8th International Conference on Modelling, Identification and Control (ICMIC-2016)
Algiers, Algeria- November 15-17, 2016

Modeling on Evaluation Object Extraction in E-commerce Corpus based on Semantic Feature

Kai Gao[1], Shan-shan Zhang[1], Shu Su[1] and Mei Li[1]
[1] School of Information Science & Engineering, Hebei University of Science and Technology, China, 050000
gaokai@hebust.edu.cn
zshanshanmiss@gmail.com
sushupro@gmail.com
41786659@qq.com

***Abstract*—As a newly shopping tool, electronic commerce has been drawing more and more attention of researchers. According to the characteristics of comments diversity, it is necessary to extract evaluation object which is an important component of sentiment information. This paper explores Conditional Random Field (CRF) to do evaluation objects extraction. After observing generally used features in sentiment extraction, this paper conclude all the features into four categories, i.e. word Segmentation, Part-of-speech Tagging (POS), Dependency Parsing, Semantic Dependency Parsing. What's more, focusing on the introduction of new feature semantic dependency is a very vital item in our research. In the experiment, we examine the various features and combinations in the extraction task performance, and make a detailed comparative study. The experimental results confirm that adding the feature of semantic dependency has better performance in terms of the evaluation object extraction.**

Keywords*—CRF; Semantic Dependency Parsing; Dependency Parsing; Part-of-speech Tagging; Word Segmentation; Evaluation Object

I. INTRODUCTION

With the rapid development of Internet and electronic commerce, Internet has become increasingly popular for the important role in opinion sharing and online shopping. People begin to use varieties of web services, publishing and dissemination of information, especially for some shopping websites, such as taobao (i.e., https://www.taobao.com), jingdong (i.e., https://www.jd.com) etc. Paying more attention to mine the hidden information behind the big data is necessary. Although these reviews information is unstructured, it still contains the objective evaluation of user and feedbacks of the products. Furthermore, a growing number of potential shoppers will regard these reviews information as an important reference of purchase products. So mining the useful information behind shopping comments is helpful to understand the opinions of this product. Being unstructured information, it is difficult to mine useful information, so it is necessary to model and extract evaluation objects. Evaluation object is an object that emotional expression of the comment text oriented to. At present, the commonly used approaches on extracting evaluation objects are lexical analysis, syntactic analysis and sequence annotation, etc.

This paper explores Conditional Random Field (CRF) based evaluation objects extraction. The essence of this approach is converting the extraction of evaluation objects into the sequence labeling. Due to CRF could be seen as a supervised learning method, feature selection and annotated training corpus have a great influence on the result of experiment. After employing frequently used features in sentiment extraction, this paper summarizes all the features into four categories, i.e. word Segmentation, Part-of-speech Tagging (POS), Dependency Parsing, Semantic Dependency Parsing. Later, combining with CRF++, this paper implements the extraction of evaluation objects.

II. RELATED WORK

The extraction of evaluation object plays an important role in grasping sentence's emotional tendencies. Generally, the commonly used methods are lexical, syntactic and sequence annotation, etc. Zhou et al. (2013) presents an approach of rule-based to do evaluation objects extraction. Compared with the method of rule-based, Song et al. (2013) presents an extraction approach on product comment targets, which combines pattern matching with semi-supervised learning. Based on word-based translation model (WTM), Liu et al. (2012) applies the relationship of evaluation collocation to extract opinion target. Above all, because of the rule-based method requires a lot of manual intervention, syntax-based approach has been proposed. Based on the syntax, Chen et al. (2013) presents a method to identify the evaluation collocation. Li et al. (2015) extracts sentiment feature based on dependency parsing[6] and summarizes six dependencies modes of evaluation units. According to the interdependence between words, Zhang et al. (2016) proposes a method based on dependency parsing and binary tree. In order to obtain the target, some researchers propose the hybrid method of syntax and other algorithm. Liu et al. (2010) proposes the method, which applies syntactic analysis to obtain the candidate and then combines PMI and NN filtering algorithm to choose the target. Xu et al. (2011) presents a new method based on the contextual clues. This method not only takes account of the syntactic analysis and point of mutual information (PMI), but also builds related dictionaries to extract the evaluation objects. According to the characteristics of the product reviews, Yang et al. (2015) employs a bidirectional bootstrapping approach to extract the evaluation object. Based on context-sensitive, this method decides the target among candidate opinion targets, which is obtained by POS template and initial seeds. Geng et

