Detecting Spam Review through Sentiment Analysis

Qingxi Peng and Ming Zhong*
State Key Lab of Software Engineering, Wuhan University, Wuhan, China
pengqingxi@gmail.com, clock@whu.edu.cn

***Abstract*—Online review can help people getting more information about store and product. The potential customers tend to make decision according to it. However, driven by profit, spammers post spurious reviews to mislead the customers by promoting or demoting target store. Previous studies mainly utilize rating as indicator for the detection. However, these studies ignore an important problem that the rating will not necessarily represent the sentiment accurately. In this paper, we first incorporate the sentiment analysis techniques into review spam detection. The proposed method compute sentiment score from the natural language text by a shallow dependency parser. We further discuss the relationship between sentiment score and spam reviews. A series of discriminative rules are established through intuitive observation. In the end, this paper establishes a time series combined with discriminative rules to detect the spam store and spam review efficiently. Experimental results show that the proposed methods in this paper have good detection result and outperform existing methods.**

***Index Terms*—Spam review; Sentiment Analysis; Product Review; Time Series**

I. Introduction

With the recent proliferation of online shopping, customer usually publish reviews to the store and product after online shopping. At the same time, the online reviews become an important approach for the potential customer to know about the store and product. They usually check the online reviews to make decision whether buy the product or not. Meanwhile, sellers and manufacturers are carrying out investigation of online reviews for decision making. However, the reviews grow rapidly, which make it difficult for the potential customer to read the reviews carefully one by one. It also makes it difficult for the seller to keep track and understand customer sentiment. Consequently, sentiment analysis has become a popular topic of many researchers. Many works have been done to summarize the sentiment of online reviews [1-4]. These works examine the polarity of the sentiment in the review, and give sentiment digest and summarization.

In another research field, spam review arise public concern [5-10]. Driven by profit or fame, some people want to influence the user's idea by spurious reviews. They (spammers) publish spurious reviews to promote or demote target online store, inducing users buy or not to buy something from particular store. More and more spam reviews emerge in major review websites such as Epinion.com, Resellerrating.com, and Shopzilla.com. Therefore, detecting such spurious reviews and spammers become a pressing issue. However, previous studies such as [6-10] mainly take advantage of rating or reviewer behavior to detect the spam review. Rating is regarded as representation of reviewer's sentiment orientation. Generally, the 5 star represent the high satisfaction while 1 star means poor satisfaction. Nevertheless these methods have shortcoming by using rating score as indicator. First, the rating will not necessarily completely represent the sentiment of the reviewer. There exist some positive reviews with low rating and some negative reviews with high rating. All these case belong to the inconsistent of the reviews. At the same time, even though two reviews have same ratings, the different content will produce different influence to the reader. The case mentioned above should not be regarded as noisy data. Potential customer will make decision after reading the content carefully. Therefore, the natural language text is more important than rating score to the readers.

Previous works mainly use rating score and the feedback score as indicator to detect the spam reviews. General speaking, the sentiment polarity has never been used to detect the spam reviews before. This paper is the first time incorporating sentiment analysis techniques into spam review detection. Compared with rating score, the content of the reviews will represent more accurate sentiment of the reviewer. Therefore it will indeed influence the potential customer. Three main observations about the rating and content of the reviews are listed below.

(1) The inconsistency between rating and sentiment polarity exist in reviews. There is large number of reviews whose rating and the sentiment polarity is inconsistent. There are two reviews listed in table 1 which are collected from the whole lexicon dataset. It is obviously that the rating and the sentiment polarity of two reviews are contradicted. We choose 1000 reviews randomly from the dataset, and find that 132 reviews belong to this inconsistent case.

(2) The sentiment strength expressed in the review varies considerably. According to [4], the sentiment strength of review "The picture quality is very good" and

*Corresponding Author: Ming Zhong, State Key Lab of Software Engineering, Wuhan University, Wuhan, China 430072.

doi:10.4304/jsw.9.8.2065-2072

the review "I like the camera" are different. Although all the sentiment polarity is positive, the former review has specific description of the product feature. The latter review just has rough evaluation. Therefore they have different sentiment strength.

TABLE 1
RATING AND CONTENT OF THE REVIEWS

Review	Rating	Review Content
A	5	Great! Too bad shipping costs are so high since that essentially doubled the price
B	1	Very good. Very prompt sending when in stock. I really like the way they tell you when you order a book you already have. Keep up the good work.

