Aspects Opinion Mining Based on Word Embedding and Dependency Parsing

Shuting Zhi
School of Computing, Xi'an University of Posts and Telecommunications, Xi'an, China, 710121

Xiaoge Li*
School of Computing, Xi'an University of Posts and Telecommunications, Xi'an, China, 710121
lixg@xupt.edu.cn

Jinan Zhang
Shaanxi Information Center, Xi'an, China, 710121

Xian Fan
School of Computing, Xi'an University of Posts and Telecommunications, Xi'an, China, 710121

Liping Du
School of Computing, Xi'an University of Posts and Telecommunications, Xi'an, China, 710121
Goodwill Information Technology Co. Ltd, Beijing, China, 100000

Zhongyang Li
Faulty of Mathematics, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1

ABSTRACT

Aspects opinion mining is one of the most important issues in the field of Natural Language Processing. This paper proposes an algorithm of aspects opinion mining by combining word embedding and dependency parsing. Firstly, training word embedding and constructing sentiment and aspect lexicon by word embedding. Secondly, using dependency parser to discover the phrases that have dependencies. Thirdly, filtering these phrases according to the sentiment and aspect lexicon, and obtaining language patterns of aspects. Lastly, using these language patterns of aspects to discover all emotion words of every aspect, and computing the sentimental orientation of every aspect. The experimental results on a reviews corpus of a video software show that the precision, recall and F-score of our algorithm achieves to 73.17%, 76.60%, and 74.85% respectively.

CCS Concepts

• **Information systems→Information retrieval** • **Retrieval tasks and goals→Sentiment analysis.**

Keywords

Aspect lexicon; Language pattern; Sentiment analysis; Sentiment lexicon

ICAIP 2017, August 25–27, 2017, Bangkok, Thailand

ACM ISBN 978-1-4503-5295-6/17/08…$15.00

DOI: https://doi.org/10.1145/3133264.3133305

1. INTRODUCTION

Nowadays, aspects opinion mining is one of the most important issues in the field of Natural Language Processing. Since 2003, Nasukawa T et al.[1] put forward the concept of sentiment analysis. Many researchers [2-5] have done a lot of deep researches on sentiment analysis. There are three types of sentiment analysis: (1) text level based; (2) sentence level based; (3) aspect level based. Where sentiment analysis of aspect level is a more fine-grained emotion analysis task, which analyzes the user's emotional tendency to a specific aspect [6]. Reference [7] first proposed a sentiment analysis method based on aspect level. All the high frequency words extracted from user reviews are treated as aspects to mine the emotional tendency of each aspect. Reference [8-9] mined aspect opinion based on Bootstrapping. Reference [10] mined aspect opinion based on translation model of word, and they treated aspect opinion mining as the word calibration problem using a monolingual word calibration model. Reference [11] presented a more refined Bootstrapping, which used the automatic rule to cut out the error results and update the rules to construct sentiment and aspect lexicons. Reference [12] proposed a method for aspect opinion mining based on dependency parsing.

This paper proposes an algorithm for aspect opinion mining based on word embedding and dependency parsing for network reviews. We conduct our experiment on the reviews of a video software to prove the effectiveness of our algorithm.

The rest of this paper is structured as follows. In section 2, we review related algorithms, word embedding and dependency parsing. In section 3, we introduce our system of aspect opinion mining based on word embedding and dependency parsing. Section 4 provides the details of our experiment and discuss the results. We conclude by summarizing our contribution and indicating directions for future research in section 5.[1]

2. RELATED ALGORITHMS

2.1 Word Embedding

Word embedding is an unsupervised technology used to transform a word into a vector, and the distance between vectors can express the semantic similarity between the words. There are two models to get word embedding: Topic-based model and Language-based model.

Based on the topic model, each word is represented as a probability distribution in different subjects. The typical topic model is based on LDA(Latent Dirichlet Allocation). Reference[13] and Reference[14] used the LDA model to learn a predefined number of topics from a large scale corpus, and the probability distribution of each word in different subjects is the semantic vector of the word. Based on the language model, word embedding is generally obtained as a by-product. Reference[15] first used the neural network to build a language model. Because of the many parameters of the model and high training costs, Mikolov et al. proposed the word2vec framework for simplifying the model. There are two basic models in the framework: Continuous Bag-of-Words(CBOW) model, and Continuous Skip-gram (Skip-gram) model. CBOW model predicts the current word based on the context of the current word; Skip-gram model predicts the context of the current word based on the current word.

