Two Step Graph-based Semi-supervised Learning for Online Auction Fraud Detection

Phiradet Bangcharoensap[1], Hayato Kobayashi[2], Nobuyuki Shimizu[2], Satoshi Yamauchi[2], and Tsuyoshi Murata[1]

[1] Tokyo Institute of Technology, Meguro, Tokyo 152-8552, Japan
phiradet.b@ai.cs.titech.ac.jp, murata@cs.titech.ac.jp
[2] Yahoo Japan Corporation, Minato, Tokyo 107-6211, Japan
{hakobaya,nobushim,satyamau}@yahoo-corp.jp

Abstract. We analyze a social graph of online auction users and propose an online auction fraud detection approach. In this paper, fraudsters are those who participate in their own auction in order to drive up the final price. They tend to frequently bid in auctions hosted by fraudulent sellers, who work in the same collusion group. Our graph-based semi-supervised learning approach for online auction fraud detection is based on this social interaction of fraudsters. Auction users and their transactions are represented as a social interaction graph. Given a small set of known fraudsters, our aim was to detect more fraudsters based on the hypothesis that strong edges between fraudsters frequently exist in online auction social graphs. Detecting fraudsters who work in collusion with known fraudsters was our primary goal. We also found that *weighted degree centrality* is a distinct feature that separates fraudsters and legitimate users. We actively used this fact to detect fraud. To this end, we extended the *modified adsorption* model by incorporating the weighted degree centrality of nodes. The results, from real world data, show that by integrating the weighted degree centrality to the model can significantly improve accuracy.

Keywords: Online Auction Fraud Detection, Graph-based Semi-supervised Learning, Weighted Degree Centrality

1 Introduction

Over the last decade, online auctions have quickly become popular e-commerce services. The extensive profits attract many users to commit fraud in online auction websites. Online auction fraud is increasingly recognized as one of serious global concerns.

Generally, online auction fraud can be categorized into three types, according to the time when the fraudulent activity is committed: pre-auction, in-auction, and post-auction [6]. Pre-auction fraud occurs prior to an auction, for example selling of low quality product. Post-auction frauds are committed afterwards, such as non-delivery of products. Both pre-auction and post-auction frauds can

be directly verified with physical evidence. The remaining type of fraud is in-auction, which is the main target of this research. There are many kinds of in-auction fraud, as shown in Figure 1. The main focus of this research was *competitive shilling* in which fraudsters participate in their own auction as bidders with another user ID in order to drive up the final price. When such a fraud takes place, a legitimate winner has to pay more than a reasonable final price. Hereafter, the term *fraud* refers to this definition.

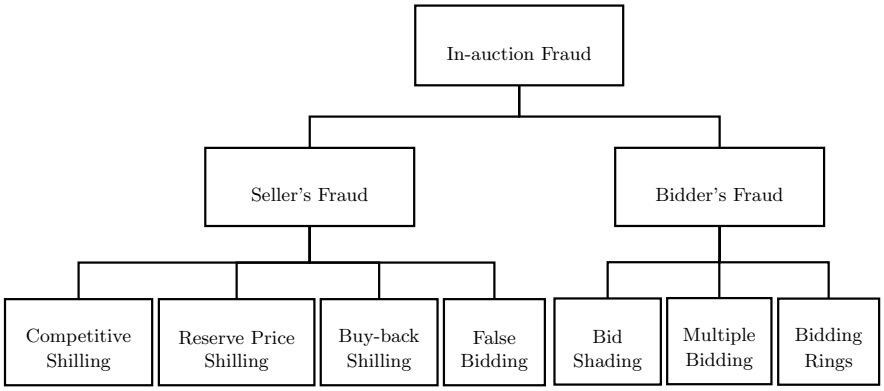


Fig. 1: Categorization of in-auction fraud [6]

Fraudulent bidders tend to frequently bid in auctions hosted by a particular seller(s) working in the same collusion group. Therefore, there is a very high tendency that a connections between fraudsters exist. In network science, we call this phenomenon *Homophily*. If we represent auction users and their activities as a social graph, groups of fraudulent users working in collusion should have strong links. It is analogous to one of the key assumptions of graph-based semi-supervised learning models (graph-based SSL). In general, graph-based SSL models try to assign a similar label to adjacent nodes. In other words, the models try to maintain the *smoothness* between adjacent nodes. This analogy, between *homophily* and *smoothness*, motivates us to investigate the potential of graph-based SSL models in online auction fraud detection.

There are two main contributions of this study. First, to the best of our knowledge, no study has been conducted to apply a graph-based SSL model for solving this serious Internet crime. We discuss the application of the state-of-the-art graph-based SSL model called modified adsorption (MAD) [14] to detect auction fraud. Furthermore, we found that the sum of the interactions between nodes with their neighbors can be used to distinguish between legitimate users and fraudsters. This sum is called *weighted degree centrality*. We argue that the weighted centrality of fraudsters is considerably higher than that of typical users. This fact alone sheds lights on the behaviors and social interactions of auction users, contributing to our understanding of the Web and its users. Even

though MAD involves edge weights as one type of information for propagating labels, a higher total weight of edges does not imply a higher likelihood of there being a fraud in the context of MAD. Therefore, we extended the model by incorporating the weighted degree centrality in the sense that it can be used to detect fraud. Our extended model, called the 2-STEP model emphasizes that a higher weighted degree centrality implies a higher chance of being fraudsters. This is our second contribution which involves the social behavior to detect auction frauds. According to experiments on real world data, our 2-STEP model significantly increases result accuracy.

