Expert Systems with Applications 41 (2014) 3027–3040

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Detecting online auction shilling frauds using supervised learning

Sidney Tsang *, Yun Sing Koh, Gillian Dobbie, Shafiq Alam

Department of Computer Science, The University of Auckland, New Zealand

ARTICLE INFO

Keywords:
Supervised fraud detection
Online auction fraud
Agent-based simulation

ABSTRACT

Online auction sites are a target for fraud due to their anonymity, number of potential targets and low likelihood of identification. Researchers have developed methods for identifying fraud. However, these methods must be individually tailored for each type of fraud, since each differs in the characteristics important for their identification. Using supervised learning methods, it is possible to produce classifiers for specific types of fraud by providing a dataset where instances with behaviours of interest are assigned to a separate class. However this requires multiple labelled datasets: one for each fraud type of interest. It is difficult to use real-world datasets for this purpose since they are difficult to label, often limited in size, and contain zero or multiple suspicious behaviours that may or may not be under investigation.

The aims of this work are to: (1) demonstrate the approach of using supervised learning together with a validated synthetic data generator to create fraud detection models that are experimentally more accurate than existing methods and that is effective over real data, and (2) to evaluate a set of features for use in general fraud detection is shown to further improve the performance of the created detection models.

The approach is as follows: the data generator is an agent-based simulation modelled on users in commercial online auction data. The simulation is extended using fraud agents which model a known type of online auction fraud called competitive shilling. These agents are added to the simulation to produce the synthetic datasets. Features extracted from this data are used as training data for supervised learning. Using this approach, we optimise an existing fraud detection algorithm, and produce classifiers capable of detecting shilling fraud.

Experimental results with synthetic data show the new models have significant improvements in detection accuracy. Results with commercial data show the models identify users with suspicious behaviour.

1. Introduction

Online auction sites such as (eBay) and (TradeMe) allow goods and services to be bought and sold online anonymously. The most common type of online auction is the English auction (Menezes & Monteiro, 2005), where bids are placed in ascending order, are publicly observable, and the winner is the final bidder with the highest bid. In 2011, there were 90 million active users in eBay (Shen & Sundaresan, 2011), with more than 170 million concurrent auctions (Auction Count Charts).

The anonymity and simplicity of creating multiple aliases allows unsuspecting users to be exploited by dishonest users. This exploitation can take many forms, including shilling, non-delivery, misrepresentation, or by the sale of stolen goods (Dong, Shatz, & Xu, 2009). Dishonest users will also disguise themselves to avoid detection by imitating normal behaviours (Chang & Chang, 2011), making fraudulent behaviour difficult to define. Previous work has noted that legitimate users often appear to behave irrationally (Mizuta & Steiglitz, 2000), and previous attempts at clustering users into predefined types according to their bidding behaviour have failed to label the majority of users (Shah, Joshi, Sureka, & Wurman, 2003). The range of potential fraudulent behaviour together with the number and range of legitimate behaviours makes it difficult to differentiate between fraudulent and legitimate users. The class imbalance in auction data, where the number of legitimate actions outnumber the fraudulent, makes the accurate classification of users as fraudulent or legitimate non-trivial.

Past research in online auction fraud has focused on detecting specific fraudulent behaviours using a range of techniques, including decision trees (Chang & Chang, 2011; Almendra, 2013), clustering (Chang & Chang, 2010), regression models (Kauffman & Wood, 2003; Chae, Shim, Cho, & Lee, 2007), statistical methods (Trevathan & Read, 2007a), model checking (Xu & Cheng, 2007; Xu, Bates, & Shatz, 2009), and graph mining methods (Pandit, Chau, Wang, & Faloutsos, 2007). The general approach in these works involve identifying the type of behaviour

* Corresponding author. Tel.: +64 210537428.
E-mail addresses: stsa027@aucklanduni.ac.nz (S. Tsang), ykoh@cs.auckland.ac.nz (Y.S. Koh), gill@cs.auckland.ac.nz (G. Dobbie), sala038@aucklanduni.ac.nz (S. Alam).

0957-4174/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2013.10.033

or fraud of interest, then selecting a set of features that are hypothesised to be able to differentiate between users with normal and suspicious behaviour. A fraud detection algorithm is then developed using the selected feature set. The algorithm is evaluated using commercial auction datasets without knowledge of ground truth, or by using synthetic datasets without guarantee of its similarity with real data. Both types of datasets reduce the reliability of any conclusions drawn about method accuracy or effectiveness (Tsang, Dobbie, & Koh, 2012a).

In this work, synthetic data is used together with supervised learning methods to develop classification models for fraud detection. Supervised learning methods allow classifiers, which can detect different types of frauds or behaviours of interest, to be trained given an appropriate training set. The synthetic data used in this work is generated using a validated agent-based simulation (Tsang, Dobbie, & Koh, 2012b,chap. 11), which has been extended to generate data containing specific fraudulent behaviours. The type of fraudulent behaviours added determines the types of frauds the resulting model can detect. An appropriate training set for supervised learning methods is created in three steps: first, define an agent-type that represents the fraud type of interest; second, generate synthetic data using the defined agent; and third, transform the generated synthetic data, which is a sequence of auctions and bids, into values for a set of user-defined features. This transformed synthetic dataset is then used as a training set for the selected supervised learning technique. This approach allows models for detecting specific types of fraud to be constructed more easily than in previous work, and with improved accuracy, as shown by the experimental results in ection 4. To our knowledge, no previous work has combined the use of a data generator and supervised learning methods to develop fraud detection methods.

