Electronic Commerce Research and Applications 11 (2012) 346–360

Contents lists available at SciVerse ScienceDirect

Electronic Commerce Research and Applications

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

An effective early fraud detection method for online auctions

Wen-Hsi Chang [a,*], Jau-Shien Chang [b]

[a] Graduate Institute of Management Sciences, TamKang University, Taipei County, Taiwan 25137, Republic of China
[b] Department of Information Management, TamKang University, Taipei County, Taiwan 25137, Republic of China

ARTICLE INFO

Article history:
Received 13 March 2011
Received in revised form 22 February 2012
Accepted 22 February 2012
Available online 1 March 2012

Keywords:
Classification
Electronic auctions
Electronic commerce
Fraud detection

ABSTRACT

While online auctions continue to increase, so does the incidence of online auction fraud. To avoid discovery, fraudsters often disguise themselves as honest members by imitating normal trading behaviors. Therefore, maintaining vigilance is not sufficient to prevent fraud. Participants in online auctions need a more proactive approach to protect their profits, such as an early fraud detection system. In practice, both accuracy and timeliness are equally important when designing an effective detection system. An instant but incorrect message to the users is not acceptable. However, a lengthy detection procedure is also unsatisfactory in assisting traders to place timely bids. The detection result would be more helpful if it can report potential fraudsters as early as possible. This study proposes a new early fraud detection method that considers accuracy and timeliness simultaneously. To determine the most appropriate attributes that distinguish between normal traders and fraudsters, a modified wrapper procedure is developed to select a subset of attributes from a large candidate attribute pool. Using these attributes, a *complement phased modeling procedure* is then proposed to extract the features of the latest part of traders' transaction histories, reducing the time and resources needed for modeling and data collection. An early fraud detection model can be obtained by constructing decision trees or by instance-based learning. Our experimental results show that the performance of the selected attributes is superior to other attribute sets, while the hybrid complement phased models markedly improve the accuracy of fraud detection.

1. Introduction

The Internet has changed the way the people interact with each other. Family and friends can instantly and conveniently get in touch with each other in ways that were unimaginable only a few decades earlier. Such speed and convenience also opened e-commerce to both businesses and individuals around the world. Nowhere has this been more obvious or lucrative than in the case of online auctions, where millions of transactions can occur in the blink of an eye. Physical goods, as well as service packages, are traded as online commodities without the limitations of time and physical location. Because the Internet offers many opportunities to interact with strangers (Resnick et al. 2000), this anonymity combined with the convenience of the Internet allows the online auction to become prosperous. For example, eBay—the largest worldwide auction site—posted US$9.2 billion in revenue and US$1.8 billion net income for 2010 (eBay 2010), while the total revenue of Taiwan's online auction market reached NT$15.3 million (III 2010).

Unfortunately, such vast profits also attract the attention of criminals who use fraud to cash in on the lucrative online trading market. According to annual reports of the Internet Complaint Center, online auction fraud ranks as one of the top two serious Internet crimes in recent years, contributing to a risky situation for online auction participants (NW3C 2010). Most online auction houses realize that fraud corrodes not only their trustworthiness but also the prosperity of the entire market. For instance, multiple online identities are easy to create: a fraudster could use his many accounts to execute sophisticated schemes, while camouflaging his malicious intent and evading traditional detection methods that merely examine individual identities (Chau 2011). As more and more inexperienced traders become targeted victims, they begin to distrust the market, resulting in fewer buyers and fewer sellers (Gavish and Tucci 2008). To help promote trust in the online market, auction houses developed reputation systems to assist users in evaluating potential trading partners.

Reputation systems help buyers decide whether to purchase a product based on a feedback score. After each trade, both the seller and the buyer can leave ratings and feedback comments on the other party. Over time, these comments and feedback accumulate in the trader's transaction history, of which the feedback score is one part. This kind of reputation system is simple and easy to understand. It uses +1, 0, and −1 to denote the level of satisfaction for a trade. However, this kind of scoring mechanism has some drawbacks (Rubin et al. 2005, Buchegger and Boudec 2003). For

* Corresponding author.
E-mail address: wenhsi.chang@gmail.com (W.H. Chang).

1567-4223/$ - see front matter
doi:10.1016/j.elerap.2012.02.005

instance, the level of satisfaction cannot express participants' precise thoughts or insights. In addition, a buyer may hesitate to give negative feedback to the seller to avoid receiving vengeful feedback in return. While reputation systems provide a certain degree of protection, they are not enough to protect traders from fraudulent schemes. Most online auction houses adopt passive approaches to the coordination of reputation systems and management policies that could address fraudulent schemes. However, if users had more proactive approaches, such as an automatic fraud detector, online trading could be safer.