The remainder of this paper is divided into six sections. We begin by giving details about the data we obtained from an online auction site in Section 2. In Section 3, we describe our proposed approach. We explain the performance evaluation of our approach in Section 4. Next, we discuss the results from the evaluation and discuss a possible extension for future work in Section 5. We describe related work in Section 6. Finally, we conclude our paper in Section 7.

2 Data Description

We first discuss the details of the data used in this research. The data are auction transactions, a set of known fraudsters, and a set of trustworthy users. The transactions were transformed into a social interaction graph. Section 2.2 gives the formal definition of the graph.

2.1 Resources

The following three sub-sections give more detail about the data used in this research.

Online Auction Transaction We obtained online auction transactions from YAHUOKU![3], which is one of the largest online auction sites in Japan and operated by Yahoo Japan Corporation. The dataset contains comprehensive bidding and selling activities on the website. Each record is a five-element tuple — (selling time, product ID, seller ID, bidding time, bidder ID). All user IDs are anonymized for preserving privacy. It is possible that the seller ID or bidder ID of two different transactions are identical. One user ID can be either of a seller or bidder. There are around 16 million transactions with around 3 million products and 2 million users.

Set of Known Fraudsters This set consists of IDs of users suspected to be fraudsters according to the definition of competitive shilling in Section 1. It is important to note that this set includes only a partial set of possible fraudsters because it is almost impossible to extract all fraudsters in this dataset due to its

[3] http://auctions.yahoo.co.jp/

size. The detection of fraudsters incurs high costs because it is performed manually. Approximately, 550 users are listed as fraudsters. The detailed description about ground-truth labeling cannot be disclosed since it is confidential business information. The set is used for training models and measuring performance. The *Set of Known Fraudsters* is referred to as $\mathcal{F}$.

Set of Store Users Some online auction user IDs are registered as official stores. These users can be placed on a whitelist, which contains trustworthy users who are unlikely to commit fraud. In our auction transactions, around 10,000 accounts are registered as stores. The *Set of Store Users* is referred to as $\mathcal{S}$.

2.2 Graph Construction

In this research, we represent online auction transactions as a weighted undirected graph $G(V, E, \mathbf{W})$, where V is the set of n nodes, E is the set of edges, and $\mathbf{W} \in \mathbb{R}^{n \times n}$ is the edge weight matrix ($\mathbb{R}$ is the set of real numbers). A node $v \in V$ represents an auction user ID. The set $E \subset V \times V$ represents interaction between nodes. An edge $e = (u, v) \in E$ indicates that u has a bid on an auction hosted by v, or vice versa. Each edge weight $\mathbf{W}_{uv} \in \mathbb{R}_+$ reflects the total number of u's products that are bidded by v. To remove noise, users participating in less than five transactions were removed. Finally, the graph contains around 0.8 million nodes and 3 million edges.

The largest group of nodes contains users who have never hosted any auction, but have only bid. This group of nodes occupies around 70% of the entire graph. Let us define this group of users as *bidder*. Another group contains users who have never bid on any auction, but only hosted. Approximately 15% of nodes fit into this category. We call this category *seller*. There is no link between nodes within the *bidder* and within *seller* groups. The last group contains users who both host and bid — *mixed*. Nodes in *mixed* can link with the previous two groups and within *mixed* themselves. Suppose we represent the graph as a directed graph whose edges originating from a *bidder* to a *seller*. It is obvious that the *seller* should not have any outgoing edges. Our model is based on an information propagation model. The propagation process cannot flow information to the entire graph when the graph contains many sink nodes. Therefore, we represent the transactions as an undirected graph. Figure 2 shows the degree distribution of our graph.

3 Proposed Approach

In this section, we now formally define the problem we want to solve and propose solutions.

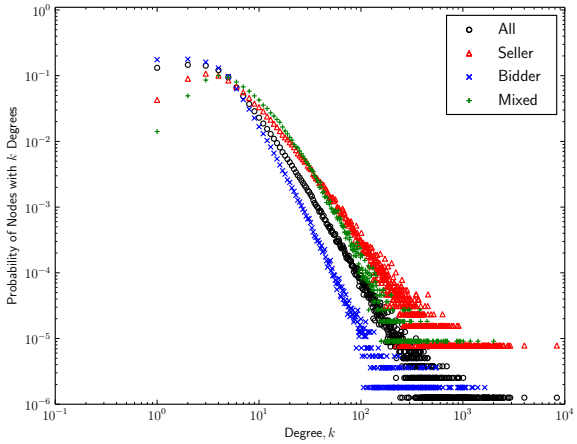


Fig. 2: Degree (k) distribution of the YAHUOKU! network

3.1 Problem Definition

The definition of the problem that we try to solve in this paper is summarized as below. We construct a weighted undirected graph $G(V, E, \mathbf{W})$ from a real world online auction dataset as described in Section 2.2. We assign a score, indicating likelihood of committing fraud activity, to all nodes. The score assignment is based on label propagation approach as described in Section 3.3. Nodes are ranked according to the score in descending order.

