2018 17th IEEE International Conference on Machine Learning and Applications

Auction Fraud Classification Based on Clustering and Sampling Techniques

Farzana Anowar
Department of Computer Science
University of Regina
Regina, Canada
fad469@uregina.ca

Samira Sadaoui
Department of Computer Science
University of Regina
Regina, Canada
sadaouis@uregina.ca

Malek Mouhoub
Department of Computer Science
University of Regina
Regina, Canada
mouhoubm@uregina.ca

***Abstract*—Online auctions created a very attractive environment for dishonest moneymakers who can commit different types of fraud. Shill Bidding (SB) is the most predominant auction fraud and also the most difficult to detect because of its similarity to usual bidding behavior. Based on a newly produced SB dataset, in this study, we devise a fraud classification model that is able to efficiently differentiate between honest and malicious bidders. First, we label the SB data by combining a hierarchical clustering technique and a semi-automated labeling approach. To solve the imbalanced learning problem, we apply several advanced data sampling methods and compare their performance using the SVM model. As a result, we develop an optimal SB classifier that exhibits satisfactory detection and misclassification rates.**

Keywords—auction fraud, shill bidding, supervised classification, hierarchical clustering, data labeling, data sampling.

I. Introduction

As reported by the Internet Crime Complaint Center of the Federal bureau of investigation, online auctions represent one of the top cybercrimes. In fact, e-auctions are vulnerable to dishonest moneymakers who can commit different types of malicious activities, such as pre-auction, post-auction and in-auction fraud. This vulnerability exists because of several reasons, such as high anonymity of users, flexibility in bidding, and cheap auction services. Commercial auction sites provide a tremendous volume of public data, and researchers can employ machine learning (ML) methods to effectively examine data of auctions and users in order to determine for examples the trends and online behavior of users as well as the popularity of products in the market. In our study, we focus on Shill Bidding (SB) because this kind of fraud does not leave concrete evidence. SB, which has been found to be predominant across many auctions, is very difficult to detect because of its resemblance to normal bidding behavior [1, 8]. Several empirical studies already demonstrated the presence of SB activities in different commercial auction sites [1, 2], and a number of lawsuits have been filed against fraudulent sellers and auction companies as well [3].

Our objective is to devise a SB classification model that is able to efficiently distinguish between normal and dishonest bidders. Nevertheless, unlike other fraud detection applications, such as fraud of credit cards, telecommunication and insurance, SB training datasets are very lacking because they are difficult to obtain. This difficulty is caused by several factors, such as determining relevant SB strategies, defining metrics to measure SB patterns, scraping data from auction sites, preprocessing auction and user data, and finally measuring the SB patterns from the collected data.

In the previous work [3], the authors crawled a large number of auctions from eBay for one of the most popular products in 2017. After that, they produced a high-quality SB (unlabeled) training data. In this present study, our aim is to develop a robust fraud classifier based on this new SB dataset. For this purpose, we combine unsupervised and supervised learning methods. First, with the help of a hierarchical clustering technique, we label the SB training data by using a specific labeling strategy [6].

Furthermore, any fraud training dataset has an imbalanced class distribution, and learning from imbalanced data deteriorates the predictive performance as shown in [4]. Additionally, when dealing with imbalanced data, the baseline classification models favor the majority class. This is a serious issue in fraud detection applications because on one hand the minority/fraud class tends to be misclassified as normal, and on the other hand the fraud class has a very high misclassification cost. Even though some SB classification models have been developed in the past, most of them did not address the imbalanced learning problem, which is one of the ten challenges in ML research [5]. To tackle this problem, we apply several advanced and intelligent data sampling techniques to our labelled SB dataset, and as a result we produce several SB training datasets. Finally, we conduct several empirical investigations to determine the optimal SB classifier. To this end, we first develop several fraud classifiers with the different sampled SB datasets. Based on the Randomized Search Cross Validation, we assess the predictive performance of the classifiers in terms of the detection and misclassification rates of fraudulent bidders.