From the perspective of crime prevention, the capability of early warning is indispensible for fraud detection (Burge et al. 1997, Dohono 2004). The detection procedure must not only identify fraud that has already occurred, but it must warn traders of potential fraudsters. The identification of a fraudster should not rely only on behavioral features that occur once the fraud has been activated. Even a less-experienced trader can distinguish between a normal trader and a fraudster if the fraudster's transaction history reveals more negative ratings. An early fraud detection system would provide a method that would alert users before a fraud is activated. Auction houses can both help and benefit from an effective early fraud detection system. In spite of possible misjudgment on suspicious accounts, the auction houses can mark potential fraudsters that are under surveillance as early as possible. As a result, the quality of services of auction houses will be improved by eliminating potential fraudulent events.

Prior research has proposed online auction fraud detection (Chau and Faloutsos 2005, Ku et al. 2007, Pandit et al. 2007, Zhang et al. 2008). The typical detection procedure used in the previous work consists of two fundamental steps: (1) a set of attributes is devised and their values are extracted from the transaction histories to distinguish between normal traders and fraudsters; and (2) a detection model based on these attributes is built by machine learning techniques, such as decision trees or instance-based learning. In terms of devising a set of attributes, for example, the speed of obtaining feedback scores is an effective attribute to identify a fraudster who builds his reputation rapidly with fake transactions. In general, the detection accuracy is strongly related to the effectiveness of the attribute set and the appropriateness of the modeling method. Further, the cost of detection is affected by how efficiently the attributes can be extracted from the transaction histories. Previous work in this area has provided some level of progress in this area; however, some problems still exist and need to be resolved.

- Latent fraudulent behavior is usually sophisticated and opaque, often including several fraudsters working together to scam buyers. For example, fraudsters can help each other by pretending to be buyers and leaving a lot of positive feedback to persuade other traders that the fraudster is actually a reliable trading partner.
- Using a larger number of measured attributes incurs more computation effort and is not necessarily helpful in raising fraud detection accuracy. In fact, the detection accuracy will be degraded if irrelevant attributes are incorporated into model construction. Additionally, even though some algorithms, such as the naïve-Bayes algorithm, are robust with respect to irrelevant attributes, the performance may degrade quickly if correlated attributes are added (Kohavi and John 1997). Applying the expectation-maximization (EM) algorithm might be helpful in fraud detection that calculates the fraudster cluster probabilities as expected class values and maximizes the likelihood of the distributions.
- Data retrieval for long-lived accounts is limited, and constructing a complete transaction history is impossible. Even if such retrieval were possible, analyzing such a huge amount of data would be too time-consuming. Therefore, it is practical to construct a more parsimonious detection model that does not require navigating numerous complex web pages to differentiate between legitimate users and fraudsters. Meanwhile, the detection accuracy can be similar to or better than those generated by complicated and costly steps.

To detect fraud effectively and efficiently in online auctions, this study aims to develop a detection method with higher accuracy but lower cost. First, we focus on devising a concise set of measured attributes for early fraud detection. For this purpose, a rich attribute set comprised of 44 elements is evaluated using a modified wrapper procedure. Based on the results, 10 of the 44 attributes are chosen. A *complement phased modeling procedure* is then proposed to further improve the detection accuracy and to reduce the cost of model construction. To test the effectiveness of the proposed procedure, 2475 real transaction records were collected from Yahoo! in Taiwan for analysis. The results demonstrate that the set of 10 selected attributes is superior to that of two larger attribute sets. This finding suggests that detection accuracy can be maintained or even improved by using fewer attributes in the model. Moreover, the hybrid complement phased modeling can materially improve the detection accuracy to more than 90%.

The rest of this paper includes a literature review followed by a section on the methods of measured attribute selection that is used in this study. Section 4 discusses how to apply complement phased modeling to construct early detection models with the selected measured attributes. Section 5 presents the experimental results, followed by the conclusion and suggestions for future work in the final section.

2. Literature

To facilitate further discussion, we next introduction concepts and techniques related to this study, including the literature on fraudulent schemes in existing reputation systems, attribute selection procedures, modeling algorithms, and early detection methods.

2.1. Fraud and its countermeasures in online auctions

The reputation systems implemented by most auction houses are too simple. After a transaction is completed, the traders can only be rated on a "negative", "neutral," and "positive" scale. Even dishonest sellers and incompetent buyers want to have positive feedback, so some of them try to increase their reputation with deceptive means (Maranzato et al. 2010a,b). Swindlers could attract buyers by fabricating transaction records that inflate their feedback scores in order to hide their malicious intent (Wang and Chiu 2005). Another common fraudulent trick is to first do several small transactions in order earn more positive rating score, but then cheat later on the first expensive commodity. Using multiple identities is also a common type of fraud in online auctions. A fraudster first creates multiple identities, dividing them into two groups (fraudsters and accomplices). Then, the fraudsters use the accomplices to artificially boost their reputations by leaving positive ratings (Pandit et al. 2007). In addition to the positive rating manipulation, the negative feedback can be biased by dishonest traders as well. Some buyers and sellers who deserve negative feedback ratings threaten their trading partners into leaving positive ratings, regardless of the actual experience.[1] Obviously, when

[1] This kind of feedback extortion has become a problem within eBay. In January 2008, eBay changed its feedback management policy so that buyers could leave negative feedback on the seller, but not the other way around (Goodrich and Kerschbaum 2011).