Input: A weighted undirected graph $G(V, E, \mathbf{W})$, a set of known fraudsters $\mathcal{F}$, and a set of store users $\mathcal{S}$.
Output: An ordered list of $\{v | v \in V \wedge v \notin \mathcal{F} \cup \mathcal{S}\}$.
Goal: Actual fraudulent nodes are expected to be ranked at the top of the output ordered list.
Approach: Propagating label information from seed known fraudsters and legitimate users to unknown users.

3.2 Modified Adsorption (MAD)

MAD is a graph-based semi-supervised learning model proposed by Talukdar and Crammer [14]. This research adopted the modified adsorption (MAD) to propagate information from known fraudsters to the whole graph. We used the Junto Label Propagation Toolkit[4] as an implementation of MAD.

[4] https://github.com/parthatalukdar/junto/

The MAD takes as the input a weighted undirected graph. The weight of edges represents the degree of similarity or correlation between nodes. A few nodes, or instances, are labeled — called seed nodes. MAD propagates labels from the few seed nodes to all nodes. Finally, all nodes are assigned a score indicating the likelihood of being each label. To deal with noisy initial labels, MAD allows the initial labels to change.

The model trade offs between three requirements: *accuracy* — the initial labels of seed nodes should be retained, *smoothness* — similar labels should be assigned to neighbor nodes, and *regularity* — output labels should be as uninformative as possible. We denote $\mathcal{L}$ as the set of m possible labels, and $\mathbf{M}_l$ as the l^{th} column of any matrix $\mathbf{M}$. Given a weighted undirected graph $G(V, E, \mathbf{W})$, where $|V| = n$, these three requirements can be expressed as a convex optimization problem as

$$\min_{\hat{\mathbf{Y}}} \sum_{l \in \mathcal{L}} \Big[\mu_1 (\mathbf{Y}_l - \hat{\mathbf{Y}}_l)^T \mathbf{S} (\mathbf{Y}_l - \hat{\mathbf{Y}}_l)) + \mu_2 \hat{\mathbf{Y}}_l^T \mathbf{L} \hat{\mathbf{Y}}_l + \mu_3 \left\| \hat{\mathbf{Y}}_l - \mathbf{R}_l \right\|^2 \Big], \quad (1)$$

where μ_1, μ_2, and μ_3 are hyperparameters, $\mathbf{Y} \in \mathbb{R}_+^{n \times (m+1)}$ stores initial label information, $\hat{\mathbf{Y}} \in \mathbb{R}_+^{n \times (m+1)}$ stores the output soft label assignment, $\mathbf{S} \in \mathbb{R}^{n \times n}$ indicates the position of seed nodes, $\mathbf{L} \in \mathbb{R}^{n \times n}$ is the Laplacian derived from the given G, and $\mathbf{R} \in \mathbb{R}^{n \times (m+1)}$ is the per-node label prior matrix, which is strongly related to an abandon action in random-walk. Each row of the matrices is associated with each $v \in V$. MAD introduces a *dummy* label in addition to the labels in $\mathcal{L}$, then each column of $\mathbf{Y}$, $\hat{\mathbf{Y}}$, and $\mathbf{R}$ is associated with $l \in \mathcal{L}$ and the *dummy* label. The intuition of *dummy* is that it is the exceptional case of all possible labels $\mathcal{L}$. The score of *dummy* is high when weights of edges originating from the node tends to be uniformly distributed, in other words the entropy of the weights is high. In terms of scalability, it has been proven that MAD is parallelizable in MapReduce [15].

3.3 MAD for Online Auction Fraud Detection

In this section, we discuss how to apply MAD to auction fraud detection. The key idea is to use the information from the set of known fraudsters and set of store sellers to assign initial labels. We denote the set of possible labels $\mathcal{L}$ as {*fraud*, *legitimate*}. We denote the 1st, 2nd, and 3rd column of $\mathbf{Y}$, $\hat{\mathbf{Y}}$, and $\mathbf{R}$ as associated with labels *fraud*, *legitimate*, and *dummy*, respectively.

MAD allows embedding world knowledge to a model by putting weights over initial labels. Precisely, it assigns a non-negative number to an element $\mathbf{Y}_{vl}$ in matrix $\mathbf{Y}$, where $\mathbf{Y}_{vl}$ is the v^{th} row and l^{th} column of $\mathbf{Y}$. In other words, $\mathbf{Y}_{vl}$ is the weight of node v over the label l. We create $\mathbf{Y}$ with conditions as

$$\mathbf{Y}_{vl} = \begin{cases} \alpha & \text{if } l = 1, v \in \mathcal{F}; \\ \beta & \text{if } l = 2, v \in \mathcal{S}; \\ 0 & \text{otherwise,} \end{cases} \quad (2)$$

