Identification model of commodity false reviews based on integrated features

Li Jing
(Guangxi University of Finance and Economics, Guangxi, Nanning, 530003)
709740726@qq.com

Abstract—**More and more consumers like online shopping, and false reviews of goods mislead consumers to some extent, so it is necessary to design a false review detection method. Based on the integration characteristics of consumer history reviews, this paper proposes an online false comment detection model. The model introduces time series and combines the static and dynamic features of the reviews to detect false comments. Finally, the model is used to test the comment data of three kinds of Amazon products. The results show that the model has high recognition accuracy.**

Keywords-fake comment; feature extraction; integration; classifier; SVM

I. INTRODUCTION

With the development of Internet, especially in e-commerce, more and more consumers prefer online shopping and consumers comment the purchased product more easily. These product comment information provides valuable information resources for producers and potential consumers. Due to certain kind of interest relationship, there may be some unreal and fake contents. These fake comments affect reference value of comment information to some extent so as to mislead consumers so it is essential to detect fake comments. The most fundamental comment information is the commenting information. It is significant to mine comment information and utilize comment content information to detect fake comments.

This paper will transform recognition of fake comment target commodity into the behavior recognition of commodity abnormal scores. First, there are four types of commodity fake scoring status. Due to interruption which is caused by previous 3 types of fake scores on overall score distribution of commodity, this paper puts forward normal distribution fitting-based target commodity recognition. For the fourth kind of fake comment status, since its fake comment distribution appears stage feature, this paper puts forward fusing feature analysis-based recognition method. Finally, through simulating experiment, we prove the effectiveness of this method.

II. SCORING BEHAVIOR ANALYSIS OF FAKE COMMENTS

There has not been studies in target commodity recognition of fake comment till now. Then this paper applies scoring system of comment website to realize target commodity recognition of fake comment. Scoring system in comment website refers that current e-commerce platform offers a scoring mechanism and scores commodity from 1 to 5 during content comment. This score stands for consumers' commodity emotional attitude. 4 and 5 refer that consumers are very satisfactory to commodity. 1 and 2 refer that consumers are very unsatisfactory to commodity. 3 refers consumers' basic approval. There are two types of fake comments. One of them is active fake comment which contains boasting content. Its purpose is to promote sale and this comment will usually be scored 5. The other is negative fake comment and this comment is the competitor depreciates and suppresses fake comment on purpose. This kind of comment will usually scored 1. Thus, there are lots of 5 or 1 scoring behaviors and this is one important reference for commodity in fake comment.

Fake comment should play an important role in inducing consumers so it needs certain amount. Little amount of fake comment is difficult to affect consumers' decision-making. In order that commodity obtains more fake comments, professional writing team can be employed or good-comment for returning cash can be performed for consumers. This will result in several fake scoring status. The first fake scoring is good-comment for returning cash activity will induce consumers to score high marks. This process will proceed in the whole process of commodity sale, that is to say, fake comment will always mix with genuine comment. The other fake scoring status is to hire unrelated people to offer high scores. As far as some new products from sale to the end, they are mixed with unrelated people's high scores and genuine scores. Similar to the first status, the second fake scoring status will not last long because unrelated people's scoring is a short-term behavior. There are some commodities in normal comment stage while there are some other commodities in unrelated people's employment and high scoring stage. However, high scoring is very obvious to interrupt normal scoring commodity to form the third kind of fake scoring status. Finally, some commodities have fake scores by employing network workers periodically but there is repairing function in ghostwriters' high scoring. Thus, the high scores of ghostwriters will not largely affect overall scoring interruption of commodity to form the fourth fake score status.