(3) The sentiment strength differs when two reviews have different number of sentences. For instance, two reviews have the same rating scores. However, one review has lots of sentences, and has detail description about the shopping procedure, usage of the product and service quality as well. The other review has just one or two sentences with simple comment without trustful fact to support. It is obviously that the former will be convincing than the latter.

According to the analysis mentioned above, the content of reviews is more important than rating score. In addition, we can also take advantage of the contradiction of the rating score and the content of the review.

In this paper, we incorporate sentiment analysis techniques into the spam review detection for the first time. The sentiment score is computed from natural language text by a shallow dependency parser. The relationship between sentiment score and spam review is discussed, and a set of discrimination model are proposed. In the end, time series is established to detect the spam reviews.

Our contributions in this paper are then as follows:

1. We propose a method that computes the sentiment score of the natural language text by a shallow dependency parser.

2. We propose a set of discriminative rules by intuitive observation to find unexpected patterns in product review.

3. A time series analysis method is established to detect the spam reviews.

To our knowledge, this is first study detecting spam reviews by means of sentiment analysis. At the same time, the sentiment strength other than polarity is calculated in this paper. The remainder of the paper is organized as follows. Section 2 discusses related work in the area of sentiment analysis and spam review detection. In Section 3, we propose three steps include computing sentiment score, building discriminative rules and establishing time series. In section 4 we present an evaluation on the spam review detection on the real life dataset. Finally, conclusions and directions for future work are given in section5.

II. RELATED WORKS

A. *Sentiment Analysis*

User-generated-content in the Web such as review, blog and microblog usually express the author's emotion. The extraction, analysis and summarization of sentiment become an important research field. In [1], the features of the product are extracted. A summarization system has been built to show the user's sentiment to every feature of the product. In [2], Ding et al. present a holistic lexicon-based approach to analyze the sentiment of both explicit and implicit aspect of product, and realize an opinion mining system. In [3], sentiment analysis is classified at word and sentence level. In [4], comparative sentence was identified, and comparative relationships were extracted from the identified comparative sentence. A series of definition about sentiment, sentiment holder and sentiment polarity are presented in [11]. In [12], SentiWordNet lexical resource has been applied in automatic sentiment classification. In [13], signals have been incorporated into unsupervised sentiment analysis to model two main categories: emotion indication and emotion correlation. In [14], Twitter as a corpus has been automatically collected, and sentiment classifier has been built to determine sentiment polarity. In [15], sentiment analysis has been combined with K-means clustering and support vector machine (SVM) to develop unsupervised text mining approach in the forum hotspot detection and forecast. In [16], classifying sentiment in microblogs has been found easier than in blogs and reviews. In [17], a concept-level sentiment analysis has been seamlessly integrated into lexicon-based opinion mining.

In order to study the sentiment analysis, it is important to extract product features. Two approaches have been proposed to solve this problem. One method is based on dependency parser. According to previous studies, product features are almost noun phrase. Therefore, product features can be extracted by phrase dependency parsing [18-20]. The other method is based on probabilistic topic model such as LDA and PLSA. Through unsupervised learning, not only product features but also their corresponding sentiment can be extracted simultaneously [21-23].

In general, existing studies mainly focus on the analysis, summarization and visualization of the sentiment.

B. *Spam Review Detection*

Spam review detection is an important task in opinion mining. The problem of spam review detection is presented for the first time in [5]. In [6], duplicate and near duplicate reviews are assumed to be fake reviews. Supervised learning has been employed to detect spam review. In [7], the impact of single reviewer to the online store, and anomaly pattern of rating are analyzed to detect the spam reviews. In [8], several characteristic behaviors of review spammers are identified, and these behaviors are modeled so as to detect the spammers. In [9], the unusual review patterns which can represent suspicious behaviors are identified, and unexpected rules are formulated. A novel concept of review graph is proposed in [10], which capture the relationships among all reviewers, reviews and stores that the reviewers have reviewed as a heterogeneous graph. Then an iterative computation model is proposed to identify suspicious reviewers. In [24], three approaches to detect deceptive opinion spam by integrating work from psychology and computational linguistics. The inconsistency problem between the evaluation score and review content is studied, and credibility of customers is detected in [25]. In [26], a frequent item set mining method is proposed to find a set of candidate spam groups. Several behavioral models derived from the collusion phenomenon among fake reviewers and relation models are been used to detect the spam groups. The previous studies mentioned above generally detect the spam review by means of rating score. Sentiment analysis has never been used in spam review detection.