where α and β are parameters, $\mathcal{F}$ is the set of known fraudsters, and $\mathcal{S}$ is the set of store users. The parameters α and β reflect the degree of association between v and the labels *fraud* and *legitimate*, respectively. Thus, seed nodes are $V_l = \{v|v \in V \wedge v \in \mathcal{F} \cup \mathcal{S}\}$. Let V_u denotes unlabeled nodes such that $V_u = V - V_l$, typically $|V_l| \ll |V_u|$. As previously mentioned, MAD outputs a soft label matrix $\hat{\mathbf{Y}}$ as a result of the label propagation process. Now, we assign each v a score reflecting the likelihood of being fraudsters, called *fraud score*, as

$$\varphi(v, \hat{\mathbf{Y}}) = \frac{\hat{\mathbf{Y}}_{v1}}{\sum_{l=1}^{m+1} \hat{\mathbf{Y}}_{vl}}. \tag{3}$$

Finally, we sort all $v \in V$ according to the *fraud score* in descending order. Then, users working in collusion with the known fraudsters are expected to be ranked at the top of the output ordered list. In Eq. 3, the score associated with the *dummy* label is a part of the denominator. Another alternative of the *fraud score* can be derived without the contribution of the *dummy* label. It is defined as

$$\bar{\varphi}(v, \hat{\mathbf{Y}}) = \frac{\hat{\mathbf{Y}}_{v1}}{\sum_{l=1}^{m} \hat{\mathbf{Y}}_{vl}}. \tag{4}$$

In Section 4.3, we compare the performance of $\varphi(.)$ and $\bar{\varphi}(.)$. We discuss the results of this comparison in Section 5.

3.4 2-STEP Model

We now present an extension of the approach described in the previous section. We give another alternative definition of *fraud score*. We found that fraudsters tend to have many heavy links to their neighbors. Let us define the *weighted degree centrality* of v as the sum of edge weights originating from v, as

$$k_w(v) = \sum_{u \in N(v)} \mathbf{W}_{uv}, \tag{5}$$

where $\mathbf{W}_{uv}$ is weight of the edge (u, v), and $N(v)$ is the set of neighbors of v. The $k_w(v)$ is high if v has a heavy link(s). Figure 3 shows the fraction of nodes having weighted degree centrality k_w. It can be observed that fraudsters have a higher probability $p(k_w)$ than legitimate users when weighted degree centrality is high ($k_w > 20$). In contrast, fraudsters have lower probability when the centrality is low. Fraudulent users in the same group always interact together in order to inflate the auction or reputation. Many heavy links appear between them. In general, few popular legitimate users gain a great deal of attraction. Therefore, there is a higher tendency for fraudsters to have high weighted degree centrality, as shown in Figure 3.

MAD uses information from edge weight $\mathbf{W}_{uv}$ in the second term of Eq. 1. However, it does not use the weighted degree centrality. Suppose there is a fraudulent bidder, b, who has one heavy link to a fraudulent seller, s. Some legitimate

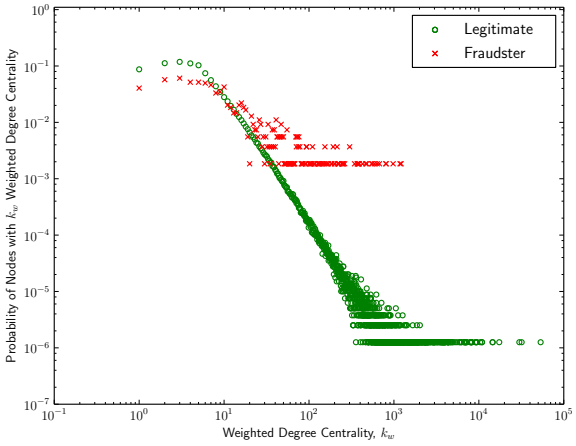


Fig. 3: Weighted degree centrality (k_w) distribution of legitimate users and known fraudsters

users connect with the fraudulent seller s. In this case, the confidence of the *fraud* label of s, $\hat{\mathbf{Y}}_{s1}$, decreases because the legitimate information propagates to s. Because s, who is the only neighbor of b, is interfered by legitimate users, then b's score of the *fraud* label, $\hat{\mathbf{Y}}_{b1}$, is low as well. Therefore, this fraudulent bidder will not be ranked at the top of the ranking results. If we incorporate the observed behavior that users who have heavy edge(s) tends to be fraudsters, the tendency of being fraudsters of b increases.

Therefore, we provide another definition of *fraud score*. This definition combines the result from MAD and the *weighted degree centrality*. We modify the aforementioned definition of the *weighted degree centrality* by penalizing each edge weighted with the $\varphi(.)$ of the neighbor. This definition of *fraud score* can be mathematically defined as

$$\rho(v, \hat{\mathbf{Y}}) = \varphi(v, \hat{\mathbf{Y}}) + \frac{\gamma}{|N(v)|} \sum_{u \in N(v)} \mathbf{W}_{uv}\, \varphi(u, \hat{\mathbf{Y}}), \tag{6}$$

where γ is a parameter, $\varphi(.)$ is defined in Eq. 3, W_{uv} is the weight of the edge (u, v), and $N(v)$ is the set of neighbors of v. The $\varphi(.)$ can be replaced with $\bar{\varphi}(.)$. We call this extension *2-STEP* model.

