International Journal of Emerging Engineering Research and Technology
Volume 2, Issue 7, October 2014, PP 57-63
ISSN 2349-4395 (Print) & ISSN 2349-4409 (Online)

Online Auction Fraud Detection

Shivani Deorukhakar [1], Namrata Khabiya[2], Ashwini Kulkarni[3], Abhishek Thorat[4]

[1,2,3,4] Computer Dept RMD Sinhgad School of Engg,
Pune, India (B.E)

Abstract: *In this project we are considering the problem of building online fraud detection in e-commerce websites. Increasing use of internet, online shopping and online auction have gained more importance. And at the same the criminals are also taking benefits to perform undesired activities to gain illegal profit. Hence fraud-detection moderation systems are commonly applied in practice to detect and prevent such illegal and fraud activities. In this project, we propose an online probity model framework which takes online feature selection, coefficient bounds from human knowledge and multiple instances learning into account simultaneously. By observations on a real-world online auction fraud detection data we show that this model can potentially detect more frauds and significantly reduce customer complaints compared to several base line models and the human-tuned rule-based system.*

Keywords: *Online Auction, Fraud-Detection, Online Modeling, Online Feature Selection, Multiple Instance Learning*

1. INTRODUCTION

Since the increasing use of internet mainly in electronic commerce, this is commonly called as e-commerce as in becomes more and more popular. Whenever we want to buy any electronic gadget we search for the best deal and features in the market t, it is followed by safe online purchase .The scene is not only applicable to electronic gadgets alone it also contains a wide range of products like home appliances ,books ,travelling packages or any software.it has become possible to buy any product u desire without actually touching the product physically .In traditional online shopping business model sellers as in sell their products or services at pre-set price, where buyers can choose what product best suites them which is of good deal .Online auction however is a different business model where the items are sold through price bidding. Usually while starting bidding we set an initial amount which is none other than starting bidding price among the bidder capable buyers in auction bid with each other. The buyer who bids with highest amount is called last bidder or winner. To assure safe bidding against fraud and to give confidence to online auction services as in E-commerce sites provide refund to victims for those who loss up to a certain amount. Online auction services and e-commerce sites adopt following approaches to control and prevent fraud. To buy a certain product from the online auction website they are to be validated with e-mail, SMS, or phone call, PAN card, license number etc. In this project, we build an online auction modeling for fraud detection, where hundreds and thousands of new auction cases are created every day. Human experts are also willing to test and see the results of online feature selection to monitor the effectiveness and stability as in of the current set of features, so as to understand the scope of frauds done by fraudulent sellers and further add or remove some features.

2. SCOPE OF PROJECT

This Project investigates the entry criteria for providing a new auction service through the real options approach, where the entry criteria is established by using an Online auctioning system designed for the use of normal users(individuals), Industrialists, Entrepreneurs, Organizations and Academicians under transaction rate uncertainty.

- Customer must have a unique and authorized User Id and password to login to the system

- If a wrong password is given more than three times in continuous, that account will be locked and the customer will not be able to use it. When an invalid password is entered a warning is given to the user that his account is going to get locked.
- After the valid user logs in his details like PAN card number, license number etc. are verified.
- FAQ test is also provided for end users benefit.
- For safe bidding user is required to pay initial amount for confirmation.

3. RELATED WORK

Online auction fraud is a threat to e-commerce. There are articles on websites to teach people how to avoid online auction fraud. Reputation systems are used extensively by websites to detect auction frauds, although many of them use different approaches. Other than reputation systems, machine learned models have been applied to moderation systems for detecting fraud. [2] Developed another simple approach that uses social network analysis. Other approaches proposed an offline regression modeling framework for the auction fraud detection moderation system which incorporates domain knowledge such as coefficient bounds and multiple instance learning. In this paper we treat the fraud detection problem as a binary classification problem .In this project Bayesian probit online model framework used to support binary response with details of model fitting via Gibbs sampling. The most frequently used models for binary classification include, probit regression [6], support vector machine (SVM) [7]. (SSVS) uses "spike and slab" prior so that the posterior of the coefficients have some probability being 0. Another approach is to consider the variable selection problem as model selection, i.e. put priors. When applied the probalities involved in Bayes's theorem may have any number of probality interpretations. In one of these interpretations the the theorem is used directly as part of a particular approach to statistical inference. in particular, with the Bayesian interpretation of probability, the theorem expresses how a subjective degree of belief should rationally change to account for evidence: this is Bayesian inference, which is fundamental to Bayesian statistics. However, Bayes's theorem has applications in a wide range of calculations involving probabilities, not just in Bayesian inference.