III. The Proposed Methods

Our goal in this paper is to incorporate sentiment analysis into the spam review detection. To achieve this aim, 3 tasks should be considered. The first task is to generate a sentiment lexicon and compute the sentiment score by a shallow dependency parser. The second task is to setup a set of discriminate rules. The third task is to establish a time series method to detect the spam reviews. We proposed different algorithms in the tasks mentioned above. For ease of presentation, we give the notations listed in Table 2.

A. Sentiment Lexicon Generation

In the opinion mining and sentiment analysis research field, there are many sentiment lexicons such as SentiWordNet and MPQA [27, 28].

TABLE 2
NOTATIONS

Notation	Definition
f_i	the i -th feature of product in a review
w_j	the j -th word of a feature
$o(w)$	the sentiment polarity of a word. $o(w) \in [-1,1]$
V	word vocabulary
c_N	the number of negation word in in one feature
d	a review
$\mid d \mid$	the number of sentence of review d
$o(d)$	the sentiment score of review d
$e(d)$	the rating of a review
ST	all the reviews of one store

We choose it as general sentiment lexicon respectively. However, product reviews have their characteristics in natural language text. To improve the accuracy, sentiment lexicon special for product should be prepared in this task.

We extract the reviews from Resellerrating.com from October, 26, 2012 to November 20, 2012. 2000 reviews with four stars to five stars are regarded as positive samples while 500 reviews with one star to two stars as negative samples. We choose Information Gain for feature selection.

Equation (1) is the definition of information entropy. Equation (2) is the Information Gain.

$$entropy(D) = -\sum_{j=1}^{|c|} P(c_j)\log_2 P(c_j), \sum_{j=1}^{|c|} P(c_j) = 1 \qquad (1)$$

$$IG(t) = entropy(D) - entropy(c \mid t) \qquad (2)$$

The Information Gain of the sentiment words in the product reviews are calculated after preprocessing. Then the sentiment words are listed in descending order respectively. We choose 220 positive words and 130 negative words special for product manually. The whole lexicon has been built by SentiWordNet plus sentiment word special for product and MPQA plus sentiment word special for product respectively.

B. Computing Sentiment Score

In this section, the task is to determine the sentiment strength of each review. We first extract the features of each product. First, several definitions are presented as below.

Definition 1 (Sentiment Score, SS): Sentiment score (denoted by $o(d)$) is a score of a review d , which means the sentiment polarity of review. For ease of understanding and computation, we limit the range of $o(d)$ in [-1, 1].

Definition 2 (Sentiment Ratio, SR): Sentiment ratio (denoted by $r(d)$) is a ratio of sentiment sentence to all sentences, $r(d) \in [0,1]$.

Definition 3 (Difference of Sentiment Polarity, DSP): Difference of sentiment polarity (denoted by $f(d)$) represent if there is inconsistency in the rating score and actual sentiment score,

$$f(d) = \begin{cases} 0, & o(d) * e(d) > 0 \\ 1, & \text{else} \end{cases}$$

The extraction methods are proposed in lots of previous works [18-22]. In this section, we first adopt the method in [18] to extract the features and its corresponding sentiment words. The features are extracted and aggregated manually. The complex sentence with multiple features is ignored in this paper. Note that in this paper, one sentence only has one feature. Totally there are nearly 15 features, and averagely by 40 sentiment words accompanied with each one. The strength of the sentiment word is also influenced by the distance between it and the feature. We propose an equation computing feature score considering the negation words. All the word scores are summed up using the following equation (3):

$$score(f) = \sum_{w_j \in f \wedge w_j \in V} (-1)^{c_N} \frac{o(w_j)}{dis(w_j, f)} \qquad (3)$$

In equation (3), $o(w_j)$ means the sentiment polarity of the word w_j. A positive word is assigned the sentiment polarity score of +1, and a negative word is assigned the sentiment polarity score of -1. c_N means the number of negation word in one feature. If there is no negation word, c_N equals 0. While there is one negation word, the polarity of the sentiment is reversed through multiply by -1. $dis(w_j, f)$ means the distance of feature f and sentiment word w_j. When the sentiment word is far from the feature word, the strength of the sentiment weakens, and vice versa. The rating of each review is a linear combination of different feature involved in it.

