conditional Random Fields (CRFs) and syntax tree pruning. A parallel tri-training method based on MapReduce is used to label corpus semi-autonomously. CRF model is used to extract positive/negative opinions and the target of opinions from comment sentences. Syntax tree pruning is used to prune the irrelevant target of opinions and extract the correct appraisal expressions. Finally, a visual product attribute report was generated. Through extensive experiment, the accuracy of the proposed method on sentiment elements and appraisal expression can reach 89% approximately. Which shows our method outperforms traditional methods on both sentiment analysis accuracy and training performance.

INTRODUCTION
In the age of web 2.0, more and more people begin to communicate and express emotion on Internet. Especially, the development of e-commerce has brought great convenience to customers and merchants. Customers can buy what they want at home but they are not sure about the quality, which may make them disappointed at the end. Merchants get rid of the fussy trade procedure and geographic restrictions, but they cannot get feedback from customer directly. Therefore, sentiment analysis for product comments has attracted more and more attention. With the help of high-quality sentiment analysis, customers can make better choices when shopping online and merchants can improve product quality or develop a new marketing strategy.

Traditional sentiment analysis [1,2] is target-independent, namely, it analyzes sentiment polarity of an article or a sentence without considering the target of sentiment, which cannot satisfy the need of customers and merchants. Most existing target-dependent studies [4,5] are based on linguistic template rules and ignore the multi-target and multi-opinion sentences, some of these targets are irrelevant to comments, i.e. Samsung will appear in the Iphone comments, we need to delete these irrelevant targets. In this paper, our goal is to identify sentiment elements by CRF and delete the irrelevant target by syntax tree pruning, finally, the correct sentiment element and appraisal expressions are got.

The rest of this paper is organized as follows: In section 2, we review related work. Section 3, we describe the sentiment elements extraction based on CRF. We introduce the extraction of appraisal expressions in section 4 and the visualization of appraisal expression in section 5. We report in section 6 experiment result and give our conclusion in section 7.

RELATED WORKS

2.1 Target-independent Sentiment Analysis

The target-independent sentiment analysis technology mainly includes sentiment lexicon based method and machine learning based method.

Sentiment lexicon based methods have been studies in several papers. Das et al. investigated the relation between stock and sentiment from stock comments by building sentiment lexicon [1]. They counted the numbers of positive sentiment and negative sentiment respectively, and then calculate the sentiment polarity of stock comments based on these numbers. Jinan et al. used three different scoring strategies to analyze sentiment of the sentence, including TF-IDF, Latent Dirichlet Allocation (LDA) and compute the difference between positive opinion and negative opinion [2]. Nowadays the comment sentences contain lots of network vocabulary, spoken word, and some words that sentiment lexicon does not contain, which makes the sentiment lexicon based method difficult to work.

The methods based on Machine Learning regard sentiment analysis as a special text classification, which used lots of labeled corpus to train machine learning models and then used the models to analyze the unlabeled corpus. Pang et al. extracted unigram, bigram and word part of speech as classification features, and chose Naive Bayes, Maximum entropy and support vector machine(SVM) as classification models [3].

2.2 Target-dependent Sentiment Analysis

Compared to target-independent sentiment analysis, the target-dependent sentiment analysis mainly focuses on extracting sentiment elements (positive/negative opinion and the target of opinion). Hu et al. proposed an association rules based method to extract noun and noun phrase [4]. They considered noun and noun phrase as attribute words (comment targets) and adjective as opinions. Popescu et al. improved the method in [4] by proposing a method based on PMI, which can get rid of the noun and noun phrase that do not belong to attribute words by computing the PMI between noun and meronymy discriminator [5]. However these methods are all unsupervised methods and the extraction of attribute words and opinion words are separate, which means they ignore the relation between attribute words and opinion words.

Some studies have proposed supervised learning to analyze web sentiment. Jin et al. proposed a novel lexicalized Hidden Markov Models (HMMs) based learning framework for web opinion mining which regarded the extraction of opinion words and attribute words as a sequence labeling task [6]. But it is a generative model and hard to integrate the various features. In our work we choose a discriminative model CRF to mine opinions from comments. Zhang et al. proposed CRF based on syntactic tree structure [7]. Besides linear-chain structure CRF, conjunction structure CRF and syntactic tree structure CRF are also investigated [8]. These two methods modified CRF model to improve extraction precision of opinion words and attribute words.