4 Experiments

We evaluated the performance of the models described in the previous sections. This section describes the evaluation methodology and reports the results.

4.1 Evaluation Metric

The output of the proposed approach is an ordered list of nodes. Thus, the normalized discounted cumulative gain (NDCG) [7] — a well-known evaluation metric for the information retrieval task — is used as an evaluation metric. It is defined as

$$\begin{aligned} \mathrm{NDCG} &= \frac{\mathrm{DCG}}{\mathrm{IDCG}}, \\ \mathrm{DCG} &= \sum_{i=1}^{p} \frac{2^{\mathrm{r}(i)} - 1}{\log_2(i+1)}, \\ \mathrm{IDCG} &= \sum_{i=1}^{\min(p,|Q|)} \frac{1}{\log_2(i+1)}, \end{aligned} \tag{7}$$

where p is the maximum number of nodes that are considered, $|Q|$ is the number of actual fraudsters in testing data, and $\mathrm{r}(i)$ is the relevance value of the i^{th} node. In this work, the relevance value is binary, $\mathrm{r}(i) \in \{0, 1\}$. $\mathrm{r}(i)$ is 1 if and only if the i^{th} node of the output list is fraudulent. NDCG ranges from 0.0 to 1.0. NDCG assumes that it is less useful for users when a relevant instance is ranked at a lower position of the result. Thus, NDCG penalizes relevant instances logarithmically proportional to the position of the instance. A higher NDCG indicates much better performance.

4.2 Methodology

The experiments were conducted on real world data acquired from YAHUOKU!, as mentioned in Section 2. We performed 5-fold cross validation in all experiments. Known fraudsters, sellers, bidders, mixed are distributed among 5 partitions, so that the distribution of user types in each fold resembles the whole dataset. The total number of fraudsters is far less than the total number of legitimate users. In each iteration, one chunk of the known fraudsters list was treated as the testing set, Q, and the remaining were used as training set. All stores were treated as a training set in every iteration.

The straightforward manner to evaluate performance is calculating NDCG on all test nodes. As previously mentioned in Section 2.2, auction users can be categorized into three groups — *bidder*, *seller*, and *mixed*. The degree distribution in Figure 2 shows that they tend to have different behaviors. We would like to gain more insight into the performance of detecting different types of fraudsters. After a model has assigned a *fraud score* to nodes, we rank the nodes in a specific category only. The remaining categories are ignored. It should be noted that the training set, or seed nodes, still contains every type of nodes. We annotate the caption *all* to the results obtained from evaluating all types of nodes. We annotate the caption *bidder*, *seller*, or *mixed*, if the results were measured on a specific user type.

4.3 Results

We set $\alpha = \beta$ for MAD and the 2-STEP model. We investigated the effect of different α. We tried $\alpha \in \{0.2, 0.4, 0.6, 0.8, 1.0\}$ and found that there was no statistically significant difference in NDCG. We then used $\alpha = 0.4$ which gave the lowest variance in the parameter tuning experiment. We set $\mu_1 = 1$, $\mu_2 = 0.01$, and $\mu_3 = 0.01$. For the 2-STEP model, we set $\gamma = 1$.

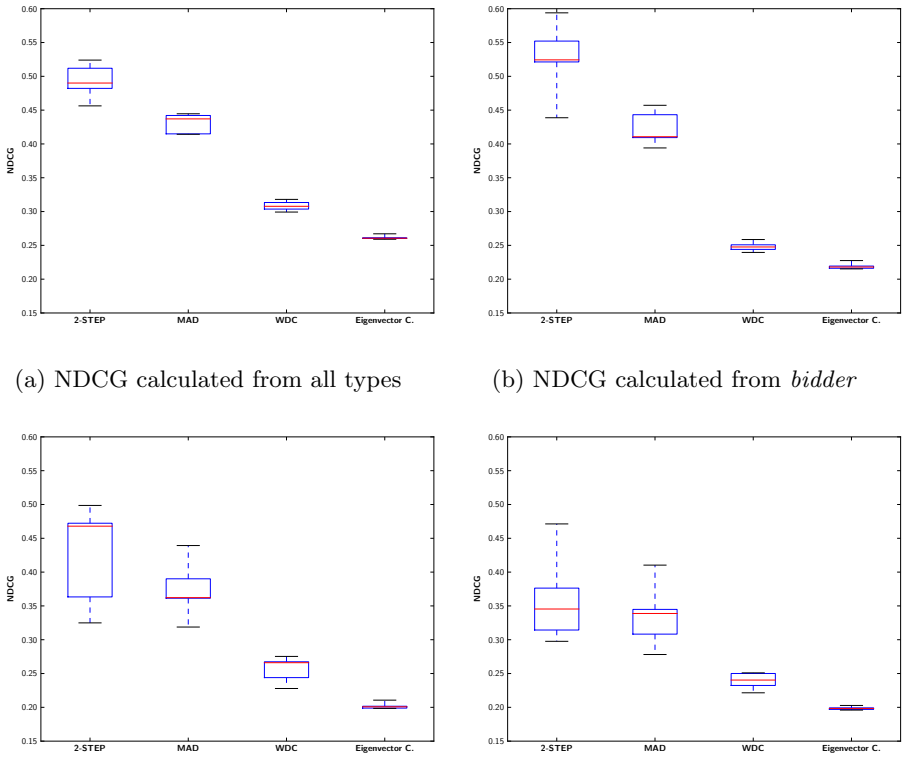
Table 1: Comparison of $\varphi(.)$ and $\bar{\varphi}(.)$ on all and separate node types ($\langle$NDCG$\rangle$ = mean of NDCG and SD = standard deviation)