$$o(d)=\frac{\sum_{f_i\in d} score(f_i)*weight(f_i)}{|d|} \tag{4}$$

$weight(f_i)$ means the weight of feature f_i. For example, there are many features such as value, service, weight, size, and portable in mobile phone. These features have different weight, and the weight is unknown. If the weight of each feature is calculated, we will also be aware of which feature the customers pay more attention to. We give empirical parameters in this paper.

We combine the equation (3) and equation (4) to get the equation (5) listed below.

$$o(d)=\frac{\sum_{f_i\in d} weight(f_i)\sum_{w_j\in f_i\wedge w_j\in V}(-1)^{c_N}\frac{o(w_j)}{dis(w_j,f_i)}}{|d|} \tag{5}$$

C. Building Discriminative Rules

In this section, we present several discriminative rules to find the unexpected patterns.

The rating of a review in the website mainly range from 1 to 5. For ease of computation, we normalize it to [-1, 1]. So the original rating scores are generally transformed to -1, -0.5, 0, 0.5 and 1.

We choose one store #215 as sample. The data include 33 reviews from October, 26, 2012 to November 20, 2012.

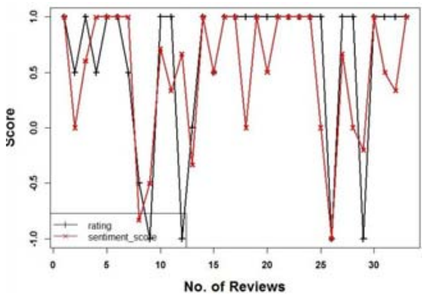

Figure 1. Rating and sentiment score in store#215

Line 1 and line 2 have same tendency. Generally speaking, the rating score and the sentiment score all represent the sentiment of the reviewer. The rating score and sentiment score of review 2, 12, 29, 31, 32 have large difference. We analyze the content of the review. The rating of review 2 is 0.5 while the sentiment score is 0. Let's have a look at the review content:

"I have received several items & have been totally satisfied with the ones I did receive. I am dissatisfied that the first bid that I won a set of rings has yet to arrive! I understand that store#215 is still trying to check on that order."

The expression of first sentence is positive. The second sentence is negative. The third sentence is objective. Although the rating is 0.5, the sentiment score of 0 represent the sentiment of reviewer more accurate.

The rating of review 12 is -1 while the sentiment score is 0.67. The review content is as follows.

"Television ordered not received yet & attempts to track have not been successful (tracking # provided is not correct) Over the years I've purchased many items with far better results. So guess I defer this review until later."

It is obvious that the review is positive to the former store and doubt to it now. The review need verification. Rating of -1 means absolute negation to the store while sentiment score of 0.65 has partly positive factors. So, the sentiment score accurately represent the sentiment of reviewers.

The rating and sentiment ratio of review #215 are listed in figure 2.

In figure 2, the sentiment ratio has a characteristic. The sentiment ratio will drop down while the sentiment score drop down.

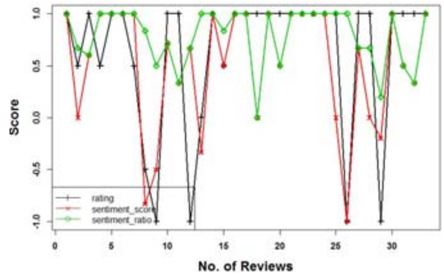

Figure 2. Rating and sentiment ratio in store#215

The reason is that reviewer will give more facts to support his opinion when he gives negative review to the store. While there are lots of negative reviews without enough facts, spam may happen.

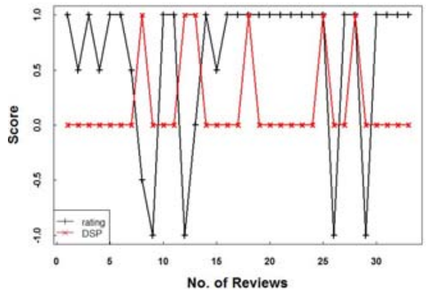

Figure 3. Rating and DSP in store#215

In figure 3, when we find the contradiction of rating and the sentiment score, the review is suspicious.