4. PROPOSED WORK

Our application is to detect online auction frauds of an auctioning site where new auction cases are posted every day. Every new case is sent to our proactive anti-fraud moderation system for a pre-screening score to assess the risk. The system has been classified into the following modules after a careful analysis.

4.1. Customer Module

A customer is one of the users who wish to shop online. For this purpose the customer will be provided with a personal account through registration. After successful registration, he will be provided with a gallery of different products from different sellers which include the product name, price, sellers' name etc. While buying a product a customer can view the percent of trustworthiness towards the product given by other users. After purchasing, a customer can also file complaint on that product where he feels uncomfortable provided with some options like-

- Products purchased by the buyer are not delivered by the seller.
- The delivered products do not match the descriptions that were posted by sellers.
- Malicious sellers may even post non-existing items with false description to deceive buyers
- General feedback as a complaint

4.2. Seller Module

The seller module includes different sellers who wish to sell their products. The seller needs to be approved by administrator after a seller submits his registration. A seller can add or delete or modify information about different items.

The different functionalities for seller are

- Can add a new a product
- Can delete a product
- Can place new offers to the product
- Can modify information related to the product such as price, basic information etc...

4.3. Admin Module

The administrative module includes an admin who acts as an intermediator between seller and the customer. An Adminis responsible to maintain the website information giving a trust to the customers. When a complaint is filed in the customer module, the admin takes the final decision whether to ban the product. If the admin feels all the products from particular seller mostly are not trusted he can also remove the seller and his related products.

4.4. Complaint Filing

Buyers can file complaints to claim loss if they are recently deceived by fraudulent sellers. The administrator views the various types of complaints and the percentage of various type complaints. The complaints values of a products increase some criteria value the administrator set the trust ability of the product as untrusted or banned. If the products set as banned, the user cannot view the products in the website.