Node Type	$\varphi(.)$		$\bar{\varphi}(.)$		p-value
	$\langle$NDCG$\rangle$	SD	$\langle$NDCG$\rangle$	SD	
All	0.431	0.015	0.406	0.019	0.002
Bidder	0.423	0.026	0.397	0.035	0.008
Seller	0.336	0.049	0.284	0.029	0.007
Mixed	0.374	0.044	0.319	0.024	0.006

In Section 3.3, we gave two alternative definitions of *fraud score*. One definition is based on information from the *dummy* label and the other is not. We conducted a two-tailed paired t-test to compare the average NDCG obtained from $\varphi(.)$ and $\bar{\varphi}(.)$. Table 1 summarizes the comparison of $\varphi(.)$ and $\bar{\varphi}(.)$ on all and individual types of fraudulent users. The result indicates that the *dummy* label has a significant advantage. From now on, we used $\varphi(.)$ as the main *fraud score* for MAD and 2-STEP.

We now compare our 2-STEP model (Section 3.4) with MAD (Section 3.3), weighted degree centrality (Eq. 5), and eigenvector centrality. In this experiment, we measured NDCG on the whole output ordered list, $p = |V_u|$. The eigenvector centrality is a well-known centrality measure defined as the principal eigenvector of a graph's adjacency matrix. PageRank is one of its variants [5]. The two centrality-based model are unsupervised. The results are summarized in Figure 4a. The centrality-based methods, weighted degree centrality and eigenvector centrality, could not precisely spot fraudsters. Our 2-STEP model outperformed MAD obviously, with 0.490 over 0.437 NDCG on average. Furthermore, Figure 4b, 4c, and 4d show the NDCG in ranking fraudulent *bidder*, *sellers*, and *mixed* separately. The results follow the same pattern as the previous result. These results imply that our extension of MAD outperformed the other models for every kind of user.

Figure 5 compares the NDCG of our 2-STEP model, MAD, and a semi-supervised model (CD). The results from the two centrality-based methods are not included in the figure because it is obvious that they cannot perform well in this online auction fraud detection. We used Viswanath et al. [17] as the baseline in this experiment. The baseline system uses a local community detection schema, Mislove's algorithm [10], to detect Sybil as mentioned in Section 6. In this experiment, we set $p \in \{100, 500\}$. We did not calculate the NDCG on the



(a) NDCG calculated from all types

(b) NDCG calculated from *bidder*

(c) NDCG calculated from *mixed*

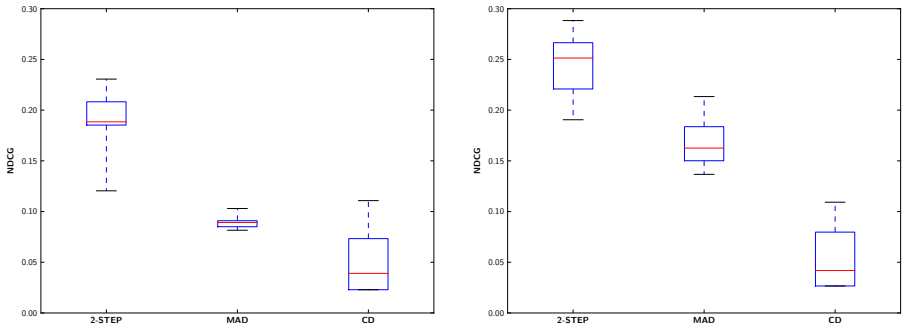
(d) NDCG calculated from *seller*

Fig. 4: NDCG of 2-STEP, MAD, weighted degree centrality (WDC), and eigenvector centrality where we set p equals the whole length of output list (box = 25^{th} and 75^{th} percentiles; central red line = median; and bar = min and max values)

whole output since the software implementation of the baseline system we used did not provide sorted results of the whole dataset. The result conforms the previous experiments that the 2-STEP model clearly outperformed MAD and the baseline.

5 Discussion

It is apparent from Table 1 that the *dummy* label has a significant advantage. As mentioned in Section 3.2, the score of *dummy* label is directly proportional to the entropy of weights of edges originating from the node. This implies that a user who uniformly interacts with others tends to be legitimate. The experimental results confirm that incorporating the entropy of edge weights can significantly



(a) NDCG where $p = 100$ (b) NDCG where $p = 500$

Fig. 5: NDCG of 2-STEP, MAD, and community detection-based model [10]

improve the system. This result is consistent with the previous computational linguistics study [15]. Our result verifies that the effect determined in computational linguistics even appears in online auction fraud detection.