A SB classifier should be able to operate in real-life scenarios. Still, past SB detection approaches did not implement any online detection strategy [6]. In our work, thanks to the auction-wise training features, we are able to suspend an auction when it is found to be infected by SB in order to avert financial loss for the auction winners. In fact, the learned classification model is to be deployed at the end of the bidding period to detect suspicious bidders before processing any payment. Nevertheless, a further investigation is necessary to confirm or reject the suspicion. When suspected bidders are found to be actual shills, actions are then taken, such as cancelling the infected auction and suspending the shill bidder's account.

II. Related Work

Very few research studies exist in the literature for SB classification as opposed to other fraud detection problems.

In [1], the authors presented “Real Time Self-Adaptive Classifier” to detect suspicious bidders using an incremental feedforward backpropagation Artificial Neural Network

978-1-5386-6805-4/18/$31.00 ©2018 IEEE
DOI 10.1109/ICMLA.2018.00061

(ANN). This classifier is trained with auctions extracted from eBay for the product '*Used Playstation 3*'. To label the training data, the authors first applied a hierarchical clustering algorithm, and then manually examined the bidding behavior in each cluster. We may mention that ANN suffers sometimes from local minima, and there are many parameters to tune, which makes ANN computationally costly. In addition, ANN does not explain the behavior of the network, which ultimately reduces its trust [18]. In [1], bidders are classified according to their participation in all the auctions. In this case, it is not possible to determine which auctions are infected by fraud. Consequently, money loss cannot be avoided.

In [7], the authors utilized the Hidden Markov Model (HMM) to detect bidding fraud by using only two parameters (the number of bids and the bid values). The general Markov model is organized into two-layered (training and detection) architecture. In the training layer, the authors utilized K-means clustering to obtain the initial probability set from the habit of bidding behavior for the authentication purpose. Then, in the detection layer, a behavioral approach has been used to analyze the bidding habits, and then depending on the outcome, users are characterized into three fraud categories (high, low and medium). Lastly, HMM is applied to these categories to detect fraudulant bidders. The authors did not specify the type of bidding fraud, and two parameters only do not offer enough information to detect fraud. Moreover, they did not show any experiments to validate the proposed framework.

Very recently, [6] proposed an SVM-based SB detection model, which overcome the issues of ANN and HMM. This research is the first one to address the class imbalance of auction fraud data. The authors applied three data sampling methods (SMOTE, SpreadSubsample and hybrid of both) to solve the imbalanced learning problem. Nevertheless, they built the SB dataset from a small number of auctions by including synthetic data for two missing attributes. For the SVM parameter tuning, the authors did not use any automated method to detect the optimal parameters to evaluate the performance metrics.

III. Overview on Bidding Fraud

Shill bidding is the most dominant in-auction fraud and also the most difficult to detect because it is similar to normal bidding behaviour [1]. Several empirical studies already demonstrated the presence of SB in different commercial auction sites [1, 16]. A shill bidder is a malicious user, the seller or his accomplice, who places many bids via phony accounts with the sole purpose of driving up the price of the product being auctioned. SB may cause a significant financial loss for purchasers in case of expensive products as well as products with an unknown value in the market. As mentioned in [8], excessive SB could lead to a market failure. We may mention that several lawsuits have been filed against fraudulent sellers and auction companies as well. As an example, in 2012, the online auction company TradeMe paid $70,000 for each victim after the investigation discovered that SB has been conducted by a motor vehicle trader from Auckland. Trade Me blocked the trader from using their website, and referred the case to the Commerce Commission for a further investigation [9]. In 2014, a lawsuit was filed against Auction.com by VRG in California claiming that the website allowed SB to happen. The bid of $5.4 million should have secured the property as the plaintiff declared, and yet the winning price was 2 million more. Auction.com was accused of helping the property owner, which is not fair for honest bidders. The California state passed a law on July 2015 requesting the property auctioneers to reveal the bids they submitted on behalf of the sellers [10].

In Table I, we present the most dominant SB patterns across many auctions [1, 2, 16]. Each pattern is unique as it represents one aspect of the bidding behavior and occurs in a certain auction stage. A pattern can be computed from the characteristics of an auction, or a bid, or a bidder. The uniqueness of SB patterns (as the training features) will help improving the predictive performance of our fraud classifier. The metrics to measure the SB patterns have been defined in [3, 16].