Ding et al. combined CRF and domain ontology, their method regarded domain ontology as semantic feature [9]. Zhu et al. combined CRF and genetic algorithm to mine opinions from comments and their method utilized the genetic algorithm to extract the best feature collections from the semantic features of emotional collections [10]. These two methods enriched and optimized features in CRF.

Zhang et al. proposed an efficient active learning method to label training corpus, which considered not only syntax tree but also conjunction word and comparisons word [11]. Yang et al. proposed a Co-CRFs model which combined Co-training and CRF [12]. The Co-training method solved the difficulty of collecting a large number of labeled training corpuses.

The above-mentioned studies made full use of various features to increase the extraction precision of opinion words and comment targets, and investigated corpus annotation semi-autonomously. But it is still difficult to annotate corpus automatically for large-scale corpus efficiently. Most important of all, most of the current methods did not deal with the multi-target and multi-sentiment sentences. Our goal is to develop an efficient and scalable corpus annotation method and remove the irrelevant comment targets and opinion words to generate a visual product report.

SENTIMENT ELEMENTS EXTRACTION BASED ON CRF

3.1 Conditional Random Fields

Conditional random fields (CRF) is a probability statistics model for sequence labeling which is first proposed in [13]. It combined HMMs and maximum entropy Markov models (MEMMs). When we input an unlabeled observation sequence, it can output corresponding labeled sequence by computing the joint probability of the whole label sequences given an observation sequence. The computing is based on the following formula:

$$\mathrm{P}(Y \mid X) = \frac{1}{Z(X)} \exp\Big(\sum_{\mathrm{i}} \sum_{k} \lambda_k f_k(y_{i-1}, y_i, X, i)\Big) \tag{1}$$

The normalization factor that makes the probability of all label sequence sum to one can be computed as:

$$\mathrm{Z}(X) = \exp\Big(\sum_{\mathrm{i}} \sum_{k} \lambda_k f_k(y_{i-1}, y_i, X, i)\Big) \tag{2}$$

In these two formulas, $\mathrm{X}(x_1, x_2, x_3 \ldots\ldots \mathrm{x}_n)$ is the observation sequence, $\mathrm{Y}(y_1, y_2, y_3 \ldots\ldots \mathrm{y}_n)$ is corresponding label sequence. $\mathrm{f}_k(y_{\mathrm{i}-1}, y_i, X, i)$ is a feature function based on our input sequence and label positions i and $\mathrm{i}-1$. λ_{k} is a weight factor of the feature function, and can be estimated by maximum likelihood. It reflects the model's confidence of the corresponding feature function. We can get the label sentence when the joint conditional probability reaches the maximum value.

3.2 Semi-automatical Corpus Annotation

Since it is expensive to label corpus manually, we propose a parallel tri-training model based on MapReduce to label corpus semi-autonomously. We collected product comments about Iphone and Levono computer from Jingdong Mall.

Tri-training is a semi-supervised learning method which is proposed by Zhou et al. [14]. Being able to label large-scale corpus using a few labeled corpus, it needs neither sufficient and redundant view, nor different classification algorithms. MapReduce is a software framework for easily writing applications which process vast

Global Journal of Engineering Science and Research Management

amounts of data in-parallel on large clusters of commodity hardware. It is made up of map and reduce functions and the input and output of two functions are all in <key, value>pairs. In this paper, we combined tri-training and MapReduce to improve the efficiency and scalability of training. The whole tri-training is iterative which means it cannot be paralleled. The every step of whole iterative process is serial and needs to handle lots of corpus. Therefore we can deal with every iterative process in parallel as following:

1) Define label sets.

Table 1. The Definition of Label Sets

Label	Corresponding Sentiment Elements
CT	Comment targets, such as color, quality and pixel, etc
PO	Positive opinion words
NO	Negative opinions words
BG	Background words that do not belong to any of above categories words

2) Label a few corpus like below:

外观/CT 很/BG 漂亮/PO ，/ BG 就/BG 是/BG 质量/CT 太/BG 差/NO 了/BG.