In this paper, word2vec is used to transform words into real vectors in n-dimensional vector space. We utilize these words embedding to build semantic lexicon and aspect lexicon.

2.2 Dependency Parsing

Dependency parsing describes the syntactic structure of a sentence by parsing the dependencies between the components. Dependency parsing is crucial for computer understanding of natural language, and it is always used in sentiment analysis, question answering system, machine translation and other Natural Language Processing tasks.

There are two main methods of dependency parsing, rule based method and statistical based method. Rule based approach is where humans write the grammar rules to build a knowledge base, and the elimination of syntactic ambiguity is by conditional constraints and checking. This method requires a lot of labor, and the network reviews will always contain a lot of new words, colloquial words, and non-standardized grammar expressions. These characteristics increase the difficulty of the artificial rules. The statistical based method utilizes the principle of statistics to mine the knowledge needed in language analysis automatically based on a large scale corpus. The basic assumption of the statistical based method is based on two points. First, the corpus is the only source of information, and all knowledge can be obtained from the method of constructing statistical models. Second, linguistic knowledge is interpreted in a statistical sense, and all parameters can be obtained automatically from the corpus by statistics or training [16].

In this paper, we use the language technology platform(LTP) developed by Harbin Institute of Technology to realize the dependency parsing of sentences. LTP is an open Chinese Natural Language Processing system; it provides rich and efficient Natural Language Processing technologies such as, Chinese word segmentation, part of speech, named entity recognition, dependency parsing and semantic role labeling[17]. There are 15 kinds of syntactic relations in the dependency parsing module of LTP, and the labels and descriptions of each syntactic relation are shown in table 1. For example, the sentence “蓝色的 UI 很好看，画面特别流畅。(The blue UI is beautiful, and the view is very smooth.)” is processed by LTP and the result of dependency parsing is shown in figure 1. We can see that there is a subject-verb relationship between "UI" and "好看(beautiful)", "画面(view)" and "流畅(smooth)".

Table 1. Types of syntax dependency in LTP

Relationship type	Tag	Describe
subject-verb	SBV	subject-verb
verb-object	VOB	verb-object
indirect-object	IOB	indirect-object
fronting-object	FOB	fronting-object
telescopic	DBL	double
centering relation	ATT	attribute
verb-adverbial structure	ADV	adverbial
verb-complement construction	CMP	complement
parallel relationship	COO	coordinate
preposition-object	POB	preposition-object
left adjunct	LAD	left adjunct
right adjunct	RAD	right adjunct
independent structure	IS	independent structure
punctuation	WP	punctuation
core relationship	HED	head

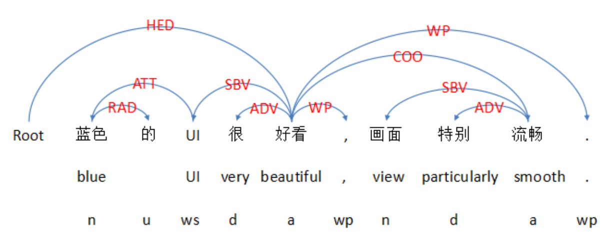

Figure 1. The result of dependency parsing

3. ASPECT OPINION MINING BASED ON WORD VECTOR AND DEPENDENCY PARSING

Figure 2 illustrates the architecture of aspect opinion mining based on words vectors and dependency parsing. The main steps of our system are as follows.

Step 1. Words segmentation for reviews.

Step 2. Training words embedding by Language-based model, i.e., transforming each word in reviews into an n-dimensional real

[1] This work is supported in part by Shaanxi science and technology innovation project foundation under Grant NO. 2016PTJS3-02, NO. 2016KTZDGY04-03 and NO. 2016PTJS3-05, and Xi’an university of posts and telecommunications graduate student innovation foundation under Grant NO. 103-602080017 and NO. 103-602080013.

vector(where n can be set according to the actual situation, such as n=50, n=200, etc.).

Step 3. Constructing sentiment lexicon and aspect lexicon based on words embedding.

Step 4. Using dependency parser to analysis the reviews, and discovering the phrases that have dependencies.

Step 5. Extracting language patterns of aspects according to sentiment lexicon and aspect lexicon.

Step 6. Discovering all emotion words of each aspect from test reviews based on these language patterns of aspects.

