A Method for the Detection of Fake Reviews based on Temporal Features of Reviews and Comments

Wenqian Liu [a], Jingsha He, Song Han, Nafei Zhu

Faculty of Information Technology Beijing University of Technology, Beijing 100124, China

[a]18515290389@163.com

Abstract. Online reviews and comments have become an important resource for various decision making processes, such as sale and buy decisions. The truthfulness of online reviews is thus critical for both buyers and sellers since fake reviews will affect customer's decisions due to misleading description and deceptive selling. This can cause financial loses for innocent customers. Fake review detection has thus attracted a lot of attention. However, most shopping websites have only focused on dealing with problematic reviews and comments. In this paper, we propose a method for the detection of outlier reviews based on reviewing records associated with products instead of just the reviews and comments. We first analyze the characteristics of such data using a crawled Amazon China dataset, revealing that the reviewing records of each product is similar for normal products. In the proposed method, we first extract the reviewing records of products to a temporal feature vector. We then develop an isolation forest algorithm to detect the outlier reviews of products based on the reviewing records of reviews and comments. We will verify the effectiveness of our proposed method and compare it to some existing temporal outlier detection methods using the crawled Amazon China dataset. We will also study the impact caused by the parameter selection of the reviewing records.

Keywords: fake reviews; products speculation; reviewing records; Isolation Forest algorithm.

1. Introduction

In recent years, many researchers have developed methods to detect fake reviews using text mining techniques. Most such work has focused on analyzing one review or one reviewer at a time without considering the potential relationship between multiple reviews or reviewers [1,2]. Han et al. analyzed burst reviews to find the outlier behavior of both reviewers and reviews [3]. Furthermore, the behavior of fake reviewers are analyzed to develop possible reviewing patterns so as to detect fake reviews [4,5].

In Section 2, we briefly review some related work in feature learning for networks. In Section 3, we empirically analyze the reviewing record and the outlier behavior. In Section 4, we describe the isolation forest based product fake review detection method and in Section 5, we evaluate the proposed method using a real dataset and compare it to several baseline methods to demonstrate the advantages of our method. Finally, in Section 6, we conclude this paper in which we also point out some promising directions for future research.

2. Related Work

2.1 Spamming Detection

In recent years, the web spam or email spam have been actively studied. For example, a survey is provided on web spam detection [6]. Email spam detection is also studied [7]. Blog spam or network spam are also intensively studied [8,9]. For the review spams, Fei et al. studied the behavior of fake reviews and provided the possible spam patterns [5].

2.1.1 Time Series Outlier Detection

Time series analysis is one of the most actively pursued approaches in outlier detection. These methods rely on the definition of a similarity function that measures the similarity between two sequences and outlier is detected using clustering. By treating all data samples as a time series feature vector, these samples can be clustered and the data sample that is the furthest to all the clusters gets the largest outlier score [10].

Parametric models can also be used to detect outliers in the unsupervised manner where anomalous instances are not specified and a summary model is constructed on the base data.. Furthermore, HMMs are interpretable but cannot scale well to pattern complicated data. Approaches that use HMMs have been proposed for outlier detection [11-13].

2.1.2 Outlier Detection for Stream Data

There is another category of methods that perform outlier detection for streaming data where the scenario becomes more complex than regular outlier detection.

When dealing with stream data, evolving prediction models are needed that will update the parameters or model components when new data arrives. For example, an online clustering method was proposed to detect outlier products [14,15]. At the same time, an approach was proposed to use dynamic Bayesian networks to model data samples that evolve over time [16]. By adding new state variables, the state of a system can be obtained.

In this paper, we propose an outlier detection method based on ideas from both stream outlier detection and time series outlier detection.

3. Trend Analysis of Product Reviews

In this section, we analyze the shopping review data crawled from Amazon. From the analysis, we can see clearly the differences in the reviews and comments of different products.

3.1 General Trend for Product Review

In this study, we use the Amazon-China dataset. the number of recorded reviews is growing. In 2006, only a few reviews were recorded. With the time goes by, more and more reviews and comments were recorded.

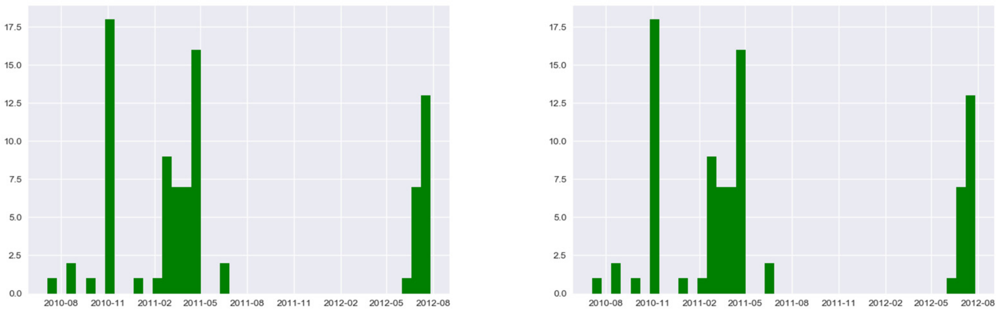
The parameters of the review data is listed in Table 1 in which the dataset contains 166,624 products and 5,055 users and the review period spans between March 2006 and August 2012. In total, there are 1,205,125 reviews.