978-0-9567157-6-0 © IEEE 2016

al. (2016) summarizes speech rule templates to obtain the seed of evaluation objects and extends new property template by Bootstrapping method to filter evaluation objects seed. According to kernel sentences, Liu et al. (2015) presents an improved method to extract evaluation collocation, which extracts evaluation collocation by combining kernel sentences and syntactic dependency. With the popularity of machine learning algorithms, more and more researchers explore CRF based evaluation extraction. Wang et al. (2012) presents a method and this method combines lexical, dependency parsing, relative-position with semantic role. On the basis of the lexical and dependency parsing, Ma et al. (2015) takes account of the feature of whether evaluation to identify in the sentence. Liu et al. (2015) presents a method of fine-grained sentiment analysis and syntax tree pruning. Liao et al. (2014) presents a new method, which incorporates lexicon with features using syntax and semantic. Based on the CRF, this method creates a domain dictionary and combines various features to extract the evaluation object.

From the above, it is obviously that the previous method is either based on the lexical or the syntactic analysis &sequence annotation approach. The traditional sequence annotation methods commonly take the basic characteristics such as lexical, part of speech into account, but they usually do not consider deeply semantic dependency relations. This paper presents a method based on CRF, combining with the lexical, POS, dependency parsing, semantic dependency parsing, to do the research on evaluation objects extraction. The experiment results show the feasible of the present approach.

III. METHODOLOGY

The progress of extracting evaluation objects based on CRF can be summarized as three main parts. First, preparing the opinion-sentences and word segmentation, Part-Of-Speech (i.e., POS) tagging, dependency parsing and semantic dependency parsing are done for the opinion-sentences. Second, dividing the opinion-sentences into two parts, which one is the training set and the other is the testing set. Last, using CRF++ to train and test the CRF model, this paper applies this method in extracting evaluation objects. At the same time, according to the results of experiment to evaluate this method, the experimental frame is shown below (see fig. 1).

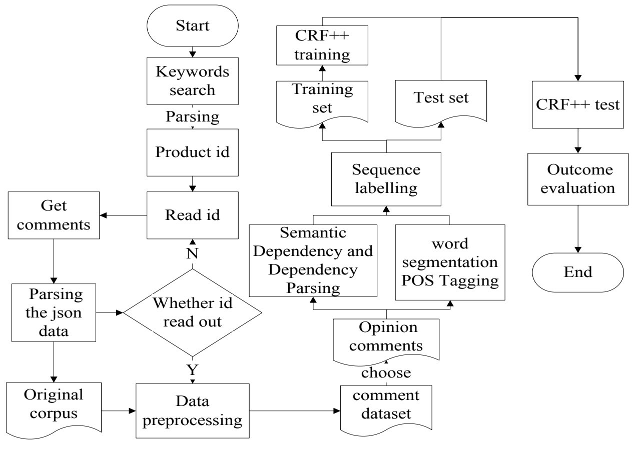

Fig .1.The experimental block diagram

A. CRF model

As one of the commonly algorithms, CRF (Conditional Random Field) is widely used in the field of natural language processing and plays an important role in syntactic analysis, named entity recognition and so on. Virtually, CRF can be seen as a reverse of the hidden Markov model (HMM). Compared with HMM, CRF can be regard as a discriminative model and HMM is considered as a generative model. Besides, CRF always presents in the structure of first-order linear chain, which is based on undirected graph G= (V, E). The structure is shown in fig. 2. $X = (X_1, X_2, ..., X_N)$ is observed sequence and Y is the corresponding output sequence.

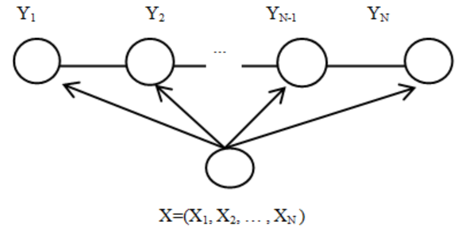

Fig .2.The structure of CRF

CRF model is used to calculate the probability between input and output sequence and gives the corresponding result of sequence X. According to the input sequence $X = (X_1, X_2, ..., X_N)$, the decision of corresponding output sequence Y need to follows the principle and the principle is shown in formula (1). Here, Z(X) can be seen as a normalization factor and the representation of Z(X) is shown in formula (2). $f_i(y_{i-1}, y_i, x, i)$ represents a feature function and λ_i can be considered as the weight of feature function f_i, which is estimated from the train set.