Step 7. Clustering aspects by cluster algorithm.

Step 8. Computing the sentimental orientation of each cluster.

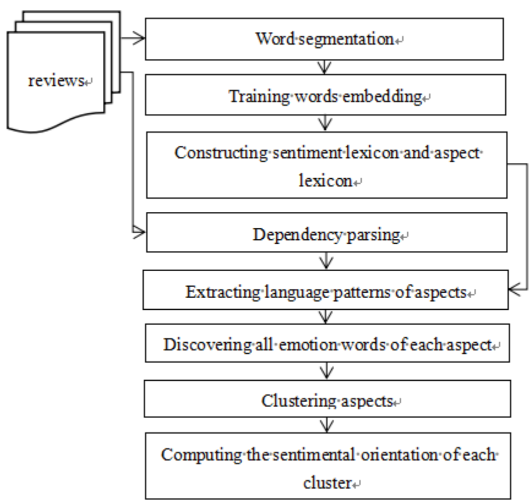

Figure 2. The architecture of aspect opinion mining based on word vector and dependency parsing

3.1 Constructing Sentiment Lexicon

The method of constructing sentiment lexicon described as follows.

Step 1. Delete the words embedding that the words are contained in the stop words list.

Step 2. 10 emotional words are selected as initial set of sentiment seeds.

Step 3. Calculating the distance between each seed and other words based on words embedding.

Step 4. Adding the first 5 words that are closest to the seed to expand the set of sentiment seeds;

Step 5. Executing iteratively step 4 and step 5, until the size of the set of sentiment seeds is no longer increasing.

Step 6. Judging the polarity of emotion words according to the score of each review.

In order to reduce the amount of calculation in step 2 and improve the accuracy of the sentiment lexicon, we delete the words embedding included in the stop word list[18] in step 1, because sentiment words are usually adjectives, adverbs, nouns or verbs[19] and the words in stop list are usually meaningless words or words that are not easily used to form a word.

In this paper, each review has a score, allowing us to judge the polarity of emotion words according to the score of each review in step 6. We think if the score of a review is bigger than a certain value, then the polarity of the review is positive; otherwise, the polarity of the review is negative. We can infer that positive review contains positive emotional words, and negative review contains negative emotional words. Of course, we also considered the influence of negative emotional words.

3.2 Constructing Aspect Lexicon

The method of constructing aspect lexicon is similar to constructing the sentiment lexicon. The only difference is that aspect words are not needed to judge the polarity.

3.3 Extracting language patterns of aspects

Extracting language patterns of aspects is divided three steps. Firstly, extracting all the phrases that have dependencies after reviews are processed by dependency parser. Secondly, filtering the phrases out when each word of the phrase is contained in sentiment lexicon or aspect lexicon. Lastly, and accordingly to aspect lexicon and the result of dependency parser, we transform the phrases met with the requirement to language patterns of aspects. For example, "画面(view)" and "流畅(smooth)" have a subject-verb dependency. While “画面(view)" is contained in aspect lexicon, and "流畅(smooth)" is contained in sentiment lexicon. This allows us to get a language pattern of the aspect "画面(view)", and the language pattern is "aspect={画面(view)} subject={流畅(smooth)} type={subject-verb}".

3.4 Discovering all emotion words of each aspect

According to the language patterns of aspects and sentiment lexicon, we can get all emotional words for each aspect. For example, using the language pattern "aspect={画面(view)} subject={流畅(smooth)} type={subject-verb}", we can find these verbs have the subject-verb dependency with aspect “画面(view)". If these verbs are contained in the sentiment lexicon, we treat these verbs as the emotional words of "画面(view)".

3.5 Clustering aspects

An aspect could be described by many words. In order to describe the emotional tendency of an aspect accurately, we cluster the words describing aspects by cluster algorithm.

3.6 Computing the sentimental orientation of each aspect

For computing the sentimental orientation of each aspect, we assign the weight of of a positive emotional word as 1 and the weight of a negative emotional word as -1. Collect all emotional words of an aspect and compute the sum of the weight of all emotional words. If the sum is positive, then the sentimental orientation of the aspect is positive; otherwise the sentimental orientation of the aspect is negative.

4. EXPERIMENT

The experiment includes three parts: 1) experimental data; 2) evaluation criteria; 3) experimental results and analysis.

