A Research of Taobao Cheater Detection

Baohua Dong, Qihua Liu, Yue Fu, Liyi Zhang

A Research of Taobao Cheater Detection

Baohua Dong[1], Qihua Liu[2], Yue Fu[1], Liyi Zhang[1]✉

[1]School of Information Management, Wuhan University, Wuhan 430072, P.R. China
baohua@whu.edu.cn fuyue412@gmail.com lyzhang@whu.edu.cn
[2]School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330013, P.R.China
qh_liu@163.com

Abstract. This paper focuses on Taobao cheater detection. At present the phenomenon of fake trading is widespread in Taobao, which makes it difficult for consumers to distinguish between true and fake product reviews. To solve this problem, we collect a total number of 50,285 historical review data from 100 cheaters and 100 real buyers to create a dataset. By using these data, we extract 8 features from three dimensions that are reviewer, commodity, and review. Then we use the SVM algorithm to construct the classification model and choose the RFB kernel function, which has a better performance to identify the cheater. The precision of the final classification model we built to identify the cheater reaches up to 89%. The experimental result shows that extracting features from the historical review data can recognize the cheaters effectively. It can be applied to the recognition of the cheaters in Taobao.

Keywords: Fake trading, Product review, SVM, Cheater detection

1 Introduction

According to the 33rd China Internet network development state statistic report issued by China Internet Network Information Center (CNNIC), by the end of December 2013, the number of online consumers in China has reached 302 million [1]. Such a huge online shopping market makes numerous entrepreneurs see business opportunities, thousands of new stores set up on Taobao every day. However, many of them have little online traffic due to low credibility. Therefore, many newly opened stores try to improve their reputation in various ways. The most typical one is fake trading. Fake trading refers to the cheating behavior that some merchants in e-commerce platform improve sales, store ratings and credit score by improper means. Fake trading has developed into a huge industry. We can find hundreds of third party fake trading platforms and thousands of QQ or YY groups that are serviced for fake trading.

Consumers usually read the reviews before buying. However, due to the existence of fake trading, it is difficult for them to judge the reality of reviews. Thus, reading reviews while shopping online is a double-edged sword, real reviews make them understand the products better, but when a spam review comes, they will be

misguided. Thus, finding spam reviews timely is of great significance. This paper adopts the method of identifying cheater to identify spam reviews. If a reviewer is cheater, the reviews he has published can be regarded as spam reviews.

2 Related work

Online reviews research has been a hot topic in recent years, especially spam reviews that attract attention of many scholars at home and abroad. Researches on spam reviews mainly concentrate on text classification and consider the recognition process as the classification of spam and true reviews through manual label, extracting features from the text and using machine learning methods to identify spam reviews automatically. However, due to the judgment of true or fake of product reviews is relevant to psychology, philosophy and many other fields, in addition, it involves the process of natural language understanding and opinion extraction of the review text, therefore, it is difficult to detect fake reviews based on review content, the effect is not very good either. Thus, scholars began to focus on the behavioral characteristics of reviewers based on which we can determine whether a review is a spam review. This method is often used to find spammers. It holds that, if a user is a spammer, then his reviews have strong possibility to be spam reviews. Jindal et al [2] take reviewer's behavior into consideration and analyze the possibility of the reviewer to be a spammer through finding abnormal review patterns. If a user repeatedly publishes positive reviews, then there is strong possibility that he is a spammer. Wu et al [3, 4] use the proportion of positive singletons (singleton refers to the only review that a reviewer had published) in all the reviews and the time aggregation degree of these singletons to analyze reviewer's suspicious behavior. Wang et al [5] discover the reinforcement relations of reviews' trustiness, reviews' honesty, and stores' reliability. They use such relations to discover suspicious spammers. Mukherjee et al [6] use frequent pattern mining to find groups of reviewers who frequently write reviews together, and then they construct features to find the most likely groups of spammers. They also construct a graph modeling the relations between groups of spammers, spammers and products for group spammer ranking [7]. Zhang [8] focuses on detecting the credibility of customers by analyzing online shopping and review behavior, and then they re-score the reviews for products and shops. Spammers can be taken as the special case of their work, which had very low credibility. Lim et al [9] propose a behavioral approach to detect review spammers who try to manipulate review ratings on some target products or product groups.

However, because these studies basically define the behavior of spam reviewers artificially, then determine the spammers or spam reviews. The precision of this method is hard to estimate. In addition, the former studies mainly focus on reviews of one or several products; few studies focus on the historical review data of the reviewers that exist on a platform such as Taobao. For this reason, we collect cheaters' account information from the third party fake trading platforms and real buyers' account information from popular stores on Tmall. We get the historical review data of these buyers on Taobao using crawler software, and then extract

features from three aspects including reviewer, commodity, and review to identify the cheaters.