Table 1. Fake review dataset

Information	Value
Products	166,624
User	5,055
Time period	Mar, 2006 ~Aug, 2012
Number of reviews	1,205,125
Frequency	507.2 reviews per day

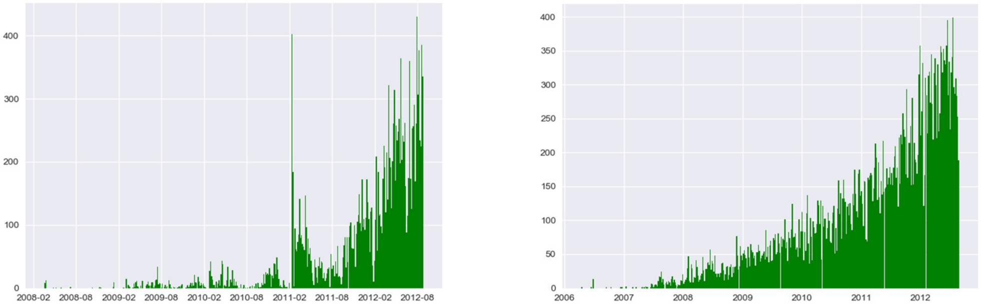
3.2 The Trend in Product Reviews

We select two products and analyze the temporal pattern of reviews, which is shown in Figure 1. For the product in Figure 1(a), we can see more clearly that the pattern of trend is not continuous. In Figure 1(c), we summarize the number of suspicious reviews while in Figure1(d), we plot the number of normal reviews.



(a) Suspect reviews

(b) Normal reviews



(c) Suspect reviews (d) Normal reviews

Figure 1. Pattern of reviews of two typical selected products

The basic data unit in our method is the product. We focus on studying the review patterns of all products to obtain outlier products. Since our method is an unsupervised review detection method, fake reviews more likely happen along with other fake reviews and the number of products in one specific time slot could be very large. According to this characteristics, we only collect the reviews within one day as our unit description of one product. For example, for product P_i, the description of the review pattern is defined as:

$$V_{p_i} = [p_{i_1}, p_{i_2}, p_{i_3}, \cdots, p_{i_E}] \tag{1}$$

where i indicates the date and thus i_1 indicates the first date with reviews in the dataset (March 2006) and i_E indicates the last date with reviews in the dataset (August 2012), and V_{P_i} is the combination feature vector that represents the review records of product P_i . Each element in the vector is a date-review indicator and the date-review indicator p_{ij} represents the number of reviews given in a certain date i_1.

3.3 Temporal Feature Extraction based on Reviews and Comments

We view the product reviews statically, and the review records of each product can be processed as an *N*-dimensional vector. The general form of the temporal feature can be described as:

$$Z_{p_i} = \{z(1)_{P_i}, z(2)_{P_i}, \cdots, z(t)_{P_i}, z(t+1)_{P_i}, \cdots, z(N)_{P_i}\} \tag{2}$$

where $z(t)_{P_i} \in R^N \ (t \geq 1)$ indicates the number of reviews in time slot t for product Pi and N is the total number of time slots to be processed.

Since the data for the reviews ranges from 2006 to 2012, if we define a time slot as one year, there will be seven time slots in total. It is also important to define all the products with the same dimension. For instance, if the time slot is M days and there are N time slots in total, then *z(t)* will be:

$$z(t)_{P_i} = \sum_{m=1}^{M} p_{i_{t*M+m'}} \tag{3}$$

where t is the t^{th} time slot of the feature and subscript t*M+m' indicates the date time of the specific review.

We can then describe all products using a matrix:

$$Z = [Z_{P_1}, Z_{P_2}, \cdots, Z_{P_P}]_{P \times N} \tag{4}$$

where P is the number of products in total. Isolation forest algorithm can be used to process the data.

3.4 Isolation Forest Algorithm for Outlier Detection

For the temporal feature vector Z, the initial outlier detection model is developed in which we build the isolation trees in terms of the bootstrap sampling from dataset Z. The ensemble detection model E is composed of L number of iTrees, namely,

$$E = \{E_1, E_2, E_3, \cdots, E_L\} \tag{5}$$

which is built from the data in the *i*th time slot.