4.1 Experimental Data

The experimental data contains 6000 reviews of a video software. We treat the former 5000 as training data for constructing sentiment lexicon and aspect lexicon, and extracting language patterns of aspects. We treat the remaining 1000 as testing data for evaluating the performance of our system.

Table 2 and table 3 describes the initial set of sentiment seeds and the initial set of aspect seeds respectively.

Table 2. Sentiment seeds

Sentiment seeds	喜欢(like), 好用(good), 推荐(recommend), 方便(convenient), 流畅(smooth), 清晰(clear), 垃圾(bad), 闪退(flash back), 黑屏(blank screen), 漂亮(beautiful)

Table 3. Aspect seeds

Aspect seeds	产品(product), 功能(function), 内容(content), 广告(advertisement), 播放器(player), 应用(application), 版本(version), 速度(speed), 界面(view), 流量(network traffic)

4.2 Evaluation Criteria

In this paper, the evaluation metrics are the accuracy(P), recall(R), and F-score(F). The definitions are as follows:

$$P = \frac{\textit{the number of accuracy comment units}}{\textit{the number of comment units}} \times 100\%$$

$$R = \frac{\textit{the number of accuracy comment units}}{\textit{the number of comment units in testing data}} \times 100\%$$

$$F = \frac{2 \times P \times R}{P + R} \times 100\%$$

4.3 Experimental Results and Analysis

In order to validate the performance of our method, experiments are conducted based on the reviews of a video software and we also analyze the following aspects: (1) the influence of the size of the initial set of sentiment seeds on constructing sentiment lexicon; (2) training words embedding based on different models and different strategies, and analyzing the performance of the sentiment lexicon; (3) comparing the performance of aspect opinion mining on short text based on a general sentiment lexicon to a domain sentiment lexicon; (4) constructing aspect lexicon by bootstrapping and by words embedding respectively, and comparing the performance of the two aspect lexicons.

4.3.1 The choice of sentiment seeds

Figure 3 illustrates the influence of different size of the initial set of sentiment seeds on constructing the sentiment lexicon. From figure 3, we can see that whether the size of the initial set of sentiment seeds is equal to 10, 20, 30, 40, or 50. After 4 iterations, the algorithm converges and the sizes of sentiment lexicon are hardly same. The result shows that the size of the initial set of sentiment seeds is not sensitive on constructing sentiment lexicon, and the effectiveness and robustness of this method based on words embedding are verified.

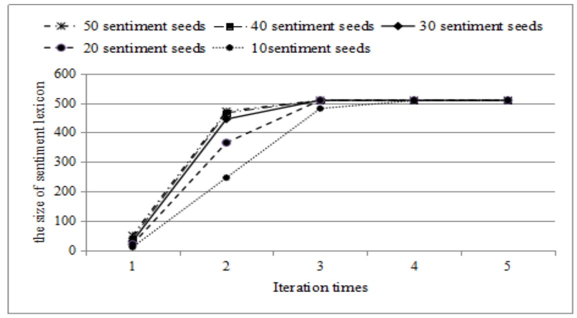

Figure 3. The iteration times on different size of the initial set of sentiment seeds on constructing sentiment lexicon

4.3.2 Constructing sentiment lexicon based on different models and different strategies

Implementing word2vec has two models and two strategies. The two models are the CBOW model and the Skip-gram model. The two strategies are Hierarchical Softmax (HS) and Negative Sampling (NS). Figure 4 describes the performance of sentiment lexicon when using different models and different strategies to implement word2vec. We use different word2vec to train word vectors and to construct the sentiment lexicon. Figure 4 shows the best performance of sentiment lexicon is achieved when Skip-gram model and HS strategy are used to implement word2vec.

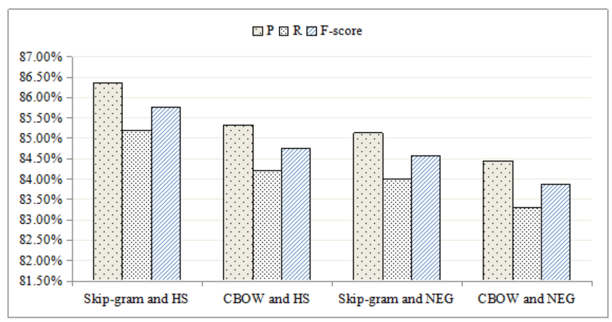

Figure 4. The performance of sentiment lexicon when using different models and different strategies to implement word2vec

From figure 4, we can see that the performance of sentiment lexicon is best when using the Skip-gram model and the HS strategy. In general, the Skip-gram model and the HS strategy can overcome the low frequency, as well as give us the fastest CBOW model.