(Appearance is very pretty, but the quality is too poor.)

Then utilize random sampling algorithms to extract three data sets D1, D2, D3 from the labeled corpus.

3) Choose Naive Bayes as the classification algorithm to train D1, D2, D3, and then generate three different annotation models M1, M2, M3.

4) Use M1, M2, M3 to label segmented corpus which contains two parallel processes.

One process is the parallel of sentence annotation. Our method needs to compute the probability of each word belong to any label categories. With the increasing of training corpus, the compute process become more and more complex, which is worth to be parallelized. The process contains Map and Reduce phrases. The input and output of each phrase are all in the form of <key, value> pairs. In map phrase, the input of map function is word and label category, and the generated result pairs are <word, probability>. The reduce function receives the result pairs and combines the pairs that have common key, forming a list <word, probability1, probability2, probability3, probability4>. At last, the label category that has the highest probability to label the word is extracted. Through the process, every word is labeled.

The other process is the parallel of three annotation models. In map phrase, we input word and annotation models where the key is word and the value is annotation models. Pairs like <word, label category> are generated in this phrase. In reduce phrase, the reduce function combines the pairs that have common key to form a list <word, category1, category2, category3>. We employ the voting method to label the word. For example if the label category1 and label category2 are the same, we will use category1 to label the word, and put the labeled word into data set D3. In the end, we record the labeled word and train the new data sets D1, D2, D3 again. We repeat the above process until all corpus are labeled.

3.3 Feature Selection

In our work, we select word, word part of speech (POS), dependency parsing ,domain ontology and sentiment information feature as our CRF feature.

Word: The feature is the word of segmented comments which is our observation sequence in CRF.

POS: The feature is the part of speech of current word. The POS of sentiment elements that we want to extract are different. Generally, comment targets are noun, opinion words are adjective. POS is the key to identify sentiment elements and we can extract this feature by Stanford Parser.

Dependency parsing: Previous studies [15] have used this feature. It is difficult to identify some special words in a sentence. But if we can identify the word that has relation with the special words, the special words will be identified easily. We define the feature as shown in table 2. We will refer to this feature as DeP in our result table.

Domain ontology: This feature can identify opinion target of different types,such as product attributes, product name, etc.

Sentiment information: In order to distinguish positive opinion and negative opinion better, we use sentiment lexicon to match them.

In table 3 we list the feature of domain ontology and sentiment information. We will refer to these feature as Do and SeI in our result table.

Table 2. Dependency Parsing Feature

Dependency Parsing	Father node	The father node of current word
	Pos of father node	Pos of father node
	Dependency relation	The relation between current word and its father node

Table 3. Feature of Domain Ontology and Sentiment Information

Feature	Feature Information	Representation
Domain Ontology	Product attribute	Employ 1 as the feature
	Product brand	Employ -1 as the feature
	Other words	Employ 0 as the feature
Sentiment Information	Positive opinion	Employ 1 as the feature
	Negative opinion	Employ -1 as the feature
	Other words	Employ 0 as the feature

APPRAISAL EXPRESSIONS EXTRACTION

After extracting comment targets and opinion words, we need to extract appraisal expressions, namely, extract the relevant opinions for comment targets. Liu et al. proposed adjacent methods to extract appraisal expressions [16]. They centered on opinion words and identified the comment targets of opinion words in a given window. We employ similar approach to extract appraisal expressions. However, the limit of multi-target and multi-opinion sentences in product comments and window size causes low precision of appraisal expressions. So we further propose a syntax tree pruning method to extract appraisal expressions.