TABLE I. Characteristics of Shil Bidding Patterns

Name	Description	Motive	Level
Starting Price	Seller sets an unusually low starting price when compared to concurrent auctions	To attract people to the auction	-Auction -Early stage
Early Bidding	Bidder submits a bid in the early hours of the auction	To allure original bidders to take part in the auction	-Bid -Early stage
Bidding Ratio	Bidder participates much more as opposed to normal bidders	To raise the price and attract higher bids from honest participants	-Bid -Middle stage
Successive Outbidding	Bidder outbids oneself with consecutive small bids	To raise the price gradually	-Bid -Middle stage
Last Bidding	Bidder becomes inactive at the final stage of the auction	To prevent oneself from winning the auction	-Bid -Last stage
Winning Ratio	Bidder competes a lot in many auctions, but hardy wins any auctions	The target is not to win the auction but to increase the price	-Bidder -User history
Buyer Tendency	Bidder participates exclusively in auctions of few sellers rather than a diversified lot	Collusive act involving the fraudulent seller and an accomplice bidder	-Bidder -User history
Auction Bids	The number of bids in an auction with shilling is much more than that of concurrent auctions without shilling	To make the product appear more popular	-Auction -Completed Bidding

IV. Production of Shill Bidding data

In [3], the authors crawled from eBay a large number of auctions (2000) of the famous product iPhone 7 for a period of March to June 2017. They selected iPhone 7 for factors that may increase the chance of SB activities: 1) it is one of the most sold products in its category; 2) it attracted a high number of bidders and bids; 3) it has a good price range with the average of $578.64 (US currency). Certainly, more the product price is high, more possibility of fraud [8]; 4) it has different bidding durations: 1, 3, 5, 7 and 10 days. All these durations were considered because in long duration, a shill bidder may easily mimic normal behavior [8], and in short duration, fraudulent sellers may receive positive feedback ratings [11].

To obtain a high-quality auction dataset, the authors conducted an intensive and time-consuming preprocessing operation by 1) removing irrelevant and duplicated attributes as well as missing values (like IDs of bidders and sellers), and inconsistent values, 2) merging several attributes into a single attribute, 3) converting several attributes into proper formats, and 4) generating IDs for the auctions. Table II provides statistics about the preprocessed auction dataset of iPhone 7.

TABLE II. Preprocessed Auction Dataset of iPhone 7

Number of Auctions	807
Number of Records	15145
Number of Bidder IDs	1054
Number of Seller IDs	647
Average Winning Price	578.64
Average Auction Duration	7

Next, the authors implemented the SB metrics, and then measured each metric against each bidder in each auction. The resulting SB training dataset has a total of 6321 instances. An instance denotes the conduct of a bidder in a certain auction. It is a vector of ten elements: Auction ID, Bidder ID and the eight SB classification features. We examined the whole SB dataset for outliers (wrong values of features) and did not find any.

V. Labelling Fraud Data Via Clustering

Labeling any training datasets is a very tedious task. Most of the time, this is done manually, which is very time consuming. To make this task easier, we first cluster SB data into meaningful groups. A robust clustering technique produces clusters where the distance is maximized between the inter-clusters and minimized in the intra-cluster [12]. Therefore, a cluster is a group of data points that have the maximum similarity between them and dissimilarity to the data belonging to other clusters. Numerous clustering techniques have been proposed by researchers for data analysis, like K-means, Hierarchical Clustering (HC), SOM, EM, DBSCAN, BIRCH and CURE. Each clustering algorithm has its own pros and cons, and which one to select depends on factors like the size and dimensionality of the dataset. After analyzing these clustering techniques, we choose HC for several reasons: 1) in terms of cluster quality, HC is very efficient for datasets of medium size and dimensionality, like ours; 2) HC handles noisy data effectively; 3) HC does not necessitate the user to specify the number of clusters, and this is a significant advantage over flat clustering methods. An inappropriate choice of K may result in low cluster quality; 4) many researches [1, 6] have successfully used HC for partitioning fraud data.