4.3.3 The necessity of domain sentiment lexicon

Aspect opinion mining is done based on the general sentiment lexicon of National Taiwan University and the domain sentiment lexicon constructed on reviews. The general sentiment lexicon of National Taiwan University contains 2810 positive sentimental words and 8276 negative sentimental words, and the domain sentiment lexicon constructed on reviews contains 268 positive sentimental words and 307 negative sentimental words. Experimenting based on the testing data, figure 5 shows the recall of the general sentiment lexicon is lower than the domain sentiment lexicon. The reasoning is that the coverage rate of the general sentiment lexicon is very low compared to the professional sentiment words in a specific field. Therefore it is necessary to construct domain sentiment lexicon.

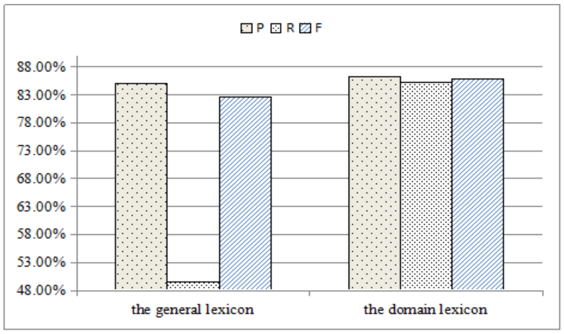

Figure 5. Comparing the general sentiment lexicon and the domain sentiment lexicon

4.3.4 Constructing aspect lexicon

Constructing aspect lexicon based on Bootstrapping and based on words embedding respectively, figure 6 shows the accuracy, recall rate, and F-score of Bootstrapping is 44.81%, 93.62%, and 60.61% respectively. The accuracy, recall rate, and F-score of words embedding is 91.87%, 96.17%, and 93.97% respectively. So the performance of words embedding is better than the performance of Bootstrapping on construction aspect lexicon. From figure 6, we can see that consideration of the semantics will be better when constructing aspect lexicon.

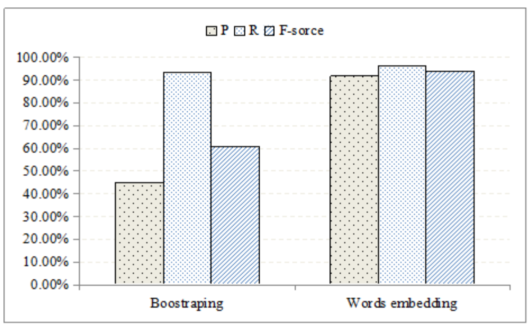

Figure 6. Constructing aspect lexicon based on Bootstrapping and words embedding

4.3.5 Aspect opinion mining

We have chosen the best performance of the lexicons - sentiment lexicon based on words embedding and aspect lexicon based on words embedding. We can combine two lexicons and dependency parser to extract language patterns of aspects. After combining the language patterns of aspects and the aspect lexicon, we could then discover all of the emotional words for each aspect from the test reviews. Next, clustering aspects and computing the sentimental orientation of each cluster. Finally, figure 7 shows the accuracy, recall and F-score of our system are 73.17%, 76.60%, and 74.85% respectively. Hence, we can see that our system is effective.

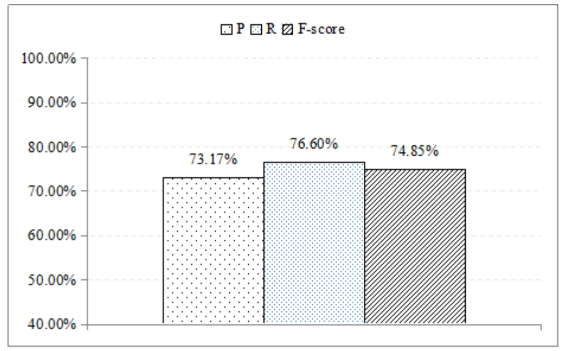

Figure 7. The result of aspect opinion mining

5. CONCLUSIONS

In this paper, we propose an algorithm for aspect opinion mining based on words embedding and dependency parsing for network reviews. We prove the effectiveness of our system based on reviews of a video software, and we also analyzed the following aspects:(1) the influence of the size of the initial set of sentiment seeds on constructing sentiment lexicon; (2) training words embedding based on different models and different strategies, and analyzing the performance of the sentiment lexicon; (3) comparing the performance of aspect opinion mining on short text based on a general sentiment lexicon to a domain sentiment lexicon; (4) constructing aspect lexicon by bootstrapping and by words embedding respectively; comparing the performance of the two aspect lexicons.