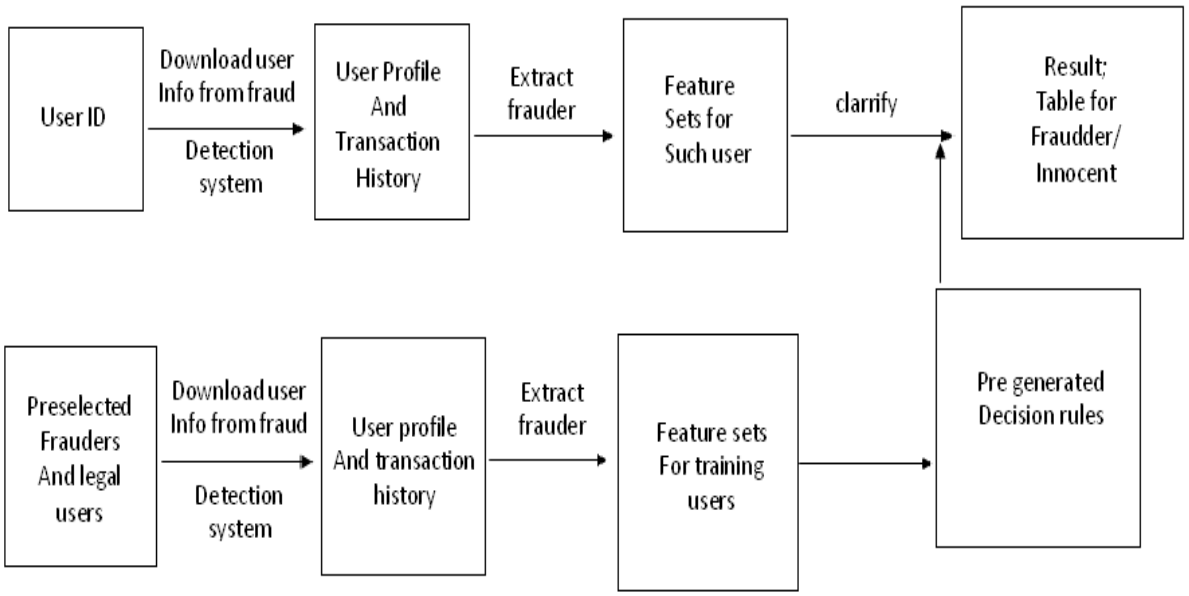

Fig1. *Process of detection fraud in system*

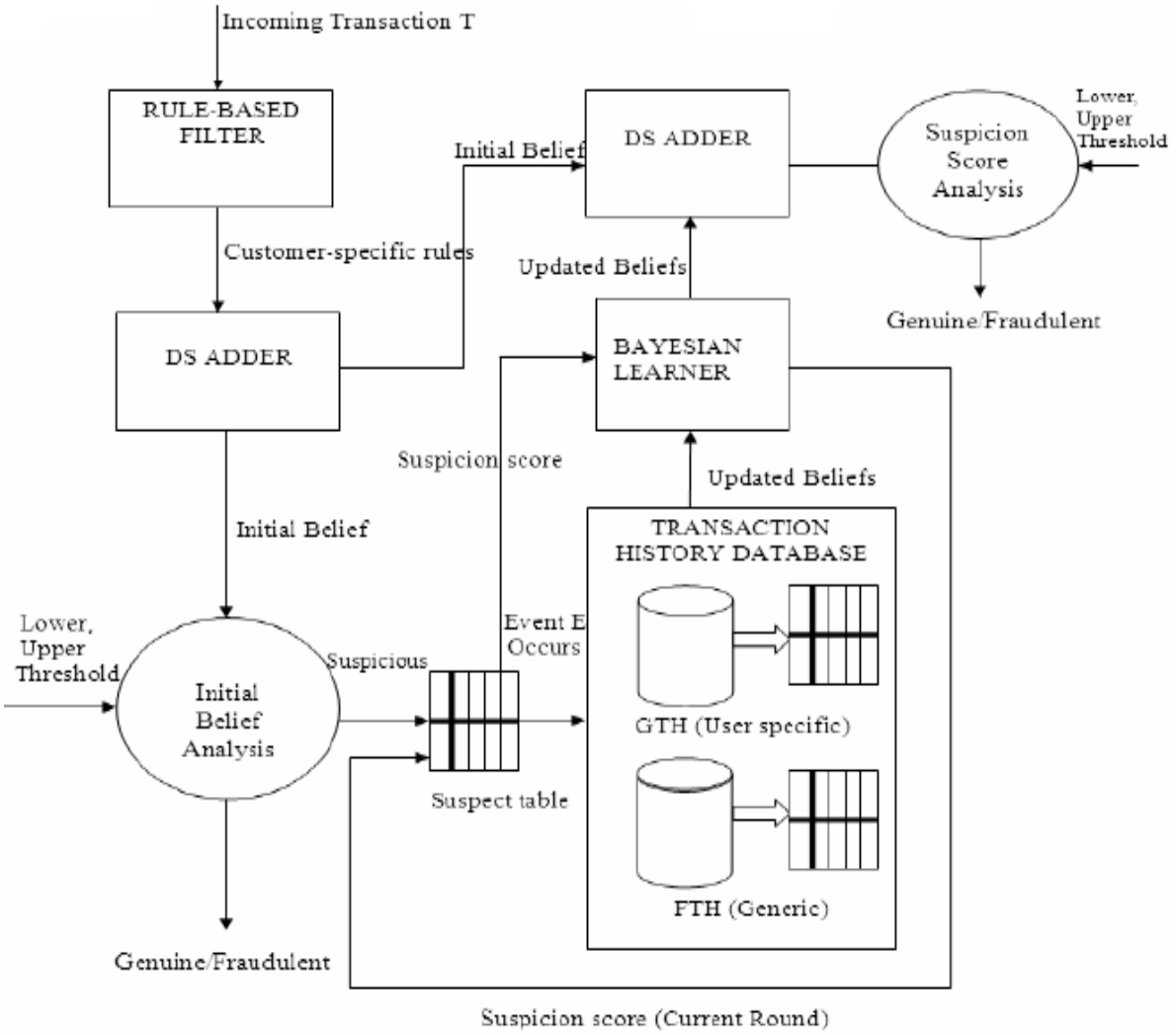

Fig2. *Block diagram of the proposed fraud detection system*

Using these specific attributes in our proactive moderation system for fraud detection, we build our Bayesian online modeling framework with details of model fitting via Gibbs sampling. Also extending it features with a selection bias fitting increases its adaptation to offline models too. The selection bias fitting approach is represented in Initial Belief Analysis.

5. SYSTEM FEATURES

The current system is featured by:

5.1. Rule-Based Features

Human experts with years of experience created many rules to detect whether a user is fraud or not. An example of such rules is “blacklist”, i.e. whether the user has been detected or complained as fraud before. Each rule can be regarded as a binary classification that indicates the fraud likeliness.

5.2. Linear Scoring Function

The existing system only supports linear models. Given a set of coefficients (weights) on features, the fraud score is computed as the weighted sum of the feature values.

5.3. Selective Labeling

If the fraud score is above a certain criteria, the case will enter a queue for further investigation by human experts. Once it is reviewed, the final result is evaluated and marked as Boolean, i.e. fraud or

clean. Cases with higher scores have high priorities in the queue. The cases whose fraud score are below the criteria are determined as clean by the system without any human judgment.

5.4. Fraud Churn

Once one case is labeled as fraud by human experts, it is very likely that the seller is not trustable and may be also selling other frauds; hence all the items submitted by the same seller are labeled as fraud too. The fraudulent seller along with his/her cases will be removed from the website immediately once detected.

5.5. User Feedback

Buyers can file complaints to claim loss if they are recently deceived by fraudulent sellers.

6. ALGORITHM

6.1. Online Probit Regression

Consider splitting the continuous time into many equal- size intervals. For each time interval we may observe multiple expert labeled cases indicating whether they are fraud or clear user. At time interval t suppose there are nt observations. Let us denote the i-th binary observation as yit. If yit = 1, the case is fraud; otherwise it is non-fraud. Let the feature set of case i at time t be xit. The probit model [3] can be written as