$$P\{Y \mid X, \lambda\} = \frac{1}{Z(X)} \exp\left\{\sum_j \lambda_i \sum_j f_i(y_{i-1}, y_i, x, i)\right\} = \frac{1}{Z(X)} \exp\left\{\sum_j \lambda_j F_j(y, x)\right\} \quad (1)$$

$$Z(X) = \sum_k \exp\left\{\sum_k \lambda_k F_k(y, x)\right\} \quad (2)$$

B. Feature selection

Generally, the traditional feature such as lexical, POS, and dependency parsing are frequently considered. This paper also appends the deeply semantic relation between the various language units into the feature set. The goal of the semantic dependency parsing is to avoid the restriction of syntactic structure, which is the important distinction between the semantic dependency parsing and dependency parsing. This paper presents a LTP (i.e., http://www.ltp-cloud.com) based dependency parsing to get the path of the features and examines the various features, and the results can be divided into four categories, namely, semantic roles, relationship between events and semantic attached tags. In order to get a deeper understanding of semantic dependency parsing characteristic, the result of different syntactic structure is shown in this paper.

For example, as for the given Chinese sentence "*Wo1 Shuai1 Le1 Shou3 Ji1,Jing4 Ran2 Hai2 Neng2 Kai1 Ji1.*" *(我摔了手机，竟然还能开机 i.e., I broke mobile phone and it can be switched on.)*, the corresponding the result of semantic dependency parsing is shown in fig. 3. By changing the syntax structure, the comparative results of "*Wo1 Ba3 Shou3 Ji1 Shuai1 Le1 ,Jing4 Ran2 Hai2 Neng2 Kai1 Ji1.*" *(我把手机摔了，竟然还能开机 i.e., I throw mobile phone and it can be switched on.)* and "*Shou3 Ji1 Bei4 Wo1 Shuai1 Le1,Jing4 Ran2 Hai2 Neng2 Kai1 Ji1.*" *(手机被我摔了，竟然还能开机 i.e., Mobile phone can be switched on, which is broken by me.)* are shown in fig. 4. Compared the result of original sentence with the latter results, it is clear that three different structure of sentences convey the same semantic information. After observing the main relationship among language units, we conclude all the semantic information into two categories, i.e. the sender of action and receiver. In the above example, the word "I" implements the action of "broke" and this action is done for "mobile phone".

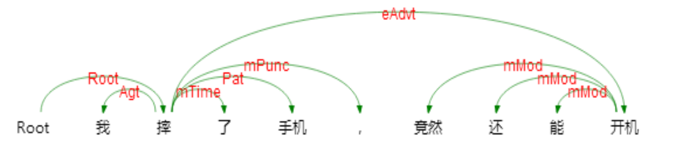

Fig .3.The result of semantic dependency parsing

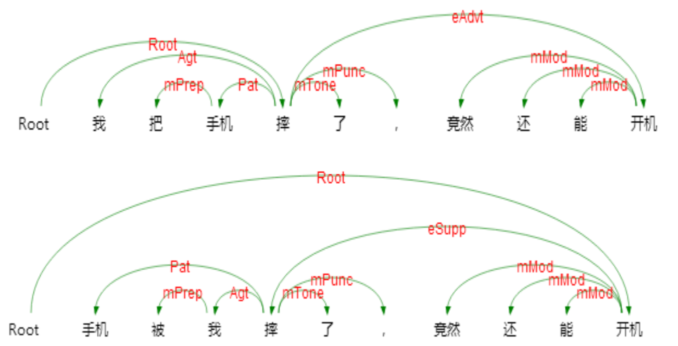

Fig .4.The latter result of semantic dependency parsing

In order to judge the effect of the feature, this paper designs the eight groups of experiments. The experiment combinations are shown in table I. Here, the parameter "WS" means the word segmentation, "POS" means the Part-of-speech tagging, "DP" means the dependency parsing and "SDP" means semantic dependency parsing.

TABLE I. DETAILS OF EIGHT GROUPS OF EXPERIMENT

ID	Feature combinations of the experiment
01	WS
02	WS and POS
03	WS and DP
04	WS and SDP
05	WS , POS and DP
06	WS, POS, and SDP
07	WS, DP and SDP
08	WS, POS, DP and SDP

C. Tagging rules and the feature template

As a kind of machine learning algorithm, CRF++ heavily relies on the annotated data. In order to obtain a more accurate model, this paper chooses the opinion-sentence from the product comments and annotates the evaluation objects among the opinion-sentence. After employing the characters of comments, this paper shows that an evaluation object is likely to be composed of several words. So this paper presents an improved approach based on the model of IOB and the detail of this method is shown in table II.