The future work is to continue improving the performance of our system, which includes improving the quality of the sentiment lexicon and the aspect lexicon, as well as considering the opinion holder's constraints, and finally, mining more accurate aspect opinions.

6. ACKNOWLEDGMENTS

We thank the anonymous reviewers for their constructive comments. And we thank for some support from Shaanxi science and technology innovation project foundation and Xi'an university of posts and telecommunications graduate student innovation foundation.

7. REFERENCES

[1] Nasukawa T and Yi J. 2003. Sentiment analysis: Capturing favorability using nature language processing. *Proceedings of the 2nd international conference on Knowledge capture, ACM*, 70-77.

[2] Orimaye S O, Alhashmi S M and Siew E G. 2013. Performance and trends in recent opinion retrieval techniques. *Knowledge Engineering Review*, 30(1):1-30.

[3] Pang B and Lee L. 2008. Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2(1-2):1-135.

[4] Medhat W, Hassan A and Korashy H. 2014. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4):1093-1113.

[5] Liu B, Zhang L. 2013. A Survey of Opinion Mining and Sentiment Analysis// *Mining Text Data*. *Springer US*, 459-526.

[6] M Hu, B Liu. 2004. Mining and summarizing customer reviews// *Tenth ACM SIGKDD International Conference on*

Knowledge Discovery and Data Mining, Seattle, Washington, Usa, August. DBLP, 168-177.

[7] M Hu, B Liu. 2004. Mining opinion features in customer reviews// *National Conference on Artifical Intelligence.* AAAI Press, 755-760.

[8] J Zhu, H Wang, M Zhu, BK Tsou and M Ma. 2011. Aspect-Based Opinion Polling from Customer Reviews. *Affective Computing IEEE Transactions on*, 2(1):37-49.

[9] Bagheri A, Saraee M and Jong F D. 2013. An Unsupervised Aspect Detection Model for Sentiment Analysis of Reviews. *Lecture Notes in Computer Science*, 7934:140-151.

[10] K Liu, L Xu and J Zhao. 2012. Opinion target extraction using word-based translation model// *Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*. 1346-1356.

[11] Q Zhao, H Wang, P Lv and C Zhang. 2014. A Bootstrapping Based Refinement Framework for Mining Opinion Words and Targets// *ACM International Conference*, 1995-1998.

[12] Y Wu, Q Zhang, X Huangand L Wu. 2009. Phrase dependency parsing for opinion mining// *Conference on Empirical Methods in Natural Language Processing*, 1533-1541.

[13] Mikolov T, Yih W and Zweig G. 2013. *Linguistic Regularities in continuous space word representations*. HLT-NAACL, 746-751.

[14] Boly-Grader J and Resnik P. 2010. *Holistic sentiment analysis across languages: Multilingual supervised latent Dirichlet allocation*. EMNLP, 45-55.

[15] Bengio Y, Schwenk H, Senécal J, Morin F, Gauvain J and Resnik P. 2006. *Neural Probabilistic Language Models*. Springer Berlin Heidelberg, 137-186.

[16] Qiang Zhou. 1995. *Natural Language Processing technology based on corpus and statistics*. Computer science, 22(4): 36-40.

[17] Zhenghua Li, Min Zhang, Wanxiang Che, Ting Liu, Wenliang Chen, and Haizhou Li. 2011. Joint Models for Chinese POS Tagging and Dependency Parsing. In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, 1180-1191.

[18] Liping Du, Xiaoge Li, Gen Yu, Chunli Liu and Rui Liu. 2016. *New word detection based on an improved PMI algorithm for enhancing segmentation system*. Acta scientiarum naturalium universitatis pekinensis, 52(1):35-40.

[19] Mikolov T, Karafiát M, Burget L, Cernocký J and Khudanpur S. 2010. *Recurrent neural network based language model*. Interspeech, Conference of the International Speech Communication Association, Makuhari, Chiba, Japan, September, 1045-1048.