From the intuitive observation listed above, we can get three discriminative rules to find suspicious reviews in one store. Suppose one store have many sequential reviews, and i represent i-th review which is from s to s'. Here, several discriminative rules are built according to the intuitive observation discussed above.

Rule 1(DSP Rule). If $\frac{\sum_i f(d_i)}{\Delta i}>\xi$, then the store contain these reviews be spam store.

Rule 2(SR Rule). If $\frac{\sum_i r(d_i)/o(d_i)}{\Delta i} > \psi$, then the store contain these reviews be spam store.

Rule 3(SS Rule). If $\sqrt{\frac{\sum_i (o(d_i) - \overline{o(d)})^2}{\Delta i}} > \zeta$, then the store contain these reviews be spam store.

D. Establishing Time Series

In this section, we establish a time series to detect the spam store and reviews. Using time series to detect spam store is firstly proposed in [7]. It depends on rating score as indicator. However, rating score has shortcoming mentioned in section I. We combine the discriminative rules with time series method to acquire better results.

The review records of each store are formulated as d_i<store, reviewer, time, $e(d_i)$, $f(d_i)$, $r(d_i)$, $o(d_i)$>. The time series has been setup. Suppose t_0 is start time, Δt is slide time window.

Algorithm SSD: Store Spam Detection Algorithm
Input: review set $D = \{d_1, d_2 ... d_{|D|}\}$, time span Δt , limit of review number k
Output: spam store and its spam reviews
Begin
1 preprocess the review d_i as

$d_i < store, reviewer, time, e(d_i), f(d_i), r(d_i), o(d_i) >$

2 order the d_i by time
3 For each subset of reviews { d_j } $j = s \rightarrow e$ in Δt
//not enough reviews in every time span
4 if $e - s < k$ return false
5 if d_j satisfy DSP rule then return d_j
6 if d_j satisfy SR rule then return d_j
7 if d_j satisfy SS rule then return d_j
8 end For
9 Return false
End

Figure 4. Algorithm SSD: Store Spam Detection Algorithm

The n -th time window can be represent as $I_n = [t_0 + (n-1)\Delta t, t_0 + n\Delta t], I = \bigcup_{n=1}^{N} I_n$. In [5], three indicators are chosen: average rating, the number of reviews, and the ratio of singleton reviews. In this paper we propose an algorithm to detect the spam store efficiently.

The discriminative rules have been combined into the algorithm SSD which is given below. After the spam store detection, the abnormal time window is zoomed out. The discriminative rules are used again in the shrink time window to find out the spam reviews. This procedure is omitted in presentation.

IV. Experimental Results

In this section, experiments are presented to demonstrate the efficiency of the proposed methods. The product reviews are extracted from Resellerrating.com from October, 26, 2012 to November 20, 2012. We choose 5000 reviews randomly which include all the reviews mentioned in II. The stop words are removed in the dataset. Then we use Porter Stemming tool to stem all the word.

A. Sentiment Lexicon Comparison

In the Table (3), SentiWordNet+Product means the sentiment word comes from both SentiWordNet and product review. MPQA+Product means the sentiment word comes from both MPQA and product review. We simply use sentiment classification accuracy as evaluation metric. The equation (6) shows the accuracy of the product reviews.

$$accuracy = \frac{N_c}{N_t} \qquad (6)$$

In the equation (6), the N_c means the number of product reviews correctly classified, while the N_t means the total product reviews.

TABLE 3
CLASSIFICATION ACCURACY OF DIFFERENT LEXICON

Lexicon	Number of Lexicon	Classification Accuracy
SentiWordNet	6810	56.4%
MPQA	6400	58.5%
SentiWordNet + Product	7030	59.6%
MPQA + Product	6530	61.4%

In table (3), the classification accuracy by using MPQA is higher than SentiWordNet. The classification accuracy all improved by combining the general sentiment lexicon and lexicon special for product. The reason lie in many word only have sentiment factor in product reviews. We make decision that we finish the following task by means of MPQA+Product because it achieves accuracy of 61.4%.