3 Feature extraction

There are many forms of fake trading; currently the most used one is fake trading on third-party platform, which links both side (sellers who need fake trading and cheaters who engage in fake trading) through an intermediate platform. Both sides publish relevant information on the middle platform, and platform provides guarantee as a third party for both sides to reach an agreement. There are several typical fake trading platforms, such as hiwinwin.com, shuaxinyong.com, shuakewang.com, etc.

Many scholars have adopted the machine learning method, establishing spam review feature library for identification of spam reviews. Among them, most of the scholars extracted corresponding features according to the content of review, such as review sentiment polarity, scoring etc. However, since the review content belongs to natural language, which is difficult to process, and there is no significant difference between true reviews and spam reviews. Therefore, it is almost impossible to identify spam reviews from the comment content. For these reasons, many scholars begin to look for features from other sources. Mukherjee et al [10, 11] believe that the factors such as content similarity, maximum number of reviews, ratio of first reviews and review length have a significant effect on identifying spam reviews. In another paper, Mukherjee et al [12] report that opinion spammers are usually not longtime members of a site. Real reviewers, however, use their accounts from time to time to post reviews. These features have played a certain effect on identifying artificially labeled spam reviews, but whether they can be used to identify the real spam reviews remains to be verified. More importantly, the previous research object is the data of Amazon, but there are many acts of fake trading and spam reviews on Taobao. In order to identify cheaters on Taobao, this paper absorbs the previous research results, analyzing the information of cheaters on third party spam trading platform and information of normal reviewer from Tmall, and we summarized the specific characteristics of cheaters on Taobao. Review centric features include gender (F1), number of registration days (F2) and identity authentication status (F3). Commodity centric features include total number of commodity categories (F4), number of categories purchased in a single day (F5). Review centric features include average length of reviews (F6), daily number of reviews (F7), and no repetition rate of reviews (F8).

Using the above features of reviewers to constitute an eigenvector for every reviewer, calculation methods of the eigenvector are as follows:

$F1 = n,\ (n = 0,\ 1)(0: female; 1: male)$

$F2 = n,\ (n = 0,\ 1, 2\ldots\ldots)\ (n: number\ of\ registration\ days)$

$F3 = n,\ (n = 0,\ 1)\ (0: no\ authentication;\ 1: authentication)$

$F4 = n,\ (n = 1, 2\ldots\ldots,\ 16)(n: total\ number\ of\ commodity\ categories\ which\ are\ divided\ into\ 16\ classes\ according\ to\ the\ Taobao\ classification)$

$$F5 = \frac{sum\ \text{of}\ daily\ number\ of\ commodity\ categories}{total\ number\ of\ days\ that\ purchase\ behavior\ happened}$$

$$F6 = \frac{number\ of\ words\ in\ all\ reviews}{total\ number\ of\ reviews}\ \#$$

$$F7 = \frac{sum\ of\ daily\ number\ of\ reviews}{total\ number\ of\ days\ that\ purchase\ behavior\ happened}\ \#$$

$$F8 = \frac{number\ of\ no\ repetition\ reviews}{total\ number\ of\ reviews}$$

4 Methodology

4.1 SVM Model

In our study, we focus on distinguishing cheaters from real buyers. In order to solve this problem, we treat the task as a binary classification problem. Given a training data set $D = \{xi, yi\}_1^n$, we can build a model that can minimize the error in prediction of y given x (generalization error) [13]. Here $x_i \in X$ and $y_i = \{cheater, real\ buyer\}$ represents a buyer and a label, respectively. The model predicts the corresponding y and outputs the score of the prediction when it is applied to a new instance. We use SVM (Support Vector Machines) [14] as the model of classification since it is very effective to solve the binary classification problem. If an instance x (reviewer) is given，SVM assigns a score to it according to

$$f(x) = w^T x + b \qquad (1)$$

where w denotes a vector of weights and b denotes an intercept. The value of $f(x)$ presents the quality of the instance x, the higher value of $f(x)$ is, the higher quality of the instance x is. The sign of $f(x)$ is employed in our classification. If the sign is positive, then x is classified into the positive category (real buyer), otherwise into the negative category (cheater). The building of SVM needs labeled training data (in our case, the categories are “real buyer” and “cheater”). Briefly, the learning algorithm creates the “hyper plane” in (1), and the hyper plane separates the positive and negative instances in the training data with the largest margin.