P [yit = 1|xit , βt] = Φ(x′ βt), where Φ(·) is the cumulative distribution function of the standard normal distribution

N (0, 1), and βt is the unknown regression coefficient vector at time t. Through data augmentation the probit model can be expressed in a hierarchical form as follows: For each observation i at time t assume a latent random variable zit. The binary response yit can be viewed as an indicator of whether zit > 0, i.e. yit = 1 if and only if zit > 0. If zit <= 0, then yit = 0. zit can then be modeled by a linear regression.

6.2. Stochastic Search Variable Selection (SSVS)

For regression problems with many features, proper shrinkage on the regression coefficients is usually required to avoid over-fitting. For instance, two common shrinkage methods are L2 penalty (ridge regression) and L1 penalty (Lasso). Also, experts often want to monitor the importance of the rules so that they can make appropriate adjustments (e.g. change rules or add new rules). However, the fraudulent sellers change their behavioral pattern quickly: Some rule-based features not support some cases. Therefore it is necessary to build an online feature selection framework that evolves dynamically to provide both optimal performance and intuition. In this paper we embed the stochastic search variable selection (SSVS) into the online probit regression framework described in Online Probit Regression.

6.3. Coefficient Bounds

Incorporating expert domain knowledge into the model is often important and has been proved to boost the model performance. In our moderation system, the feature set x is proposed by experts with years of experience in detecting auction frauds. Most of these features are in fact “rules”, i.e., any violation of one rule should ideally increase the probability of the seller being fraud to some extent. A simple example of such rules is the “blacklist”, i.e. whether the seller has ever been detected or complained as fraud before. However, for some of such rules simply applying probit regression as described in online probit regression or logistic regression might give negative coefficients, because

given limited training data the sample size might be too small for those coefficients to converge to right values, or it can be because of the high correlation among the features. Hence we bound the coefficients of the features that are in fact binary rules, to force them to be either positive or equal to 0. Note that this approach couples very well with the SSVS all the coefficients which were negative are now pushed towards zero.