B. Sentiment Score Result

Since there is no dataset for the sentiment score evaluation, we construct a dataset for evaluation. We extract the reviews from resellerrating.com to build the dataset. There are totally 397116 reviewers, 3951 store and 418995 reviews as well. 1000 reviews are chosen from these reviews randomly as dataset. Sentiment classification is carried out to evaluate the sentiment score computation. In table (4), the statistics on rating of dataset is listed.

TABLE 4
STATISTICS ON RATING OF DATASET

Rating	# of Reviews
5	715
4	49
3	9
2	32
1	195

From table (4), we can get a conclusion that a large proportion of reviews are positive or negative reviews. Only a small proportion of reviews are neutral reviews. We split the dataset into 4 subsets. Since rating score of reviews is unreliable in our analysis in section II. We check the reviews manually. To guarantee the reliability of the experiment, three volunteers are selected as reviewers. All of them are familiar with the usage of online reviews. They work independently on the sentiment classification. In order to evaluate the three evaluators' consistency in their judgments, we compute the Cohen's Kappa values of the evaluator pairs. All the Kappa coefficients are larger than 0.75%. We also use accuracy in equation (6) as indicator. Two baselines are chosen to compare with our sentiment score method. One is rating score of the each review. The other is word counting method. The word counting method simply counts the number of positive word and the negative word. It first split the review into sentence. And then the polarity of a review $o(d)$ is defined in equation (7) as follows. n_{pos} means the number of positive word in review d while n_{neg} means the number of negative word in review d.

$$o(d) = \begin{cases} 1, & n_{pos} > n_{neg} \\ 0, & n_{pos} = n_{neg} \\ -1 & n_{pos} < n_{neg} \end{cases} \tag{7}$$

In table (5), the accuracy of sentiment classification show different result of three methods.

In this experiment, we only consider the sentiment polarity of reviews. As the result, all the reviews are transformed into positive review, negative review and neutral review as well. For the rating method, reviews with 4 or 5 stars are marked as positive reviews; reviews with 1 or 2 stars are marked as negative reviews; reviews with 3 stars are marked as neutral reviews. The sentiment score method can accurately calculate the score of the reviews. But we also transform the sentiment score to 3 grades. Reviews with score less than 0 are marked as negative reviews. Reviews with score equal 0 are marked as neutral reviews. Reviews with score large than 0 are marked as positive reviews.

TABLE 5
CLASSIFICATION ACCURACY OF DIFFERENT METHODS

Dataset	Classification Accuracy		
	Rating	Word Counting	Sentiment Score method
Subset #1	80.6%	76.7%	85.7%
Subset #2	81.5%	74.8%	86.3%
Subset #3	79.5%	75.6%	82.6%
Subset #4	80.4%	77.2%	84.5%

In table (5), the accuracy of sentiment score method is larger than that of rating. We check the difference of the methods and find out that some rating of the reviews contradicts the sentiment expressed in natural language text. These reviews have been mentioned in section I. The word counting method has low accuracy since it ignores the negation words in the reviews, and will result in more mistakes in sentiment classification.

C. *Spam Review Detection Result*

In this section, we incorporate discriminative rules into time series method to detect the spam store and the spam review. Both the number of reviews and reviewers in our data are extremely large. It is very difficult to manually labeling all the data. Evaluation based on small sampling is mainly used in information retrieval, which usually use several quires to evaluate the results from search engine. This method is also adopted in our experiment.

According to the previous researches [5-10], evaluating the detection of the spam review mainly depends on manual work. Three volunteers were recruited as evaluators to read the reviews. They use their intuitions to make the judgments. We choose the store with more than 500 reviews as samples. 30 days are determined to the time window. Our method takes advantage of algorithm SSD to detect the spam reviews. We choose the method in [7] as baseline. Firstly, multidimensional time series have been established. On each dimension, a Bayesian change point detection algorithm is employed to fit curves. Then a simple template matching algorithm is applied to detect bursty patterns. We choose top 50 suspicious stores to check it manually.

There are several challenges in spam review detection. The evaluators work independently on spam detection. The selected store is randomly ordered before they are forwarded to the evaluators. If two of the evaluators vote the store as spam store, we regard it as spam store. In table (6), manual evaluation result of the spam stores are listed.