The other advanced clustering methods, such as CURE, DBSCAN and BIRCH, are efficient in the context of large datasets. We work with hierarchical agglomerative clustering, which combines clusters into a new one at each step. We follow three phases to generate the clusters. Firstly, based on HC, we create a dendrogram, a tree diagram that illustrates the cluster arrangement. Secondly, we determine the optimal number of clusters**.** Finally, we generate the content of each cluster**.** Data clustering is based on different distance/similarity functions: Single Linkage, Complete Linkage, Average Linkage and Centroid Linkage. We produce four dendrograms w. r. t. the similarity measures and find out that Complete Linkage returns the best results**.** Next, we utilize the two testing methods, 'Average Silhouette' and 'Gap Statistic', to find out about the optimal number of clusters for our SB dataset. To do so, we try different number of clusters from 1 to 30 using Complete Linkage. The same optimal number of 10 has been returned by both testing methods. Lastly, we apply HC to the SB dataset with the optimal number of clusters and Complete Linkage in order to generate the content of the 10 distinct clusters. Here, a cluster consists of users with similar bidding behavior.

Next, we apply the labeling strategy introduced in [6]. In each cluster, we first categorize each fraud pattern into 'low' or 'high' category according to its average value of all the bidders in that cluster. Table IV provides two examples of cluster analysis and labeling. We also assign weight to each SB pattern as done in [16]. In Cluster 2, five shill patterns fall in the high category (all have an average value more than 0.7). Among them, the presence of high and medium weighted patterns indicates that bidders in this cluster are most probably shills. The other three patterns have a very low value. Now, in Cluster 3, almost all SB patterns fall in the low category (all have an average value less than 0.32), and the single high value pattern is of low weight. This means bidders in this cluster are behaving normally. In summary, according to the values and weights of the fraud patterns, we label the cluster as 'Normal' or 'Suspicious'. Table III presents all the clusters and their classes.

TABLE III. Results of Data Clustering and Labeling

Cluster Number	Cluster Size	Cluster Label
Cluster 0	4.33%	Suspicious
Cluster 1	30.28%	Normal
Cluster 2	3.32%	Suspicious
Cluster 3	13.95%	Normal
Cluster 4	5.72%	Normal
Cluster 5	0.46%	Suspicious
Cluster 6	1.80%	Suspicious
Cluster 7	23.98%	Normal
Cluster 8	8.35%	Normal
Cluster 9	7.78%	Normal

TABLE IV. DATA ANALYSIS AND LABELING

Cluster Number	Category	Weight	Size (%)	Class
Cluster 2	**Low Average Value**		3.32%	Suspicious
	Bidder Tendency (0.387139)	Medium		
	Bidding Ratio (0.344252)	Medium		
	Auction Bids (0.375692)	Low		
	High Average Value			
	Successive Outbidding (0.969048)	High		
	Last Bidding (0.892746)	Medium		
	Starting Price Average (0.70136)	Low		
	Early Bidding (0.862606)	Low		
	Winning Ratio (0.899158**)**	Medium		
Cluster 3	**Low Average Value**		13.95%	Normal
	Bidder Tendency (0.138401)	Medium		
	Bidding Ratio (0.081071)	Medium		
	Successive Outbidding (0.019841)	High		
	Auction Bids (0.31197)	Low		
	Last Bidding (0.085009)	Medium		
	Early Bidding (0.070112)	Low		
	Winning Ratio (0.224646)	Medium		
	High Average Value			
	Starting Price (0.973062**)**	Low		

VI. IMBALANCED LEARNING PROBLEM

Table V shows the class distribution of our SB training dataset with an imbalance ratio of 9:1. Here the normal class is over-represented. Learning from imbalanced data deteriorates the predictive performance as shown in [4]. Additionally, with imbalanced data, the minority class (usually the class of interest) will be misclassified because baseline classifiers favor the majority class. In fraud detection domain, having screwed class distribution is a serious concern because the fraudulent class tends to be misclassified as 'normal'. Moreover, classifiers that learn to always predict the majority class may obtain 99% accuracy, but such classifiers are not helpful in identifying the fraud class. Also, classifiers tend to obtain poor accuracy when identifying the fraud class, but it is the fraud class that carries the highest cost of misclassification.

TABLE V. ANALYSIS OF GENERATED CLUSTERS

Total Instances		**Clusters**	
Normal	***Suspicious***	***Normal***	***Suspicious***
5694 (91%)	627 (9%)	6	4
6321 (total instances)		10 (total clusters)	

Classification algorithms are most performing when the counts of instances of the two classes are roughly equal. In the context of medium sized datasets, the most popular way to deal with imbalanced data is to rebalance the class distribution. Generally speaking, data sampling techniques aim to remove instances from the majority class (under-sampling) or add instances to the minority class (over-sampling) or do both of them.