6.4. Multiple Instance Learning

The procedure of expert labeling is done in the moderation system, we noticed that experts do the labeling in a "bagged" fashion: i.e. when a new labeling process starts, an expert picks the most "suspicious" seller in the queue and looks through all of his/her cases posted in the current batch (e.g. this day); if the expert determines any of the cases to be fraud, then all of the cases from this seller are marked as fraud. In literature the models to handle such scenario are called "multiple instance learning". Suppose for each seller i at time t there are Kit number of cases. For all the Kit cases the labels should be identical, hence can be denoted as yit. For probit link function, through data augmentation denote the latent variable for the l-th case of seller i as zilt.

7. CONCLUSION

Due to increasing use of internet, online shopping and online auction has taken a major concern in electronic-commerce today. Now-days its gaining importance and this lead to frauds taking place and an innocent victim suffers without notice. Many people are being involved in this kind of activity and makes online shopping an un-safe place to buy goods and service. This produces undesirable effects not only on auction participants and customers but also on the auction mechanism itself as a resource allocation market. But implementation of Online Fraud Detection has prevented ways to detect the fault in early stages and measure to prevent it. This proves to be a simplistic approach. To prevent in-auction fraud, robust auction rules need to be proposed by economists. On the computer technology side, there is a need of airtight transaction process design to foil the efforts of fraudsters. In this project, we describe the indicators of in-auction fraud, and pointed out that because no single indicator will be accurate enough to assure the presence of in-auction fraud, a combinatorial way using multiple indicators would be more effective and precise.

ACKNOWLEDGEMENT

We thank the editors and all anonymous referees for the careful review of this paper and the suggestions for improvements they provided. We also thank Mrs. Vanita Babanne for his editorial help, which significantly improved the presentation of this material.

REFERENCES

[1] http://www.ibisworld.co.uk/market research/ecommerce-online-auctions.html?partnerid=prweb

[2] http://www.google.com/finance?catid=usTRBC:5720103013&ei=zIDSUfiZFJ6wlgOABQ

[3] USA Today. How to avoid online auction fraud. http://www.usatoday.com/tech/columnist/ 2002/05/07/ yaukey.htm, 2002.

[4] Federal Trade Commission. Internet auctions: A guide for buyers and sellers. http://www.ftc.gov/bcp/ conline/pubs/online/auctions.htm, 2004.

[5] C. Chua and J. Wareham. Fighting internet auction fraud: An assessment and proposal.Computer, 37(10):31–37, 2004.

[6] P. Resnick, K. Kuwabara, R. Zeckhauser, and E. Friedman. Reputation systems. Communications of the ACM, 43(12):45–48, 2000.

[7] P. Resnick, R. Zeckhauser, J. Swanson, and K. Lockwood. The value of reputation on ebay: A controlled experiment. Experimental Economics, 9(2):79–101, 2006

AUTHORS' BIOGRAPHY

Miss. Shivani Bandu Deorukhakar is Completed Diploma In Computer Technology from Sou. Venutai Chavan Polytechnic Pune(MSBTE Affliated) Pursuing BE Computer from RMD Sinhgad School Of Engineering,Warje Pune(Pune University)

Miss. Ashwini Ashok Kulkarni is Completed Diploma In Computer Engineering from MIT Sri Savitribai Phule Polytechnic Pune(MSBTE Affliated) Pursuing BE Computer from RMD Sinhgad School Of Engineering,Warje Pune(Pune University)

Miss. Namrata Navneet Khabiya is Completed Diploma In Computer Engineering from MIT Sri Savitribai Phule Polytechnic Pune(MSBTE Affliated) Pursuing BE Computer from RMD Sinhgad School Of Engineering,Warje Pune(Pune University)

Mr. Abhishek Appasaheb Thorat is Pursuing BE Computer from RMD Sinhgad School Of Engineering,Warje Pune(Pune University)