TABLE 6
MANUAL EVALUATION RESULT OF THE SPAM STORE

	Evaluator 1	Evaluator 1	Evaluator 1
# Spam Store			
Evaluator 1	35	33	29
Evaluator 1		31	28
Evaluator 1			35

We compute the Cohen's Kappa equation to evaluate the consistency in their judgments. According to [29, 30], the results are all above 0.65 which indicate a substantial agreement. Therefore, the judgments among evaluators are consistent and effective and thus be used as the following experiments.

D. *Case Study of Spam Review Detection*

The store#364 has been identified as spam store by using proposed method. However, it has not been identified as spam store by using rating. In time window [39, 42], the sentiment score decline sharply. At the same time, the rating also decline. The rating does not decline sharply. So, according to the baseline method, the spam store has not been detected. However, by using our method, the discriminative rule DSP can easily identify the spam store. We check the review manually, and find out that the number of negative reviews increase smoothly in the four time windows. Thus, the store has not been identified as spam store according to baseline method. However, appreciable quantity of reviews with 4 or 5 stars have negative words, which result in sharp decline of the sentiment score. Spammer may possibly try to avoid the detection, which leads to the situation mentioned above.

The next step is to identify the spam review from the spam store. We zoom out the reviews in the abnormal time window. The discriminative rules are used to detect the spam reviews.

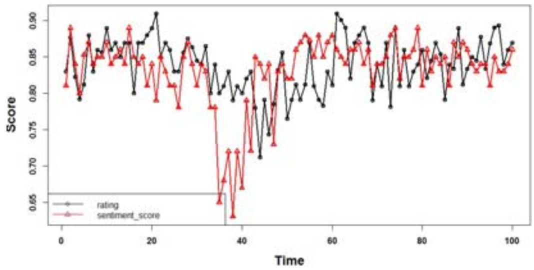

Figure 4 Rating and sentiment score of store#364

IV. Conclusion and Future Works

In this paper, we incorporate sentiment analysis techniques into the spam review detection. Firstly, a sentiment lexicon combined with general sentiment lexicon and sentiment lexicon special for product has been built. Then a method has been proposed to compute the sentiment score from the natural language text by a shallow dependency parser. A set of discriminative rules are presented through intuitive observation. The discriminative rules are combined with the time series method to find out suspicious stores. Furthermore, the spam reviews can be also identified by the discriminative rules from the abnormal time windows. The experiment and the case study demonstrate the efficiency of the proposed methods.

Future works include the improvement computation of the sentiment score in consideration of the modifier word and the expansion of the discriminative rules. In-depth researched should be explored by means of proposed method.

Acknowledgment

This work is supported by “The Fundamental Research Funds for the Central Universities”, No. 201121102020005 and National Science Foundation of China under grant No. 61202036.

References

[1] Hu Minqing, Liu Bing. Mining and summarizing customer reviews[C]// Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM Press , 2004:168-177

[2] Ding Xiaowen, LIU Bing, YU Philip S, et al. A holistic lexicon-based approach to opinion mining[C]// Proceedings of the 2008 International Conference on Web Search and Data Mining. New York: ACM Press , 2008:231-240

[3] Kim S-M, HOVY E. Determining the sentiment of opinions[C]//Proceedings of the 20th International Conference on Computational Linguistics. Philadelphia: Association for Computational Linguistics,2004:1367

[4] Jindal, Nitin, and Bing Liu. Mining comparative sentences and relations. AAAI. Vol. 22. 2006..

[5] Jindal Nitin, LIU Bing. Review spam detection[C] // Proceedings of the 16th international conference on World Wide Web. Canada. New York: ACM Press ,2007: 1189-1190.

[6] Jindal Nitin, Liu Bing. Opinion spam and analysis[C]// Proceedings of the 2008 International Conference on Web Search and Data Mining. New York: ACM Press , 2008:219-230.

[7] Xie Sihong, WANG Guan, LIN Shuyang, et al. Review spam detection via temporal pattern discovery [C] // Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery And Data Mining. New York: ACM Press , 2012:823-831.

[8] Lim Ee-Peng, Nguyen Viet-An, Jindal Nitin, et al. Detecting product review spammers using rating behaviors [C]// Proceedings of the 19th ACM international conference on Information and knowledge management. New York: ACM Press , 2010:939-948.

[9] Jindal Nitin, Liu Bing, Lim Ee-peng, et al. Finding unusual review patterns using unexpected rules[C] // Proceedings of the 19th ACM International Conference on Information and Knowledge Management. New York: ACM Press , 2010:1549-1552

[10] Wang Guan, Xie Sihong, Liu Bing,et al. Identify online store review spammers via social review graph[J]. ACM Transactions on Intelligent Systems and Technology, 2011,3(4):61.1-61.21.