SMOTE is the one of the most famous intelligent data over-sampling approaches [13]. SMOTE generates synthetic positive instances and places them between a positive instance and its K neighbors. Moreover, several extensions of SMOTE have been proposed to achieve a higher efficiency, and the most performing ones are SMOTE-ENN and SMOTE-TomekLink. Both methods combine under and over-sampling, which makes them hybrid methods [15]. Indeed, in addition to over-sampling the minority class via SMOTE, the distance-based method TomekLink and ENN under-sample the majority class by deleting instances that from TomekLink links i.e. borderline and noisy samples that affect the predictive performance. In fact, ENN eliminates more instances than TomekLink.

Regarding data under-sampling, NearMiss technique is commonly used. It keeps only the instances from the majority class whose mean distance to the K-nearest neighbor is the lowest. One major problem of using under-sampling methods is that significant information may be lost from the majority class. This can cause overly general rules since under-sampling removes instances from the majority class. Hence to overcome this problem, ClusterCentroids method has been introduced [17]. ClusterCentroids under-samples the majority class by replacing cluster of samples by the cluster of centroids using K-means algorithm. In Table VI, we apply the five sampling methods, and consequently produce five SB training datasets with different sizes.

VII. OPTIMAL CLASSIFICATION OF SHILL BIDDING

To develop the five SB classifiers, we select the Support Vector machine (SVM) model for the following important reasons:

- Through empirical studies, it has been shown that combining data sampling with SVM leads to higher classification performance [5].
- Auction data are difficult to linearly separate [1]. This issue can be easily solved with the kernel SVM.
- SVM is very efficient with medium sized training datasets like ours.

TABLE VI. SAMPLING OF SHIL BIDDING DATASET

Method	Sampling Type	Normal Bidders	Suspicious Bidders
SMOTE	Over-Sampling	5694	5694
SMOTE-ENN	Hybrid Sampling	5633	4888
SMOTE-TomekLink	Hybrid Sampling	5688	5688
NearMiss	Under-Sampling	627	627
ClusterCentroids	Under-Sampling	627	627

We use the Randomized Search Cross Validation (CV) to build the optimal fraud classifier for each of the five sampled training datasets by selecting the best kernel function and tune efficiently the two SVM hyper-parameters. Since auction data are non-linearly separable, we use the non-linear kernel RBF shown to be effective in many classification and fraud detection problems**.** Also, when we used Randomized Search to select the best kernel, it provides us with the RBF as the optimal one. Moreover, when the number of instances is high and comparatively the number of features is much smaller, like our SB training dataset, RBF is the most suitable for classification.

Next, we search for the best values of the misclassification cost parameter C and the free kernel parameter γ. This is accomplished with the help of Randomized Search CV that tries different values from the range of C and γ. The range that we define for C is [1.0, 100.0] and for γ [0.1, 0.9].

In fraud detection area, we are more concerned about the suspicious class rather than the normal class as it has a much higher misclassification cost. So, we focus on the detection rates of suspicious bidders, like Recall and Precision. Recall means how sensitive is the classifier in detecting shill bidders, and Precision means how precise is the classifier in detecting shill bidders is. These two metrics are also used to calculate another important metric known as F-measure, which has been widely used for fraud detection [11, 14].

Table VII exposes the performance results for 5-fold and 10-fold CV based on the five sampled SB datasets. As we can observe, the hybrid method SMOTE-ENN with 10-fold CV outperforms the other data sampling techniques. Hence, we obtain the best classification model with K = 10, C = 11.85 and γ = 0.14.