4.1 Building Domain Ontology

The domain ontology is a formal, explicit specification of a shared conceptualization, which can represent domain

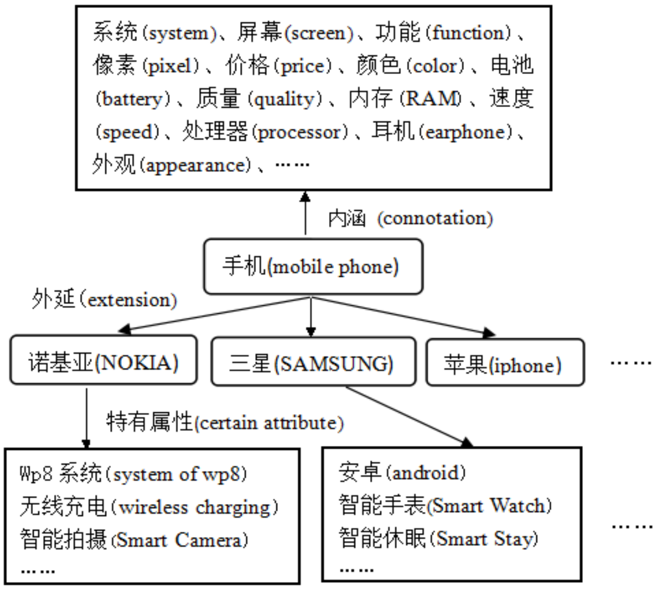
ISSN2349-4506
Impact Factor: 2.265
Global Journal of Engineering Science and Research Management

knowledge and facilitate knowledge sharing. We structure domain ontology about mobile phone in figure1 using the method of previous studies [17]. In figure 1, "mobile phone" is a concept and the node above mobile phone means connotation, namely, the mobile phone has attributes such as "screen, software, hardware ", etc. The nodes under mobile phone are extension, such as NOKIA, SAMSUNG, etc. "wp8 system" belongs to the certain attribute of NOKIA. Each entity of extension has competitive relation with other entity, e.g., the relation of NOKIA, SAMSUNG and iphone are competitive.

4.2 Building Syntactic Path Library

Building syntactic path library is to extract syntactic path between comment targets and positive/negative opinion from syntax tree. The syntax tree of sentence "三星是很漂亮，但上档次的还是苹果(SAMSUNG is very pretty, but the advanced is still iphone)" is shown in figure 2. There are two opinion words and two comment targets in the syntax tree which can form four syntactic paths. For example, "三星(SAMSUNG)" is a comment target and "漂亮(pretty)" is an opinion word. The syntactic path among the two words is "NR→NP→IP→VP→VP→VP→VA". The previous study has found a definite regularity by counting these syntactic paths in large-scale corpus, the syntactic paths of correct appraisal expression are more than wrong appraisal expression[18].

After extracting syntactic paths, we need to generalize them, namely, merge the syntactic paths that have tiny differences and generate a representative syntactic path. For example, the syntactic path NR→NP→IP→VP→VP→VP→VA can be generalized NR→NP→IP→VP→VA. When generalized all syntactic paths, we need to sort syntactic paths by their frequency, and then remove syntactic paths that have low frequency according to predefined threshold. At last, put the remaining syntactic paths into syntactic path library.



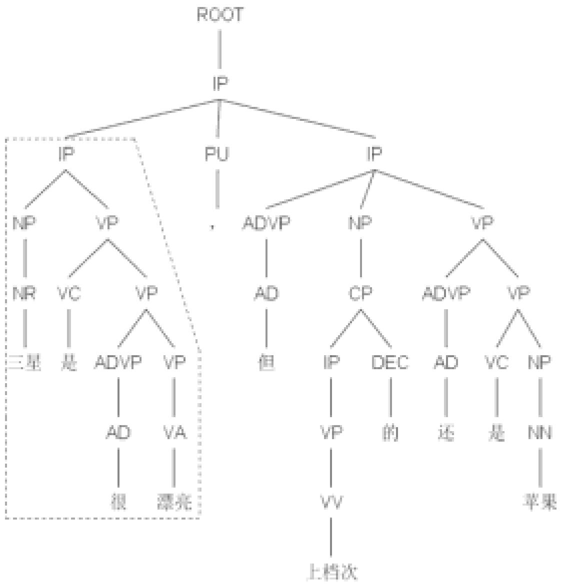


Fig. 1. Domain Ontology of Mobile Phone

Fig. 2. Syntax tree and Syntactic path

4.3 Syntax Tree Pruning

There are many multi-target and multi-opinion sentences in product comments, in which some target entities are irrelevant to product entities. We employ syntax tree pruning to remove the irrelevant target entities.