4.2 Experimental program

This study selects historical review data of cheaters and real buyers as object, using LIBSVM tool for training and testing experimental data. Specific programs are as follows: Firstly, data processing and format adjusting are done on historical review data of reviewers. We use the crawler software crawling the basic personal information and review information of reviewers on Taobao. These data are consolidated, and we build a data set of historical reviews for each reviewer. Elements

and the format of all data sets are the same. Secondly, we assign features according to reviewer, commodity, and review to construct a standard data set. We extract features from historical review of each reviewer and construct a data set using feature data of all reviewers. All data are labelled completed. Thirdly, we call LIBSVM tool and use radial basis kernel function to train samples and generate a model file. We construct a classification model using the data of 70 cheaters and 70 normal reviewers for testing. Lastly, we test and forecast the generated model using the test data set. We use the classification model to test the remaining 30 cheaters and 30 normal reviewers, in order to testing the effectiveness of the classification model.

4.3 Evaluation indexes

The most commonly used classification indexes for text classification are recall, precision and F1-measure [15]. The recall of a class X is the ratio of the number of users correctly classified to the number of users in class X. The precision of a class X is the ratio of the number of users classified correctly to the total predicted as users of class X. The F1-measure is the harmonic mean between both precision and recall, and it is usually reported to evaluate classification effectiveness. Our research can also be seen as a classification problem. To assess the effectiveness of our classification strategies, we use the standard classification index of recall, precision and F1-measure.

（1）Recall

Recall refers to the ratio of the cheaters correctly classified to the number of cheaters in our dataset. The calculation formula is as follow:

$$Recall = \frac{|TP|}{|TP| + |FN|}$$

（2）Precision

Precision refers to the probability that cheaters are correctly predicted to be cheaters. The calculation formula is as follow:

$$Precision = \frac{|TP|}{|TP| + |FP|}$$

（3）F1-measure

F1-measure refers to the harmonic mean between both precision and recall. The calculation formula is as follow:

$$F1 = \frac{2 * Recall * Precision}{(Recall + Precision)}$$

|TP| refers to the number that cheaters are correctly predicted to be cheaters, |FN| refers to the number that cheaters are incorrectly predicted to be real buyers, |FP| refers to the number that real buyers are incorrectly predicted to be cheaters.

5 Experiments and result analysis

5.1 Introduction to Dataset

In this study, we collect nearly 50285 data about historical review information of cheaters and real buyers.

The data collection method of cheaters is as follows: Firstly, we get merchants' fake order information and specified requirement from the professional third-party fake trading platforms, such as shuaxinyong.com, hiwinwin.com, shuakewang.com and so on. Secondly, we find out the cheaters' account information by comparing the specified requirements with the order information. Thirdly, we collect the cheaters' personal information by the Taobao inquiry website taodake.com. Lastly, we collect each cheater's data by using the web crawler (The system default reviews and anonymous reviews are not included in the data).

The data collection method of real buyers is as follows: Firstly, we choose the stores with great influence and high credit which don't need to improve credit ratings through fake trading, such as the official flagship store of Xiaomi. Secondly, we find out the real buyers' account information by judging the behavior of the anonymous reviewers. If the reviewer does not have abnormal behavior, that is, the reviews are objective, it can be concluded that the reviewer is a real buyer. Thirdly, we collect the real buyers' personal information through the Taobao inquiry website taodake.com. Lastly, we collect each real buyer's data by using the web crawler (The system default reviews and anonymous reviews are not included in the data).

5.2 Experiments

Our experiments take matlab2010b (matlab7.1) as the experimental platform, the support vector machine is professor Lin's libsvm-mat-3.1 version. We adjust the format of data first, and then import it into the matlab and conduct experiments. The training set includes 140 samples (70 cheaters and 70 real buyers), the test set includes 60 samples (30 cheaters and 30 real buyers), and all the samples have been labeled. The experiment uses a RBF SVM model to identify cheaters out of the data set. When the data is normalized, we use the grid search and ten-fold cross validation to optimize the SVM parameters. Then we use the optimized parameters to construct a classification model that can be used to predict the test set. The predict result is shown in Figure.1.

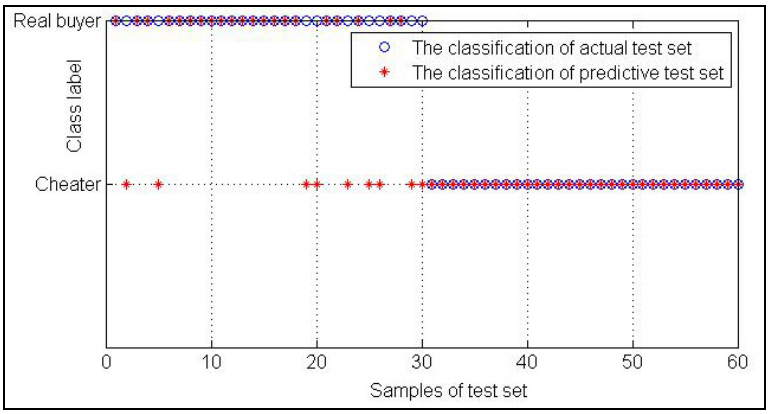

Figure 1. The actual classification and predictive classification of the test set