In the fraud detection area, researchers are more interested in the 'Recall' metric [6]. We keep in mind that we have generated five classifiers (with 5 and 10-fold CV for each), and to determine the optimal classifier, we consider the overall performance metric F-measure. We can see that SMOTE-ENN with 10-fold CV produces the best Recall with a value of 0.850. It also produces overall better Precision and F-measure with values of 0.940 and 0.890 respectively. We may note that NearMiss with 10-fold CV and SMOTE-ENN with 5-fold CV produce better precisions with values of 0.960 and 0.945 respectively than SMOTE-ENN with 10-fold. However, SMOTE-ENN with 10-fold CV has improved the F-measure by 4.5% than NearMiss and 2% than SMOTE-ENN with 5-fold CV. It has also improved Recall by 2% than NearMiss with 10-fold CV and 3% than SMOTE-ENN with 5-fold CV. It is not rare to have a precision and recall trade-off in classification problems [15]. However, this is not surprising that SMOTE-ENN outperforms the other techniques since it has already been shown in [5] that SMOTE-ENN performs outstandingly well on 11 real-world churn datasets.

TABLE VII. CLASSIFICATION PERFORMANCE RESULTS

Sampling Method	CV (K)	Precision	Recall	F-Measure
SMOTE	5	0.935	0.785	0.840
	10	0.890	0.835	0.860
SMOTE-ENN	5	0.945	0.820	0.870
	10	**0.940**	**0.850**	**0.890**
SMOTE-TomekLink	5	0.915	0.815	0.855
	10	0.855	0.830	0.845
NearMiss	5	0.860	0.825	0.845
	10	0.960	0.830	0.845
ClusterCentroids	5	0.855	0.815	0.835
	10	0.885	0.845	0.865

We can see that when CV is 10, the SVM model performs the best in every case in terms of Recall. The latter performs better with 5-fold CV. We may mention that 10-fold CV is usually employed to avoid the chance of the model over-fitting the training data. In Table VII, the Recall rate of 0.850, Precision rate of 0.940, and F-measure of 0.890 indicate that the SVM classifier gives a good justice to the fraudulent class with very satisfactory detection rates.

TABLE VIII. EVALUATION OF LEARNED MODEL

Optimal SB Classifier (CV=10, C=11.85, γ= 0.14	
Performance Measurement	**Randomized Search CV**
AUROC	0.849
False Positive Rate	0.06
False Negative Rate	0.15

As presented in Table VIII, the optimal learning SB model exhibits a very good overall performance with AUROC equals to 0.849. Additionally, we compute the two misclassification errors, False Negative Rate (FNR) and

False Positive Rate (FPR). FNR denotes the percentage of fraudulent instances that have been incorrectly identified as normal and FPR the percentage of normal instances that have been misclassified as fraudulent. We obtain FPR of 0.06, which means only 6% of normal bidders have been labeled as suspicious erroneously, and FPR of 0.15, which says that 15% of suspicious bidders have not been detected as suspicious and labeled as Normal wrongly.

VIII. Conclusion

In online auctions, before processing the payment of products (goods and services), it is imperative to detect the shill bidding fraud in order to avert financial losses for the auction winners. For this purpose, we developed an efficient and robust shill bidding classification model. There are limited studies on shill bidding because training data are difficult to obtain. Firstly, we properly applied a hierarchical clustering technique to partition training data into clusters of bidders with similar behavior. Subsequently, we utilized a semi-automated approach to label the clusters of bidders into normal or suspicious. Moreover, to solve the imbalanced learning problem, we employed several advanced data sampling methods; under-sampling (NearMiss and ClusterCentroids), over-sampling (SMOTE) and hybrid sampling (SMOTE-ENN and SMOTE-TomekLink). To compare the efficiency of these methods, we developed several SVM-based fraud classifiers by using Randomized Search Cross Validation with 5-fold and 10-fold. Lastly, we determined the optimal shill bidding classifier that exhibits a very good testing performance in terms of detection rates (Precision, Recall, F-measure and AUROC) and misclassification rates (FPR and FNR). The experimental results demonstrate that SMOTE-ENN is the most performing data sampling method.

This present work can lead to a few research possibilities, including:

- In real-life scenarios, hundreds of auctions occur simultaneously. To scale up with this huge traffic problem, our fraud classifier maybe deployed on multiple autonomous agents. More precisely, for each new auction, an agent is created dynamically to classify its participants, and then the agent is destroyed once the auction is completed or cancelled as done in [16].
- Another important research is to develop a shill bidding classifier that is able to evolve constantly with new bidding data and user trends. To this end, we will implement an adaptive fraud classifier based on SVM incremental and decremental learning.