According to Figure 4, the 2-STEP and MAD models exhibit low NDCG in ranking *seller*. The propagation method is used to satisfy the three requirements discussed in Section 3.2. One of them is *smoothness* — two adjacent nodes should be assigned similar labels. Suppose there is a fraudulent seller v who can successfully fool many legitimate users, the information from the legitimate users propagate to the fraudulent seller. In this case, the seller's *fraud* score, $\hat{\mathbf{Y}}_{v1}$, decreases relative to the *legitimate* score, $\hat{\mathbf{Y}}_{v2}$. We hypothesize that this is the primary reason fraudulent sellers are more difficult to detect. An effective method to solve this problem should be designed for future work.

We are also sure that the method is easily combined with the approach using user's information because it only use basic information of transaction. In future work, we will implement this hybrid approach in real service. The hybrid approach is expected to improve the effectiveness of the detection system in terms of not only on precision and resistance against the wrong user information.

6 Related Work

This section surveys related work focused on detecting online auction fraud. Over the last decade, online auction fraud has become one of the most serious Internet crimes; therefore, it has attracted much attention from many researchers. One of the first attempts was presented by Shah et al. [13]. They analyzed bidding strategies on eBay and revealed normal characteristics of online auction fraudsters. An association analysis was adopted to find cases of likely shilling behavior. However, they did not propose any systematic schema that can be used in large scale system. Rubin et al. [12] proposed statistical models based on observed

fraudulent behaviors. Their statistical bidder profiles are based on suspicious patterns in which shilling bidders are strongly associated with sellers, and shills rarely win auctions. Recently, supervised machine learning techniques have been used. Tsang et al. [16] used the C4.5 algorithm in WEKA to detect fraudulent bidders based on their bidding history. In general, these two research tried to propose a set of rules to detect fraudsters. However, sophisticated fraudsters usually have very flexible, adaptive, and various strategies. Therefore, it would be difficult to detect fraudsters via generalized rules — as the Vapnik's principle that when trying to solve some problem, one should not solve a more difficult problem as an intermediate step [3]. Yoshida and Ohwada [19] used a one-class support vector machine (SVM) and a decision tree to learn bidding attributes based on bidding history and user's evaluation results. In real world-wide-web situations, fraudsters can easily control and lie in their profile and rating. If a fraud detection system deeply rely on the user's inputs, these miss-leading information would easily defeat the precision of the system.

Many recent attention has focused on graph-based approach since objects in graph have long-range correlations [2]. Markov random field modeling was used to solve this problem [4, 11] in which belief propagation was used to detect near bipartite cores in an undirected graph, which was expected to be an abnormal pattern. However, fraudsters in our dataset rarely form the near bipartite core structure, then this schema could not be effectively used in our context. In 2012, Shi-Jen et al. [8] adapted PageRank and k-core clustering algorithm to detect collusive groups in online auction. As shown in Section 4.3, the eigenvector centrality, which PageRank is derived from, exhibited unpleasant results in our dataset.

Online auction fraud detection can be recognized as a member of anomaly detection problem. A large body of literature has investigated the potential of graph-based anomaly detection algorithm. Akoglu et al. discovered several rules in graph-based features from 1-step neighborhood around a node [1]. Therefore, fraud or anomaly score of a node depends on only 1-step neighbors. In contrast, our information propagation-based models can better use long-range correlations of objects in a graph. In 2010, Viswanath et al. demonstrated that community detection algorithm can be utilized to avoid multiple identity, or Sybil, attacks [17]. Sybil are malicious attackers who create multiple identities and influence the working of systems that rely upon open membership such as collaborative content rating and recommendation system. It is noticeable that Sybil and our focused fraudsters share a common behavior that they are groups of identities aiming for committing unacceptable activities. The scheme works by detecting local communities around trusted nodes because Sybil nodes tend to poorly connected to the rest of network. In another word, they assume that the *homophily* behavior tends to happen. In addition, the *homophily* behavior was employed in the area of social security [18] and accounting fraud detection [9] as well. Please refer to Akoglu et al.'s survey [2] for more extensive review about graph-based anomaly detection.

As described in Section 1, *homophily* behavior tends to occur in online auction fraud networks. There is a very high tendency that online auction fraudsters have connections together. Even if, the behavior has been widely used to solve many fraud detection problems, to the best of our knowledge, there is no publication that focused on such behavior for online auction fraud detection. The homophily behavior is analogous with one of the main principle concepts of graph-based SSL models. Therefore, this work proposed a graph-based SSL model for online auction fraud detection.

7 Conclusion

We proposed an online auction fraud detection approach involving the extension of a graph-based semi-supervised learning model. The development of our approach was motivated by the *homophily* behavior of fraudsters. We extended the *modified adsorption* model to propagate information from a small set of known fraudsters to the entire graph. We found that fraudsters tend to have many heavy interactions with neighbors. We integrated this suspicious social behavior into this extended model, which we call the 2-STEP model. The experiments, on real-world data, suggest that our approach significantly improves accuracy.