Through above analysis, fake comment in previous fake scoring status is obviously affecting general scoring interruption. The final one is fake scoring status which little affects overall scores. At present, many scholars have proposed many effective methods to detect fake comments with the perfect effect. However, the common disadvantage among these studies is that their applied features are most in static feature. Static feature refers to the extracted feature in non-sequence data in comparison with dynamic feature in this chapter. Dynamic feature is the extracted feature from sequence data in time series data. Actually, users' historical behavior contains abundant dynamic information. If these information can be effectively applied to extract dynamic feature which can describe users' behavior law as well as characteristics and to fuse current static feature which is in broad application and relative maturity, accuracy in fake

978-1-5386-8031-5/18/$31.00 ©2018 IEEE
DOI 10.1109/ICVRIS.2018.00103

comment detection will necessarily be improved. Therefore, such idea should be applied to solve fake comment detection and this paper puts forward a method which applies semi-supervision learning strategy to establish classification model with dynamic feature. Basic process and structure is shown as figure 1

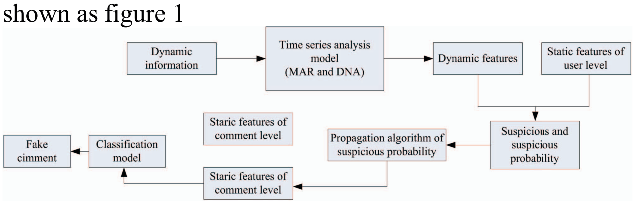

Figure 1. Structure of fake comment detecting method based on integrated features

III. Integration Feature-based Fake Comment Detection

A. Dynamic Feature Extraction

Started from commentators' perspective, we focus on studying its historical behavior data and obtain time series information from it. Information in this paper contains:

- all comments of one user will be ranked according to time sequence and successively labeled from earlier to later phase. Labels will be taken as independent variables and commenting scores are taken as dependent variables so it will form a feasible time series.
- Towards all comments of the same user to the same one product, the processing will be the same to above content and this will obtain scoring time series towards the same one product.
- All comments of one user's publishing time is taken as dependent variable and this users time series in active time can be achieved.
- Users' host IP is taken as dependent variable and user's time series in position can be obtained.

In terms of normally commenting product, there is no relationship between commenting quantity inside time window and average scores of comments.

However, for ghostwriters' fake commenting products in stages, commenting quantity and scoring means appear positive correlation or negative correlation inside unrelated people's operating time windows. That is, sudden increase of commenting quantity suddenly increases or decreases with scoring means. Therefore, time series data is constructed on commenting quantity and commenting means according to time window and fake comment target product is discovered through correlation suddenness of double time series data. This method is dynamic recognition method. According to above description, this method will divide analyzing time length T according to time window size and scoring quantity as well as scoring means will all be obtained inside each window. Finally, time series of scoring quantity and scoring means will be obtained. All time series will be analyzed. If it can determine related commenting quantity and scoring means sudden point, it indicates the product contains fake comments, that is, the target commodity is determined. The procedures in detail for detection of commodity are:

(1) Time series establishment. Given all the remarks of give commodity $C(p)=(r_1,r_2,...,r_{n_p})$ and comment hour $T=(tp_1,tp_2,...,tp_{n_p})$. n_p denotes total number of comment. The analysis time length is T and the size of time window is Δt . Then T can be divided to consecutive time window $T_n=[t_0+(n-1)t,t_0+nt]$. t_0 denotes the initial hour; for each time window we can acquire the following statistics:

$$f_1(T_n)=|\{r_i : tp_i \in T_n\}| \tag{1}$$

$$f_2(T_n)=\sum_{tp_i\in T_n}\frac{r_i}{f_1(T_n)} \tag{2}$$

$f_1(T_n)$ denotes the number of comment in T_n and $f_2(T_n)$ denote the average value of T_n .

(2) Abnormal point detection of time series. On above step, two-dimension time series $(f_1(T_n)f_2(T_n))$ is obtained through data statistics. The recognition of fake comment target commodity is transformed into abnormal point detection in two dimension time series. The points on time series are respectively fitted to broken line and abnormal points on broken line will be discovered. If related abnormal points are discovered, this indicates there are fake comments in this commodity, that is, the checked target commodity.