ISSN2349-4506
Impact Factor: 2.265
Global Journal of Engineering Science and Research Management

The pruning process can be described as below:

1) Employ positive/negative opinion and the target of opinion that extracted by CRF to constitute opinion set and comment target set.

2) Since many comment targets in target set are irrelevant to product entities, we extract these irrelevant targets by matching domain ontology, and put them into candidate pruning set. The targets that belong to domain ontology and have competition with product entities also should be putted into candidate pruning set.

3) For the targets of candidate pruning set, we extract their syntax path with opinion word, and then identify the target-related opinion word by matching syntax path library.

4) According to irrelevant targets and related opinion that step 3 extracted, we separately find their position in syntax tree, and then find their common father node. In the end, we prune the sub-tree that contains comment target and opinion word under the father node. If father node has no other sub-tree, we will prune the parent node. An example is shown in figure 2.

In this example, the comment targets are "三星(SAMSUNG)" and "苹果(Iphone)", while opinion words are "漂亮(pretty)" and "上档次(advanced)". From domain ontology, we can find "三星(SAMSUNG)" is irrelevant to product entities. Therefore we put it into candidate pruning set. The syntax path between "三星(SAMSUNG)" and "漂亮(pretty)" is NR→NP→IP→VP→VA. The syntax path between "三星(SAMSUNG)" and "上档次(advanced)" is NR→NP→IP→NP→CP→IP→VP→VV. After searching syntax path library, we can find the first syntax path has high possibility than the second syntax path. So we can consider "漂亮(pretty)" is the relevant opinion word of "三星(SAMSUNG)". We find the common parent node "IP" of both "三星(SAMSUNG)" and "苹果(iphone)", and remove the sub-tree under the parent node which contains "三星(SAMSUNG)" and "漂亮(pretty)".

After syntax pruning, the appraisal expression about product entities can be extracted by combining comment targets and opinion words that CRF identified.

VISUALIZATION OF APPRAISAL EXPRESSION

A visual product attribute report about iphone is shown in figure 3. All positive/negative opinions about one comment target are identified which is just like "review summary generation" [8].

外观（appearance）：
正面评价：漂亮，好看，大气……
(positive opinion: pretty, beautiful, marvelous)
负面评价：土，难看，一般……
(negative opinion: rustic, ugly, general)

Fig. 3. Visual attribute report of Lenovo computer

EXPERIMENT RESULTS AND DISCUSSION

6.1 Corpus Collection and Preprocessing

We collect two types of product comments from Jingdong Mall, one is Iphone comments, which contains 9735

positive comments and 6251 negative comments, the other is Lenovo computer comments, which contains 8931 positive comments and 4906 negative comments.
In addition to deleting some invalid URL and characters from these comments, we extract new words that tokenizer do not identify using NLPIR, which can improve the precision of segmentation.

In order to label corpus semi-automatically, we randomly extract 5% comment corpus from positive/negative corpus separately three times to labeling by manually. Then we use the parallel tri-training method to label the remaining unlabeled corpus and verify by manually. In the end, the two types of comments are divided into six parts separately, we select five parts as training data, the sixth part as testing data.

6.2 Construction of Syntactic Path Library

In this section, we employ the method that section 4.2 has proposed, we get syntax tree of product comments by Stanford Parser and count syntax path between opinion targets and opinion words . Threshold is 5 according to the method of previous study has proposed[18]. So the detail result as table 4.

Table 4. Statistics of Syntax Path

Number	Syntax path	Frequency
1	NN→NP→IP→VP→VA	15,172
2	NN→NP→NN	14,062
3	NN→NP→VP→VV	12,683
4	NN→NP→ADJP→JJ	12,062
5	NN→NP→IP→VP→VV	10,621

6.3 Result of Sentiment Element Extraction

In this paper, we employ CRF model to identify sentiment element, The Language Technology Platform is used to get the feature of word, pos and dependency parsing. The domain ontology feature can be got by domain knowledge, Sentiment information feature can be got through sentiment lexicon of Hownet. we get parse tree by Standford Parser. We employ five methods to identify sentiment elements, the first one is CRF based on syntactic tree structure[7], the second is CRF model combined domain ontology[9], the third is CRF model combined syntactic tree and domain ontology, the fourth added sentiment information feature on the basic of the third method, the fifth added parser tree pruning on the basic of the fourth method. Precision, Recall and F-measure are used to evaluate result. The final result as table 5.