This work focuses on a type of fraud called *competitive shilling*. Competitive shilling occurs when a user submits bids to a collaborating seller's auction to elevate the final auction price, without the intention of winning. The legitimate bidder is cheated by paying more than they otherwise would when winning the item. For example, suppose there were only two bidders in an auction: one legitimate (L), and one fraudulent (F), with bidding proceeding like so: L: \$10, F:\$11, L:\$12, F:\$13; L:\$14. If there are no additional bids, L, the legitimate bidder, pays an additional \$4 due to bids by F.

1.1. Contribution

Our contributions in this paper are as follow:

- We propose an approach for generating classification models for detecting suspicious behaviours in commercial auction data. The approach uses a synthetic data generator with supervised learning techniques. To demonstrate this approach, we define two types of fraudulent behaviours, and develop two corresponding classification models for detecting those behaviours. We show that these models, created using synthetic data, can also be applied to commercial online auction datasets to identify suspicious users. We also use supervised learning techniques to improve the performance of an existing fraud detection algorithm.
- We describe and define a set of user features for fraud detection. The set of features captures many aspects of bidding behaviour, and may be useful for developing models for the detection of different types of auction fraud.
- We present experimental results on both validated synthetic datasets and commercial datasets. Results on synthetic datasets show that our supervised approach produces classification models of greater accuracy than existing algorithms, and that this improvement is further increased by the use of our proposed feature set. Evaluation results on commercial datasets using these same models show that they are able to identify users that exhibit suspicious behaviour.

1.2. Overview

The remainder of the paper is organised as follows: helpful information about auction mechanics and data generation, auction fraud, and a shill detection algorithm is given in Section 2; the methods used to improve and develop classifiers for identifying shilling fraud is given in Section 3. Section 4 presents the evaluation procedure and results on synthetic data; and Section 5 is a case study applying the classifiers to commercial auction data. Section 6 presents related work. Finally, Section 7 concludes the paper.

2. Background

This section provides information that may be helpful in understanding subsequent sections. Section 2.1 defines a simplified auction model for the auctions discussed in this paper. Section 2.2 describes the generation and quality of synthetic data used in this paper. Section 2.3 describes the characteristics of shilling fraud. Sections 2.4 and 2.5 describes the behaviour of two types of shilling agents. Section 2.6 briefly describes the Shill Score algorithm proposed by Trevathan and Read (2007a).

2.1. Auction model

In this section we present a formal model for English auctions, which consists of a set of users $\mathcal{U}$, a set of auctions $\mathcal{A}$, a set of bids $\mathcal{B}$, and a set of rules that govern the relationship between them. The notation used in this section is also used to describe the features defined in Section 3.3 and Appendix A.1.

2.1.1. User model

A user can list auctions, and participate in auctions by bidding. For user $u \in \mathcal{U}$, let $\mathcal{S}^u = \{m | m \in \mathcal{A},\ m \text{ is listed by } u\}$, where $\mathcal{S}^u$ is the set of auctions listed by u; let $\mathcal{P}^u = \{m | m \in \mathcal{A},\ u \text{ bids in } m\}$, where $\mathcal{P}^u$ is the set of auctions u participated in; and let $\mathcal{G}^u$ represent a set of attributes associated with user u, such as account age and reputation score.

2.1.2. Auction and bid model

An auction consists of no bids, one bid, or multiple bids. For auction $a \in \mathcal{A}$, let $\mathcal{B}^a$ be the complete set of bids submitted to a, and let b^a_c, where $1 \leqslant c \leqslant |\mathcal{B}^a|$, be the cth bid submitted to a.

Each auction has a set of attributes. The set of attributes includes reserve price (RS), start time (ST), end time (ET) and duration (D). For $a \in \mathcal{A}$, let $\mathcal{W}^a = \{w^a_{RS}, w^a_{ST}, w^a_{ET}, w^a_D\}$ be the set of auction attributes for auction a.

For auction $a \in \mathcal{A}$ and user $u \in \mathcal{U}$, let $\mathcal{B}^{a,u}$ be the complete set of bids submitted to a by u, and let $b^{a,u}_c$, $1 \leqslant c \leqslant |\mathcal{B}^a|$ be the cth bid submitted to a by u.

2.1.3. Model of auction rules

In online auctions, bid values and times increase monotonically. Given b^a_c (the cth bid in auction a), we define t^a_c as the time since the auction start time (w^a_{ST}) that b^a_c was submitted, and let $\mathcal{T}^a = \{t^a_c | 1 \leqslant c \leqslant |\mathcal{B}^a|\}$ be the set of bid submission times for auction a. Correspondingly, given $b^{a,u}_c$ (the cth bid by user u in auction a), let $t^{a,u}_c$ be the time since the auction's start that $b^{a,u}_c$ was submitted. Let $\mathcal{T}^{a,u} = \{t^{a,u}_c | 1 \leqslant c \leqslant |\mathcal{B}^a|\}$ be the set of submission times for bids by user u in auction a.