TABLE II. THE IMPLICATION OF TAGGING SYMBOLS

Symbol	Implication
B_E	The start of evaluation objects
I_E	The middle of evaluation objects
E_E	The end of evaluation objects
B_S	The start of sentiment words
I_S	The middle of sentiment words
E_S	The end of sentiment words
O	Other

This paper uses the tool of CRF++ to extract the evaluation objects. As for the generation of model, it not only needs the standard form of input data, but also combines the command and the template.

Based on the CRF, there are two kinds of template: one is the Unigram template and the other is the Bigram template. The purpose of the Unigram template can be summarized as unigram feature description. In the process of performing Unigram template, the tool will generate L*N feature functions. Compared with the Unigram template, the Bigram template is used to describe bigram feature and L* L* N feature functions will be generated in performing the Bigram template. Here, the parameter "L" represents the number of class and the parameter "N" can be seen as the cardinal number of template. The most essential difference between the Unigram template and the Bigram template is the latter not only take account of the current label, but also consider the label of a moment. Obviously, it is clear that feature counts have a great influence of processing speed. In order to improve the efficiency of experiments, this paper uses the Unigram template and chooses [-2, 2] as observation windows. In the template, each row represents an extensible rule, which is summarized from the various characteristics in training set. Generally, the template is shown in table III, in which the symbol "[row, col]" represents the relative position of the current term and the symbol "row" represents the relative number of rows and the symbol "col" is on behalf of the absolute number of columns.

TABLE III. GENERAL FORMAT OF THE FEATURE TEMPLATE

```
# Unigram
U01:%x[-2,0]
U02:%x[-1,0]
U03:%x[0,0]
U04:%x[1,0]
U05:%x[-1,0]/%x[0,0]
U06:%x[1,0]/%x[2,0]
U07:%x[0,1]
U08:%x[-1,1]/%x[1,1]
# Bigram
B
```

IV. Experiment Results and Analysis

A. The dataset and preprocessing

In this paper, mobile phone review data set is used to extract the evaluation objects, which is crawled from jingdong. In detail, the datasets include total of 3046 species, containing 536596 comments. However, the evaluation object extraction is on the basis of the effective opinion-sentences, so processing the opinion-sentences is of significant. This paper adopts the method of rule-based to filter opinion-sentences. Last, this paper chooses 1050 comments as the experimental dataset from 483843 opinion-sentences and the comments are divided into two parts. One is the training set and the other is the testing set. The proportion of the training set and the testing set is 4:1.

B. Evaluation metrics

This paper adopts the following Precision, Recall and F-Measure metrics to analyze the result. The corresponding formulas are shown in (3), (4), and (5). The *system_total* represents the number of evaluation objects that are identified by CRF++ and the *system_correct* represents the number of the evaluation objects that are extracted correctly in *system_total*. Parameter *total* represents the total number of the evaluation objects that originally belong to test set.

$$Precision = \frac{system\_correct}{system\_total} \tag{3}$$

$$Recall = \frac{system\_correct}{total} \tag{4}$$

$$F\_Measure = \frac{2 * Precision * Recall}{Precision + Recall} \tag{5}$$

C. Experiment results and analysis

The evaluation objects are extracted from the effective opinion-sentences, which are chosen from the comments of mobile phone. At the same time, LTP is used to do word segmentation, POS tagging, dependency parsing and semantic dependency parsing. Extracting evaluation objects are based on the above work.

In order to verify the effect of semantic dependency parsing, this paper designs eight groups experiments, which are based on the lexical and dependency parsing experiment. Eight groups of experiments are divided into four groups contrast experiment. The one is experimental group 1, which is used to judge the effect of word segmentation and semantic dependency parsing. Experimental group 2 is used to observe the effect of semantic dependency parsing and word segmentation & part of speech. On the basis of word segmentation and dependency parsing, this paper designs experimental group 3 to analyze the impact of semantic dependency parsing. Experimental group 4 takes four conditions into account to analyze the influence.

As for the baseline group, the precision, recall and F-measure of the lexical and dependency parsing experiment are shown in Fig.5. Compared with the signal feature of word segmentation, it is clearly that any of the feature combination can make the promotion of the precision, recall and F-measure. What's more, among the four groups of experiments, the combination of word segmentation and part-of-speech can not only make the recall obtain maximal value, but also make the F-measure achieve the best effect. Above all, it can conclude that the effect of combination with part-of-speech is best. This may due to the characteristics of review data.