We can see that the model established using the data of 70 cheaters and 70 normal reviewers for training have better ability to identify cheaters, 30 cheaters were successfully predicted. Considering our ultimate goal is to kick out the spam reviewers to reduce the harm caused to consumers, although some normal reviewers are mistakenly identified as cheaters, the error is still acceptable. Overall, the classification model we developed has good prediction ability, and it can well separate cheaters and normal reviewers.

5.3 Result analysis

The experimental result of all samples is as follow:

Table 1. The final results of all samples classified by the classification model

Cheaters					
The number of cheaters correctly identified	The number of cheaters identified by the model	The actual number of cheaters	recall	precision	F1-measure
93	104	100	93%	89%	91%

As we can see, the better cheater recognition effect is due to the choice of 8 characters. The characteristic dimension is relatively high and the RBF kernel function effectively maps it into a high dimensional feature space, so the cheaters and real buyers are certainly linearly separable. What's more, the effect is better.

6 Conclusion

This paper identifies the spam reviews by identifying cheaters. We choose the history review data of 100 cheaters and 100 real buyers as our research object. From these

data we extract 8 features using SVM algorithm to construct a classification model, and the precision of the final classification model we built to identity the cheater reaches up to 89%. The experimental result shows that, extracting features from the history review data can identify the cheaters effectively. It can be applied to the recognition of the cheaters in Taobao.

Acknowledgement. The work is supported by the Natural Science Foundation of China (No. 71373192，71363022) and the MOE Project of key research Institute of Humanities & Social Science in Chinese Universities (No: 14JJD870002)

References

1. China Internet Network Information Center. The 33rd Statistical Report on Internet Development in China, 38-39 (2014).
2. Jindal, N., Liu, B., Lim, E.-P.: Finding unusual review patterns using unexpected rules. In: CIKM 2010, pp. 1549–1552. ACM, New York (2010).
3. Wu, G., Greene, D., Smyth, B., Cunningham, P.: Distortion as a validation criterion in the identification of suspicious reviews. In: Proceedings of the First Workshop on Social Media Analytics, SOMA 2010, pp. 10–13. ACM, New York (2010).
4. Xie, S., Wang, G., Lin, S., Yu, P.S.: Review spam detection via temporal pattern discovery. In: KDD, pp. 823–831 (2012).
5. Wang, G., Xie, S., Liu, B., Yu, P.S.: Identify Online Store Review Spammers via Social Review Graph. Journal of ACM Transactions on Intelligent Systems and Technology (TIST) 3(4) (September 2012).
6. Mukherjee, A., B. Liu, J. Wang, N. Glance, and N. Jindal.: Detecting Group Review Spam. In: Proceedings of the 20st International Conference on World Wide Web. ACM (2011).
7. Mukherjee, A., Liu, B., Glance, N.: Spotting fake reviewer groups in consumer reviews. In: Proceedings of the 21st International Conference on World Wide Web. ACM (2012).
8. Zhang R.: Exploiting shopping and reviewing behavior to re-score online evaluations. In: Proceedings of the 21st international conference companion on World Wide Web. (2012).
9. Lim, E.-P., Nguyen, V.-A., Jindal, N., Liu, B., Lauw, H.: Detecting Product Review Spammers using Rating Behaviors. In: Proceedings of the 19th ACM International Conference on Information and Knowledge Management (2010).
10. Mukherjee A, Kumar A, Liu B, et al.: Spotting opinion spammers using behavioral footprints. In: Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM (2013).
11. Mukherjee A, Liu B, Glance N.: Spotting fake reviewer groups in consumer reviews. In: Proceedings of the 21st international conference on World Wide Web. ACM(2012).
12. Mukherjee A, Venkataraman V, Liu B, et al.: What Yelp Fake Review Filter Might Be Doing? In: ICWSM (2013).
13. Liu, J., Cao, Y., Lin, C.Y., Huang, Y., Zhou, M.: Low-Quality Product Review Detection in Opinion Summarization. In: EMNLP-CoNLL, pp. 334–342 (2007).
14. Vapnik, V. N.: An overview of statistical learning theory, IEEE Transactions on Neural Networks 10(5), 988–999 (1999).
    Yang, Y.: An evaluation of statistical approaches to text categorization. Information Retrieval Journal (1999).