B. Fake Comment Detection Combined Dynamic Feature with Static Feature

In terms of analysis and extraction between comment content and evaluators' behavior feature before combination, we discover that suitable comment feature is the key to detect fake comment. Comment content feature and commentators' behavior feature can all effectively recognize fake comment. Meanwhile, based on previous studies, this paper integrates these two kinds of features to detect fake comment. Then, two features are taken as two mutually independent views, apply semi-supervised learning algorithm and depend on lots of unmarked samples to improve classifier performance so as to recognize fake comment.

In practical application situation, it is extremely difficult to construct a large training dataset. Furthermore, it is an incomplete task to judge one comment authenticity. Therefore, recall rate in traditional supervision learning algorithm is commonly low but AC-Learning can effectively solve this problem. AC-training algorithm requires that training dataset satisfies two abundant redundancy view conditions. First, two dataset can describe this problem. That is to say, if training samples are enough, classifier can respectively learn strong classifier from two dataset. Secondly, during given marks, two dataset are mutually and conditionally independent. It starts from a small marking evaluation dataset and constantly selects comment dataset from unmarked comment dataset to add to the marked comment dataset according to marking comment dataset. According to the extracted two features, the comment content feature and commentators' behavior feature respectively train a weak classifier on each feature. Classifier selection adopts SVM classifier since the detecting effect of SVM classifier is better than that of the largest entropy classifier and Bayesian classifier. Then, unmarked comment dataset are classified and the marked comment data add to marking comment dataset of another classifier. This will iterate and two classifiers coordinately train. Co-training algorithm is described as follows:

Algorithm AC-training(L,U)

1 repeat

2 Learn a classifier f_1 using L based on only x_1 portion of the examples x.

3 Learn a classifier f_2 using L based on only x_2 portion of the examples x.

4 Apply f_1 to classify the examples in U, for each class c_i, pick n_i examples that f_1 most confidently classifies as class c_i, and add them to L.

5 Apply f_2 to classify the examples in U, for each class c_i, pick n_i examples that f_2 most confidently classifies as class c_i, and add them to L.

6 until U becomes empty or a fixed number of iterations are reached

IV. EXPERIMENTAL ANALYSIS

A. *Experiment Data and Evaluation Index*

Comment content feature and evaluators' behavior are analyzed and extracted before the combination. In this sector, we obtained yearly comment data in representative Household items, Computer supplies and Clothes from Amazon. The experimental data set is shown as table. The comment data is marked fake comment, real comment and undetermined comment. Final comment type is judged by majority principle. For example, under condition that each comment is marked by fake comment at least two markers, this comment is fake comment.

Table 1. Statistical information in experiments

	Fake comment	Unlabeled comments	Total number
Comments	3522	6250	9766
Users	3316	5684	9067
Unique IP	1315	4526	5535
Comment of average user	1.056	1.005	1.028
Comment of each IP	2.682	1.369	1.354
Average word number of each comment	53.12	63.21	59.98

In order to comprehensively measure the methods effectiveness, on the basis of correct rate and recall rate in statistics, this paper adopts F value as final evaluating index for fake comment detection effect. The formula table of accuracy rate P, recall rate R and F value which are shown as follows:

$$P = \frac{TP}{TP + FP} \tag{3}$$

$$R = \frac{TP}{TP + TN} \tag{4}$$

$$F = \frac{2PR}{P + R} \tag{5}$$

T refers to fake comments in correct classification, N refers to fake comments quantity in wrong recognition and R refers to genuine comment quantity of the recognized fake comment.

B. *Results Analysis*

The method in this paper and feature extraction of reference are respectively applied to contrast fake comment detection effect of mixed data in three kinds of product comments of Amazon. The results are shown as figure 1. From this figure, recognition accuracy and value of the extracted features in this paper largely improve. Although the generated model files of training data from different products will affect other products' fake comment recognition, accuracy rate of average fake comment recognition is 83.2%, recall rate is 75.21% and F value is 78.63%. This indicates effectiveness of feature in this paper is perfect and its generalization in different types of products comment is better. Among fake comment information of 6 commodities, average proportion of dynamic fake comment in each comment information is large. The classification statistics is shown as figure 4. This implies importance of dynamic feature in distinguishing fake comment and normal comment.