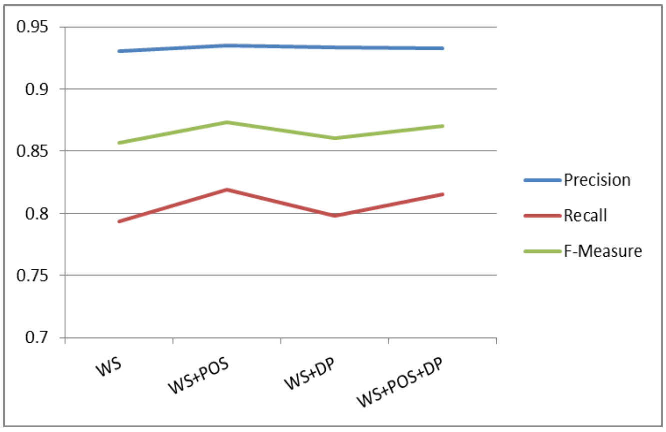

Fig .5.The result of the lexical and dependency parsing

Table IV shows the precision, recall and F-measure of the eight groups of experiment results. In experimental group 1, experimental group 3 and experimental group 4, the feature of semantic dependency parsing can not only improve the recall, but also make the F-measure promotion. In the group1 and the group 4, the feature of semantic dependency parsing improves the F-measure only 0.3%-0.4% and the recall is 1% increased. Making a comprehensive analysis in the three groups of experiments (experimental group 1, experimental group 3 and experimental group 4), the feature combination of experiment group 3 has a better performance. The F-measure of semantic dependency parsing experiment is only 0.48% improved, while the recall is 1.37% improved. In experimental group 2, the feature of semantic dependency parsing increases the precision. In total, the performance of semantic dependency parsing is better. It also fully illustrates that the deeper semantic relationships are suitable for extracting evaluation objects in this short content e-commerce.

TABLE IV. COMPARSION RESULTS OF SEMANTIC DEPENDENCY PARSING AND LEXICAL & DEPENDENCY PARSING

Group name	Feature combination	Precision	Recall	F-Measure
Experimental group1	WS	0.9308	0.7937	0.8568
	WS+SDP	0.9235	0.8074	0.8616
Experimental group2	WS+POS	0.9349	0.8192	0.8732
	WS+POS+SDP	0.9406	0.8094	0.8701
Experimental group3	WS+DP	0.9333	0.7976	0.8601
	WS+DP+SDP	0.9159	0.8349	0.8735
Experimental group4	WS+POS+DP	0.9325	0.8153	0.8700
	WS+POS+DP+SDP	0.9271	0.8251	0.8731

Fig. 6 shows the compared result of the eight groups of experiments. Among the result, this paper can conclude that the combination of word segmentation, POS and semantic dependency parsing can make the precision get maximal value. Combining with table IV, feature combination of word segmentation, dependency parsing and semantic dependency has the best performance on the recall and F-measure. Above all, it is clear that the feature of semantic dependency parsing has a good impact on evaluation objects extraction and different feature combinations of semantic dependency parsing makes different promotion among the precision, recall and the F-measure.