8 Acknowledgments

We would like to thank the Yahoo! JAPAN patrol team for their assistance and support, especially Hiroyuki Kobayashi and Yuichi Nakatsu for regarding their data preparation.

References

1. Akoglu, L., McGlohon, M., Faloutsos, C.: Oddball: Spotting anomalies in weighted graphs. In: Proceedings of the 14th Pacific-Asia Conference on Advances in Knowledge Discovery and Data Mining - Volume Part II. pp. 410–421. PAKDD'10, Springer-Verlag, Berlin, Heidelberg (2010)
2. Akoglu, L., Tong, H., Koutra, D.: Graph based anomaly detection and description: a survey. Data Mining and Knowledge Discovery pp. 1–63 (2014)
3. Chapelle, O., Schölkopf, B., Zien, A.: Semi-supervised learning. MIT, Cambridge, Mass. London (2010)
4. Chau, D.H., Pandit, S., Faloutsos, C.: Detecting fraudulent personalities in networks of online auctioneers. In: Proceedings of the 10th European Conference on Principle and Practice of Knowledge Discovery in Databases. pp. 103–114. PKDD'06, Springer-Verlag, Berlin, Heidelberg (2006)
5. Chau, D.H.P., Nachenberg, C., Wilhelm, J., Wright, A., Faloutsos, C.: Polonium: Tera-scale graph mining and inference for malware detection. In: SIAM International Conference on Data Mining (SDM). pp. 131–142 (2011)
6. Dong, F., Shatz, S.M., Xu, H.: Combating online in-auction fraud: Clues, techniques and challenges. Computer Science Review 3(4), 245–258 (2009)

7. Järvelin, K., Kekäläinen, J.: IR evaluation methods for retrieving highly relevant documents. In: Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 41–48. SIGIR '00, ACM, New York, NY, USA (2000)
8. Lin, S.J., Jheng, Y.Y., Yu, C.H.: Combining ranking concept and social network analysis to detect collusive groups in online auctions. Expert Systems with Applications 39(10), 9079–9086 (2012)
9. McGlohon, M., Bay, S., Anderle, M.G., Steier, D.M., Faloutsos, C.: Snare: A link analytic system for graph labeling and risk detection. In: Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 1265–1274. KDD '09, ACM, New York, NY, USA (2009)
10. Mislove, A., Viswanath, B., Gummadi, K.P., Druschel, P.: You are who you know: Inferring user profiles in online social networks. In: Proceedings of the 3rd ACM International Conference on Web Search and Data Mining. pp. 251–260. WSDM '10, ACM, New York, NY, USA (2010)
11. Pandit, S., Chau, D.H., Wang, S., Faloutsos, C.: Netprobe: A fast and scalable system for fraud detection in online auction networks. In: Proceedings of the 16th International Conference on World Wide Web. pp. 201–210. WWW '07, ACM, New York, NY, USA (2007)
12. Rubin, S., Christodorescu, M., Ganapathy, V., Giffin, J.T., Kruger, L., Wang, H., Kidd, N.: An auctioning reputation system based on anomaly. In: Proceedings of the 12th ACM Conference on Computer and Communications Security. pp. 270–279. CCS '05, ACM, New York, NY, USA (2005)
13. Shah, H., Joshi, N., Sureka, A., Wurman, P.: Mining ebay: Bidding strategies and shill detection. In: Zaïane, O., Srivastava, J., Spiliopoulou, M., Masand, B. (eds.) WEBKDD 2002 - Mining Web Data for Discovering Usage Patterns and Profiles SE - 2, Lecture Notes in Computer Science, vol. 2703, pp. 17–34. Springer Berlin Heidelberg (2003)
14. Talukdar, P.P., Crammer, K.: New regularized algorithms for transductive learning. In: Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases: Part II. pp. 442–457. ECML PKDD '09, Springer-Verlag, Berlin, Heidelberg (2009)
15. Talukdar, P.P., Pereira, F.: Experiments in graph-based semi-supervised learning methods for class-instance acquisition. In: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. pp. 1473–1481. ACL '10, Association for Computational Linguistics (2010)
16. Tsang, S., Koh, Y.S., Dobbie, G., Alam, S.: Detecting online auction shilling frauds using supervised learning. Expert Systems with Applications 41(6), 3027–3040 (2014)
17. Viswanath, B., Post, A., Gummadi, K.P., Mislove, A.: An analysis of social network-based sybil defenses. SIGCOMM Comput. Commun. Rev. 40(4), 363–374 (2010)
18. Vlasselaer, V.V., Akoglu, L., Eliassi-Rad, T., Snoeck, M., Baesens, B.: Guilt-by-constellation: Fraud detection by suspicious clique memberships. In: System Sciences (HICSS), 2015 48th Hawaii International Conference on. pp. 918–927 (2015)
19. Yoshida, T., Ohwada, H.: Shill bidder detection for online auctions. In: Zhang, B.T., Orgun, M. (eds.) PRICAI 2010: Trends in Artificial Intelligence SE - 33, Lecture Notes in Computer Science, vol. 6230, pp. 351–358. Springer Berlin Heidelberg (2010)