Table 5 Result of Sentiment Elements Extraction

methods	Sentiment elements	Iphone comments			Lenovo computer comments		
		P	R	F	P	R	F
CRF+Word+Pos+Dependency Parsing	Opinion target	81.3	60.5	69.4	80.4	62.1	70.1
	Positive opinion	78.2	69.9	73.8	84.1	73.4	78.4
	Negative opinion	76.1	81.5	78.7	77.3	74.2	75.7
CRF+Word+Pos+Domain Ontology	Opinion target	83.9	74.7	79.0	85.1	72.6	78.4
	Positive opinion	71.6	65.2	68.3	72.5	64.6	68.3

Global Journal of Engineering Science and Research Management

	Negative opinion	76.9	59.8	67.3	73.6	69.8	71.6
CRF+Word+Pos+Dependency Parsing+Domain Ontology	Opinion target	85.6	71.9	78.2	86.8	77.4	81.8
	Positive opinion	80.9	71.7	76.0	74.3	82.1	78.0
	Negative opinion	78.2	80.5	79.3	82.4	75.2	78.6
CRF+Word+Pos+Dependency Parsing+Domain Ontology+Sentiment Information	Opinion target	87.4	78.1	82.5	88.2	72.5	79.6
	Positive opinion	86.1	80.9	83.4	81.5	85.2	83.3
	Negative opinion	76.9	86.5	81.4	83.4	80.8	82.1
CRF+Word+Pos+Dependency Parsing+Domain Ontology+Sentiment Information+Pruning	Opinion target	88.5	89.1	88.8	90.1	86.9	88.5
	Positive opinion	86.3	91.2	88.7	92.5	85.7	89.0
	Negative opinion	92.1	86.8	89.4	93.5	92.8	93.1

Table 5 shows the evaluate result of sentiment elements for five methods, From the first method,we can see the highest precision has reached 84.1% in two product field. But recall as low as 60.5%. The second method combined CRF and domain ontology, which improved the whole recognition rate of opinion target. but the recognition rate of opinion words has reduced. It is obvious domain ontology feature is good at catching opinion target information, the first method is better than second one in identifying opinion words. So we combined two methods, the recognition rate of sentiment elements all reached about 78%. For further improving recognition rate of sentiment elements, we added the sentiment information feature in fourth method, We can see that the extraction of opinion words achieve better result than before. The precision and recall has been improved a lot. This is because sentiment information gives a specific label to every word in advance and CRF receives the information and identifies it easily. But this feature has little influence on comment targets since they have no direct relation. Therefore, on the basic of previous method, we pruned the irrelevant opinion target by syntax tree pruning in fifth method, In the end, the precision and recall of sentiment elements have improved a lot, it is proved the method we proposed is better than previous methods.

6.4 Result of Appraisal Expression

Table 6 shows the extraction result of appraisal expression in two methods. It is obviously that precision and recall of adjacent method are low. The main reason is adjacent method pays more attention on experience and ignores syntactic structure. Through syntax tree pruning, the precision and recall has been improved a lot which is because we removed the irrelevant comment target with product entities and avoided the interference. Finally, The appraisal expression we got can form the visual product report.

Table 4. Extraction result of Appraisal Expression

Methods	Iphone comments			Lenovo computer comments		
	P	R	F	P	R	F
Adjacent Method	72.3	68.5	70.3	74.6	70.9	72.7
Syntax tree Pruning	90.4	88.7	89.5	87.3	90.2	88.7

ISSN2349-4506
Impact Factor: 2.265
Global Journal of Engineering Science and Research Management

CONCLUSION

In order to extract valuable information from product comments, we propose a target-dependent sentiment analysis method based on CRF and syntax tree. We employ a parallel tri-training method based on Map-Reduce to label these product comments, which makes it has good scalability for large-scale data set. Besides, we verify the labeled corpus by manually. After labeling corpus, we integrate rich features to identify positive/negative opinion and the target of opinion by CRF. We employ syntax tree pruning to remove irrelevant opinion target and get the correct opinion target and appraisal expression. Finally, a visual product report is formed. From experimental result, we can see that our method present a certain superiority comparing to traditional method.