[11] Nitin Indurkhya, Fred J Damerau. Handbook of natural language processing[M]. 2nd. Chapman and Hall/CRC, 2010.

[12] Ohana B, Tierney B. Sentiment classification of reviews using SentiWordNet[C]//9th. IT & T Conference. 2009: 13.

[13] Hu X, Tang J, Gao H, et al. Unsupervised sentiment analysis with emotional signals[C]//Proceedings of the 22nd international conference on World Wide Web. International World Wide Web Conferences Steering Committee, 2013: 607-618.

[14] Pak A, Paroubek P. Twitter as a Corpus for Sentiment Analysis and Opinion Mining[C]//LREC. 2010.

[15] Li N, Wu D D. Using text mining and sentiment analysis for online forums hotspot detection and forecast[J]. Decision Support Systems, 2010, 48(2): 354-368.

[16] Bermingham A, Smeaton A F. Classifying sentiment in microblogs: is brevity an advantage?[C]//Proceedings of the 19th ACM international conference on Information and knowledge management. ACM, 2010: 1833-1836.

[17] Mudinas A, Zhang D, Levene M. Combining lexicon and learning based approaches for concept-level sentiment analysis[C]//Proceedings of the First International Workshop on Issues of Sentiment Discovery and Opinion Mining. ACM, 2012: 5.

[18] Zhang Q, Wu Y, Li T, et al. Mining product reviews based on shallow dependency parsing[C]//Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval. ACM, 2009: 726-727.

[19] Wu Y, Zhang Q, Huang X, et al. Phrase dependency parsing for opinion mining[C]//Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 3-Volume 3. Association for Computational Linguistics, 2009: 1533-1541.

[20] Joshi M, Penstein-Rosé C. Generalizing dependency features for opinion mining[C]//Proceedings of the ACL-IJCNLP 2009 Conference Short Papers. Association for Computational Linguistics, 2009: 313-316.

[21] Zhao, Wayne Xin, et al. Jointly modeling aspects and opinions with a MaxEnt-LDA hybrid. Proceedings of the 2010 Conference on Empirical Methods in Natural

Language Processing. Association for Computational Linguistics, 2010.
[22] Thet, Tun Thura, Jin-Cheon Na, and Christopher SG Khoo. Aspect-based sentiment analysis of movie reviews on discussion boards. Journal of Information Science 36.6 (2010): 823-848.
[23] Mei, Qiaozhu, et al. Topic sentiment mixture: modeling facets and opinions in weblogs. Proceedings of the 16th international conference on World Wide Web. ACM, 2007.
[24] Myle Ott, Yejin Choi, Claire Cardie, et al. Hancock. Finding deceptive opinion spam by any stretch of the imagination [C]//Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1, 2011: 309-319
[25] Zhang, Rong, et al. Exploiting shopping and reviewing behavior to re-score online evaluations [C]//Proceedings of the 21st international conference companion on World Wide Web. ACM, 2012.
[26] Mukherjee, Arjun, Bing Liu, and Natalie Glance. Spotting fake reviewer groups in consumer reviews.[C]// Proceedings of the 21st international conference on World Wide Web. ACM, 2012.
[27] Esuli, Andrea, and Fabrizio Sebastiani. Sentiwordnet: A publicly available lexical resource for opinion mining. Proceedings of LREC. Vol. 6. 2006.
[28] Wilson, Theresa, et al. OpinionFinder: A system for subjectivity analysis. Proceedings of HLT/EMNLP on Interactive Demonstrations. Association for Computational Linguistics, 2005.
[29] Cohen J. A coefficient of agreement for nominal scales[J]. Educational and psychological measurement, 1960, 20(1): 37-46.
[30] Cohen J. Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit[J]. Psychological bulletin, 1968, 70(4): 213.

Qingxi Peng (male) was born in 1974. He is a Phd. candidate of State Key Laboratory of Software Engineering in Wuhan University. His research interest is opinion mining and sentiment analysis.

Ming Zhong (male) was born in 1982. He is an assistant professor of State Key Laboratory of Software Engineering in Wuhan University. His research interest is data management.