In the algorithm, an iForest consists of multiple isolation trees, namely iTree. We know that iTree is created by selecting product temporal review features and the feature values randomly [17]. At each node in the isolation trees, the instances set is divided into two parts based on chosen the temporal review value. Generally, products with outlier reviews are those that have review records or review values very different from the normal products and are easier to be divided than normal products. In order to alleviate the effects imported by the random characteristics in the process of building isolation forest, the average depth of products in the forest is calculated, which can serve as the anomalous score of the products. The lower the score, the further away the product is to the normal products, making it a likely candidate to be an outlier product. Figure 7 further illustrates the algorithm.

In summary, we use isolation forest algorithm to build the isolation forest based on the product review records. Meanwhile, the outlier score can be obtained by applying the isolation forest algorithm.

The anomaly score is used to determine whether a product is an outlier product. For product P_i , the anomaly score can be calculated using formula (6).

$$S(z_{P_i}, N) = 2^{-\frac{E(h(z_{P_i}))}{c(N)}} \tag{6}$$

where

$$E(h(x)) = \frac{1}{L}\sum_{i=1}^{L} h_i(x) \tag{7}$$

In Formula (6), N denotes the sampling size in Algorithm 1, $h_i(x)$ indicates the length of the *i*th iTree, $E(h(x))$ is the average of $h(x)$ from a collection of iTrees and $c(N)$ is the average of $h(x)$ with a given N. For product P_i, the outlier score is $S(z_{P_i}, N)$. An anomaly score of high value is regarded as an outlier while low value is regarded as a normal sample. A high anomaly score actually indicates product review pattern that is different from that of normal products temporally. Thus, a high anomaly score is considered to be highly probable that the corresponding product consists of fake reviews and comments. We can therefore use Algorithm together with Formula (6) to analyze the possibility of abnormal reviews of products based on outlier scores.

4. Experiment and Analysis

We have conducted some experiment using the dataset described in Section 3. Through the experiment, we first compare our proposed method to several baseline fake reviews speculation methods to demonstrate the effectiveness of our method. We also study the performance of our method with different temporal parameter settings.

4.1 Measurement Metrics

To evaluate our method, we quantify the performance in terms of the ground-truth outlier labels and the predicted outlier labels. We use two metrics to measure the performance of our method.

4.2 Comparative Analysis and Evaluation of Experiment Results

In this section, we use three baseline methods, i.e., ARIMA, LOF and SVM, to detect outlier products in the dataset described in Section 3. Thus, we followed the same re-crawl strategy to detect whether each product is regarded as abnormal. For each product, a low anomaly score is regarded as abnormal, making the product possessing the outlier commercial behavior. The results of accuracy for all the methods are shown in Figure 2, showing that our method can better detect fake reviews. Lower accuracy implies that the detection method is not promising while higher accuracy makes it more successful for the detection of fake reviews.

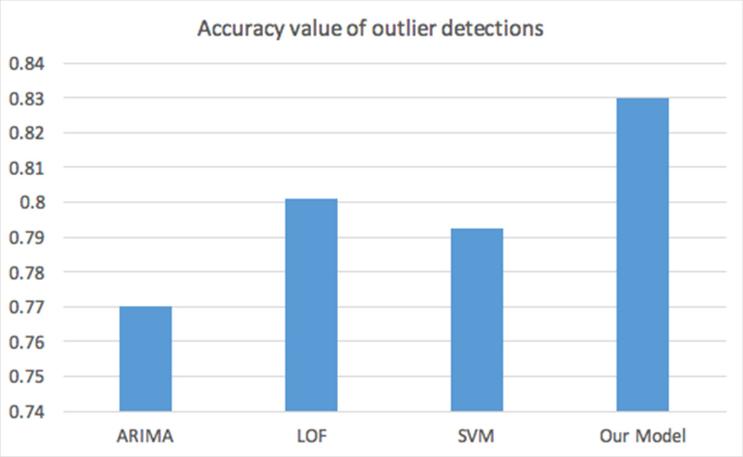


Figure 2. Comparison of accuracy

As can see in the Figure 2, our method performs better than the other three baseline methods. ARIMA can only reach 0.77 in accuracy, which may be caused by insignificant performance change in time-series. LOF is the most competitive method among the three baseline ones, which indicates that outlier can also happen locally. Isolation forest-based methods show to be superior to all the other methods.

We also compared our detection method to the three baseline methods in terms of efficiency ass shown in Figure 3. The result shows that the isolation forest method can significantly reduce the amount of running time and the tree based approach can be fast in both the training and the evaluation phases.