Table 2. statistical table of experimental results

Training set	Test set	Fake comment detection		
		P(%)	R(%)	F(%)
Household items	Bed	71.81	90.13	80.32
	Chair	90.49	69362	78.71
	Curtain	80.95	66.68	73.13
	Wardrobe	77.23	65.52	70.88
Computer supplies	Mouse	72.12	75.32	73.94
	Monitor	92.65	79.01	85.32
	Keyboard	69.58	80.74	74.78
	Main engine	71.93	85.39	78.06
Clothes	Dress	84.12	51.42	63.87
	Short boots	96.75	79.46	85.59
	Cap	89.36	74.52	81.32
	T-shirt	81.22	78.33	79.77
	Average	95.40	75.12	78.54

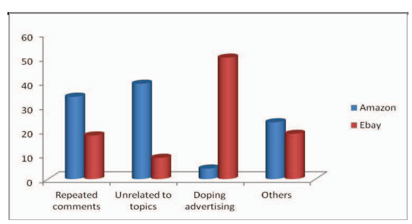

Figure 2. Structure of fake comment detecting method based on integrated features

Table 3 lists the experiment results comparison with different classifiers under condition that experiment data and the extracted features are totally the same. The classifiers involved mainly contain logistic regression model (LR), support vector machine (SVM) and AO-learning strategy. From data in table 3 and towards fake comment detection, AO-learning performance is superior to LR and SVM. It attaches more importance that AO-Learning introduction largely improves recall rate of system detection. This is because fake comment is difficult to be acquired through artificial marking so that training sample set of fake comment detection is not comprehensive and not accurate.

Table 3. Comparison result of fake comments detecting

Method	True comment			Fake comment		
	P	R	F	P	R	F
PU-LEA NB	0.868	0.759	0.801	0.768	0.884	0.823
CPU	0.885	0.752	0.851	0.781	0.902	0.835
FDMIF	0.865	0.772	0.824	0.780	0.915	0.841

V. CONCLUSIONS

When all commenting information in e-commerce platform is detecting fake comment, variety in commenting field and overall sparsity will result in decrease of recognition accuracy rate. If commodity with fake comment is firstly recognized, comment will be aimingly detected and this will largely increase recognition. This paper puts forward an abnormal scoring behavior analysis-based fake comment commodity recognition. On the basis of analyzing fake comment behavior, it adopts combination detection between static and dynamic feature to realize fake comment discovery. The experimental results show that this method can effectively detect fake comment target for online commodities.

REFERENCES

[1] Bhattarai A, Dasgupta D. A Self-Supervised Approach to Comment Spam Detection Based on Content Analysis. International Journal of Information Security & Privacy, 2011, 5(1):14-32

[2] Belikova T P, Yaroslavsky L P. Comments on ‘ ‘Image feature analysis and computer - aided diagnosis in digital radiography. I. Automated detection of microcalcifications in mammography. Medical Physics, 1989, 16(1):142

[3] J Wijngaard J, Stidham S. Forward Recursion for Markov Decision Processes with Skip-Free-to-the-Right Transitions, Part I: Theory and Algorithm. Proceedings of International Workshop on Semantic Media Adaptation and Personalization. IEEE, 200, pp.9-14

[4] Gordon A D, Cardelli L. Mobile Ambient. Lecture Notes in Computer Science, 1998, 240(99):140-155

[5] Gu X, Gu X. On the Detection of Fake Certificates via Attribute Correlation. Entropy, 2015, 17(6):3806-3837

[6] Reputation N, Forwarding M, Rate H A. A Fuzzy Collusive Attack Detection Mechanism for Reputation Aggregation in Mobile Social Networks: A Trust Relationship Based Perspective. Mobile Information Systems, 2016, 2016(4):1-16