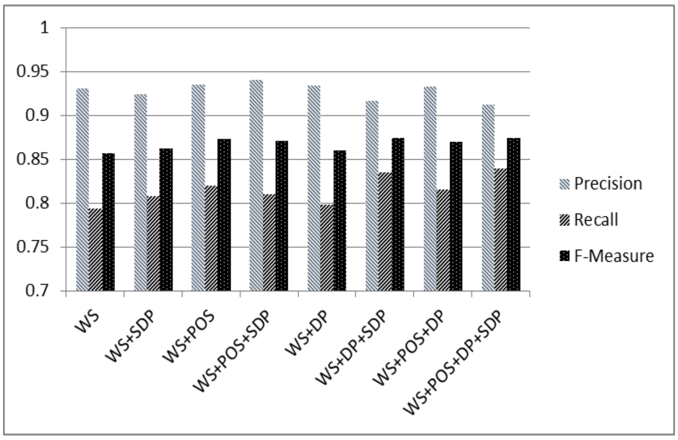

Fig .6.The result of the semantic dependency parsing

V. CONCLUSIONS AND FUTURE WORKS

Evaluation objects extraction is an important component of sentiment information. This paper applies Conditional Random Field (CRF) to extract the evaluation objects from the e-commerce corpus and summarizes all the features into four categories, i.e. word Segmentation, Part-of-speech Tagging (POS), Dependency Parsing, Semantic Dependency Parsing. Experiment results show that the introduction of new feature semantic dependencies has better performance than that of the other feature combination in both recall and F-measure. In the result of the lexical and dependency parsing experiment, the effect of features combination (Part-of-speech tagging, word segmentation and dependency parsing) is superior to the signal feature of word segmentation (WS) and any of the feature combination can make the promotion of precision, recall and F-measure. While, the result of experiment illustrates that the promotion of the recall may leads to the decline in precision. In the further works, it is necessary to explore the method to improve both precision and recall in the introduction of semantic dependencies.

ACKNOWLEDGMENT

This work is sponsored by National Science Foundation of Hebei Province (No.F2017208012), National Science Foundation of China (No.61272362) and Key Research Project for University of Hebei Province (No.ZD2014029). It is also supported by Natural Science Foundation of Hebei Education Department (No.QN2015207).

REFERENCES

[1] X.J. Zhou, X.J. Wan, J.G. Xiao. "Collective opinion target extraction in Chinese microblogs," In Proceedings of the International Conference onEmpirical Methods in Natural Language Processing, 2013, pp.1840-1850.

[2] H. Song and N.S. Shi, "Comment object extraction based on pattern matching and semi-supervised learning," Computer Engineering, vol. 39, 2013, pp.21-26.

[3] K. Liu, L. Xu, J. Zhao, "Opinion target extraction using word-based translation model," Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. Korea: Association for Computational Linguistics, 2012, 1346-1356.

[4] J. Chen, H. Zhang, F.Y. Cao, "Research on identification of evaluation collocation from chinese customer reviews," Computer Engineering and Design, vol. 34, 2013, pp.1073-1074.

[5] Y.G. Li, X.G. Zhou, W. Guo, H.G. Zhang, "Research on methods of Microblogging sentiment feature extraction based on dependency parsing," 2015 International Conference on Intelligent Systems Research and Mechatronics Engineering. China: Atlantis Press, 2015, pp.581-589.

[6] W.X. Che, M.S. Zhang, T. Liu, "Active learning for Chinese dependency parsing," Journal of Chinese Information Processing, vol. 26, 2012, pp.18-22.

[7] J.H. Zhang, M. Weng, X.L. Li, F. Liu, "Extraction of evaluation object based on dependency parsing and binary tree," Computer Technology and Development, vol. 26, 2016, pp.52-55.

[8] H.Y. Liu, Y.Y. Zhao, B. Qin, T. Liu, "Comment target extraction and sentiment classification," Journal of Chinese Information processing, vol. 24, 2010, pp.84-88.

[9] X.Q. Xu and S.P. Lin, "Research on evaluation object extraction from web document," Computer Engineering, vol. 37, 2011, pp.30-31.

[10] X.Y. Yang, G. Xu, X.W. Liao, "Context-sensitive bidirectional bootstrapping approach for opinion target extraction," Computer Engineering and Applications, vol. 51, 2015, pp.143-147.

[11] Y.S. Geng, L.S. Zhang, T. Sun, "A bootstrapping based evaluation object extraction in comments on chinese commodities," International Journal of u- and e- Service. Science and Technology, vol. 9, 2016, pp. 219-228.

[12] T. Liu, P. Zhao, H.T. Liu, X. Ji, "An improved method of evaluation collocation extraction based on kernel sentence,"Computer Engineering, vol. 41, 2015, pp.157-160.

[13] R.Y. Wang, J.P. Ju, S.S. Li, G.D. Zhou, "Feature engineering for CRFs based opinion Target Extraction," Journal Of Chinese Information processing, vol. 26, 2012, pp.56-61.

[14] X.J. Ma, S. Jin, S.T. Yang, "Opinion mining of chinese e-commerce platform products comments – empirical research based on conditional random field," J. Sys. Sci. & Math. Scis, vol. 35, 2015, pp.1327-1346.

[15] L. Liu, Y.H. Wang, H. Wei, "Fine-grained sentiment analysis oriented to product comment," Journal of Computer Applications, vol. 35, 2015, pp.3481-3486.

[16] C. Liao, C. Feng, S. Yang, H.Y. Huang, "A hybrid method of domain lexicon construction for opinion targets extraction using syntax and semantics," Social Media Processing. Beijing: Springer Berlin Heidelberg, 2014, pp.108-116.