References

[1] B. Ford, H. Xu and I. Valova, "A Real-Time Self-Adaptive Classifier for Identifying Suspicious Bidders in Online Auctions", *The Computer Journal*, vol. 56, no. 5, pp. 646-663, Elsevier, 2012.

[2] J. Trevathan and R. Wayne, "Detecting SB in online English auctions", *Handbook of research on social and organizational liabilities in information security, Chapter: 27, Publisher: IGI Global, Press Editors: Manish Gupta, Raj Sharman*, pp. 446-470, 2008**.**

[3] A. Alzahrani and S. Sadaoui, "Scraping and Preprocessing Commercial Auction Data for Fraud Classification", *Technical Report CS 2018-05*, pp. 1-17, arXiv preprint: 1806.00656, 2018.

[4] S. Ganguly and S. Sadaoui, "Classification of Imbalanced Auction Fraud Data", *Canadian Conference on Artificial Intelligence, (CAI),* pp. 84-89, Springer**,** 2017.

[5] B. Zhu, B. Baesens and S. vanden Broucke, "An empirical comparison of techniques for the class imbalance problem in churn prediction", *Information Sciences*, vol. 408, pp. 84-99**,** Elsevier, 2017.

[6] S. Ganguly and S. Sadaoui, "Online Detection of Shill Bidding Fraud Based on Machine Learning Techniques", Proc. of 31st *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems,* (IEA-AIE), pp. 303-314, Springer, 2018.

[7] P. Gupta and A. Mudra, "Online in-auction fraud detection using online hybrid model", *International Conference on Computing, Communication & Automation (ICCCA)*, pp. 901-907, IEEE, 2015.

[8] F. Dong, S. Shatz and H. Xu, "Combating online in-auction fraud: Clues, techniques and challenges", *Computer Science Review*, vol. 3, no. 4, pp. 245-258, 2009.

[9] J. Duffy, "Shill bidders - not welcome at Trade Me | Trust & Safety", *Trademe.co.nz*, 2018. [Online]. Available: https://www.trademe.co.nz/trust-safety/2012/9/29/shill-bidding. [Accessed: June 2018].

[10] R. Morgan, *Nypost.com*, 2018. [Online]. Available: https://nypost.com/2014/12/25/lawsuit-targets-googles-auction-com. [Accessed: June 2018].

[11] J. Chang and W. Chang, "Analysis of fraudulent behavior strategies in online auctions for detecting latent fraudsters", *Electronic Commerce Research and Applications*, vol. 13, no. 2, pp. 79-97, 2014.

[12] A. Sorin Sabau, "Survey of clustering based financial fraud detection research", *Informatica Economică*, vol. 161, pp. 110-122, 2012.

[13] A. Fernandez, S. Garcia, F. Herrera and N. V. Chawla, "SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary", *Journal of Artificial Intelligence Research (JAIR)*, vol. 61, pp. 863-905, 2018.

[14] S. Zhang, S. Sadaoui and M. Mouhoub, "An Empirical Analysis of Imbalanced Data Classification", *Computer and Information Science*, vol. 8, no. 1, pp. 151-162, 2015.

[15] R. Pierre, "Detecting Financial Fraud Using Machine Learning: Winning the War Against Imbalanced Data", Towards Data Science, 2018. [Online]. Available: https://towardsdatascience.com/detecting-financial-fraud-using-machine-learning-three-ways-of-winning-the-war-against-imbalanced-a03f8815cce9. [Accessed: June 2018].

[16] S. Sadaoui and X. Wang, "A dynamic stage-based fraud monitoring framework of multiple live auctions", *Applied Intelligence*, vol. 46, no. 1, pp. 197-213, SpringerLink, 2016.

[17] M. Rahman and D. Davis, "Cluster Based Under-Sampling for Unbalanced Cardiovascular Data", *Proc. of the World Congress on Engineering (WCE),* vol. 3, pp. 3-5, 2013.

[18] M. Mijwel, "Artificial Neural Networks Advantages and Disadvantages", 2018. [Online]. Available: https://www.linkedin.com/pulse/artificial-neural-networks-advantages-disadvantages-maad-m-mijwel/. [Accessed: June 2018].