In future work, we will extract more useful features to identify sentiment element and improve sentiment analysis. We will also consider how to handle the comparative sentences.

ACKNOWLEDGES

First of all, this work is supported by the National Natural Science Foundation of China (No.61371116).
Second, I would like to express my gratitude to my supervisor, Yongheng Wang, for his instructive guidance and precious suggestion on my paper. I also appreciate my classmate of our lab, for their generous help. Finally, I want to express my gratitude to my family, for their supporting all the time.

REFERENCES

[1] Das S, Chen M. Yahoo! for Amazon: Extracting market sentiment from stock message boards[C].In: Proceedings of the Asia Pacific finance association annual conference (APFA). 2001, 35: 43.
[2] Jinan F, Osama M, Sabah M, et al. Opinion Mining over Twitter space: Classifying Tweets Program matically using the R Approach[C].In: Proc.of the 7th International Conference on Digital Information Management. Macau, China: IEEE Press, 2012.
[3] Pang B, Lee L, Shivakumar V. Thumbs up? Sentiment Classification Using Machine Learning Techniques [C] //Proc. of Conference on Empirical Methods in Natural Language Processing: ACM Press, 2002.
[4] Hu M., Liu B. Mining and Summarizing Customer Reviews [C]. In: Proceedings of KDD'2004,2004: 168-177
[5] Popescu A.M., Etzioni O. Extracting Product Features and Opinions from Reviews[C]. In: Proceedings of HLT/ACL'2005, 2005: 339–346
[6] Jin W, Ho H H. A novel lexicalized HMM-based learning framework for web opinion mining[C]//Proceedings of the 26th Annual International Conference on Machine Learning. 2009: 465-472.
[7] Yue Zhang. Finer grained opinion analysis on product reviews [D]. Harbin Institute of Technology, 2013.
[8] Li F, Han C, Huang M, et al. Structure-aware review mining and summarization[C]//Proceedings of the 23rd International Conference on Computational Linguistics. Association for Computational Linguistics, 2010: 653-661.
[9] Ding S, Jiang T. Comment Target Extraction Based on Conditional Random Field & Domain Ontology[C]//Asian Language Processing (IALP), 2010 International Conference on. IEEE, 2010: 189-192
[10] Zhu J, Wang H, Mao J. Sentiment classification using genetic algorithm and Conditional Random Fields[C]//Information Management and Engineering (ICIME), 2010 The 2nd IEEE International Conference

ISSN2349-4506
Impact Factor: 2.265

Global Journal of Engineering Science and Research Management

on. IEEE, 2010: 193-196

[11] Zhang K, Xie Y, Yang Y, et al. Incorporating conditional random fields and active learning to improve sentiment identification[J]. Neural Networks, 2014

[12] Yang L, Liu G, Liu Q, et al. Analyzing Sequence Data Based on Conditional Random Fields with Co-training[C]//Computational Intelligence and Security (CIS), 2012 Eighth International Conference on. IEEE, 2012: 94-98.

[13] Lafferty, A. McCallum, F. Pereira. 2001. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In: Proc. 18th International Conf. on Machine Learning (ICML).

[14] Yuan Gao. A Survey of Semi-Supervised Algorithms with Co-Training Paradigm. Department of Computer Science and Technology, Nanjing University.

[15] Rongyang Wang, Jiupeng Ju, Shoushan Li, et al. Feature Engineering for CRFs Based Opinion Target Extraction [J]. Journal of Chinese Information Processing, 2012, 26(2):56-61.

[16] Liu K., Xu L.H., Zhao J. Opinion Target Extraction Using Word-Based Translation Model[C]. In: Proceedings of EMNLP-CoNLL'2012, 2012: 1346-1356.

[17] Efstratios K, Christos B, Theologos D. Ontology-based sentiment analysis of twitter posts[J]. Expert System with Applications, 2013, 40(10): 4065-4074.

[18] ZHAO Y,QIN B,CHE W,et al.Appraisal Expression Recognition Based on Syntactic Path.Journal of Software, 2011, 22(5): 887-898.