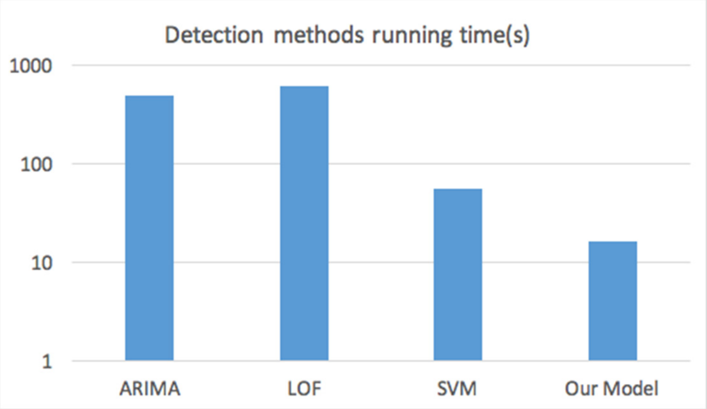


Figure 3. Comparison of efficiency

5. Conclusion

In this paper, we studied the review records of online shopping sites and proposed a novel approach to detecting fake reviews of products. This review outlier detection method detects the outlier products by temporal trends of reviews and comments. Such perspective makes our method more advantageous than some existing methods. We also compared our method with several temporal outlier detection methods to prove the effectiveness and the efficiency of our method.There are a lot of challenges in the detection of fake reviews based on review records. Our experiment did not indicate clearly when a product has the highest probability of being involved in fake reviews and comments, which is another interesting piece of future work.

Acknowledgements

The work in this paper has been supported by National Natural Science Foundation of China (61602456).

References

[1]. Streitfeld, D. Fake reviews, real problem. New York Times. http:// query. nytimes. com/ gst/ fullpage.html, 2012.

[2]. Rayana S.; Akoglu L. Collective opinion spam detection: bridging review networks and metadata. Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2015, pp. 985-994.

[3]. Catal C.; Guldan S. Product review management software based on multiple classifiers. IET Software, 2017, 11(3): 89-92.

[4]. Han S.; Prince J.; Zuo L.; Carass A. Automatic outlier detection using hidden Markov model for cerebellar lobule segmentation. Proceedings of International Conference on medical Applications in Molecular, Structural, and Functional Imaging, 2018.

[5]. Mukherjee, A.; Venkataraman, V.; Liu, B.; Glance, N.S. What yelp fake review filter might be doing? Proceedings of the 7th International AAAI Conference on Weblogs and Social Media, 2013, pp. 409-418.

[6]. Spirin, N.; Han, J. Survey on web spam detection: principles and algorithms. ACM SIGKDD Explorations Newsletter 2012, 13: 50-64.

[7]. Chirita, P.A.; Diederich, J.; Nejdl, W. MailRank: using ranking for spam detection. Proceedings of the 14th ACM International Conference on Information and Knowledge Management, 2005, pp. 373-380.

[8]. Yang, W.; Kwok, L. Improving blog spam filters via machine learning. 2017, 9: 99-121.

[9]. Tan, E.; Guo, L.; Chen, S.; Zhang, X.; Zhao, Y. Unik: Unsupervised social network spam detection. Proceedings of the 22nd ACM international conference on Information & Knowledge Management, 2013, pp. 479-488.

[10]. Takahashi, T.; Hooi, B.; Faloutsos, C. Autocyclone: automatic mining of cyclic online activities with robust tensor factorization. Proceedings of the 26th International Conference on World Wide Web, 2017, pp. 213-221.

[11]. Fei, G.; Mukherjee, A.; Liu, B.; Hsu, M.; Castellanos, M.; Ghosh, R. Exploiting burstiness in reviews for review spammer detection. Proceedings of the 7th International AAAI Conference on Weblogs and Social Media, 2013, pp. 175-184.

[12]. Ovhal, K.B.; Patange, S.S.; Shinde, R.S.; Tarange, V.K.; Kotkar, V.A. Analysis of anomaly detection techniques in video surveillance. Proceedings of the 2017 International Conference on Intelligent Sustainable Systems, 2017, pp. 596-601.

[13]. He, X.; Dai, H.; Ning, P. HMM-based malicious user detection for robust collaborative spectrum sensing. IEEE Journal on Selected Areas in Communications, 2013, 31: 2196-2208.

[14]. Yamanishi, K.; Takeuchi, J. A unifying framework for detecting outliers and change-points from nonstationary time series data. Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2002, pp. 676-681.

[15]. Aggarwal, C.C.; Yu, P.S. On clustering massive text and categorical data streams. Knowledge and Information Systems, 2009, 24: 171-196.

[16]. Smith, D.V.; Timms, G.P.; de Souza, P.A.; D'Este, C. A Bayesian framework for the automated online assessment of sensor data quality. Sensors, 2012, 12(1): 9476-9501.

[17]. Liu, F.T.; Kai, M.T.; Zhou, Z.H. Isolation forest. Proceedings of the 8th IEEE International Conference on Data Mining, 2009, pp. 413-422.