

A Run-Time Algorithm for Detecting Shill Bidding in Online Auctions

Nazia Majadi¹, Jarrod Trevathan² and Heather Gray³

¹ Griffith University, School of ICT, Brisbane, Australia, nazia.majadi@griffithuni.edu.au

² Griffith University, School of ICT, Brisbane, Australia, j.trevathan@griffith.edu.au

³ Griffith University, School of ICT, Brisbane, Australia, h.gray@griffith.edu.au

Received 11 July 2017; received in revised form 16 November 2017; accepted 27 December 2017

Abstract

Online auctions are a popular and convenient way to engage in ecommerce. However, the amount of auction fraud has increased with the rapid surge of users participating in online auctions. Shill bidding is the most prominent type of auction fraud where a seller submits bids to inflate the price of the item without the intention of winning. Mechanisms have been proposed to detect shill bidding once an auction has finished. However, if the shill bidder is not detected during the auction, an innocent bidder can potentially be cheated by the end of the auction. Therefore, it is essential to detect and verify shill bidding in a running auction and take necessary intervention steps accordingly. This paper proposes a run-time statistical algorithm, referred to as the Live Shill Score, for detecting shill bidding in online auctions and takes appropriate actions towards the suspected shill bidders (e.g., issue a warning message, suspend the auction, etc.). The Live Shill Score algorithm also uses a Post-Filtering Process to avoid misclassification of innocent bidders. Experimental results using both simulated and commercial auction data show that our proposed algorithm can potentially detect shill bidding attempts before an auction ends.

Keywords: Auction fraud, Bidding behaviour, Live shill score, Online auction, Post-filtering process, Shill bidding.

1 Introduction

An online auction is a process of selling or buying where a seller lists an item online for a set amount of time and buyers must place a bid higher than the last bid in order to win. Online auctions have made auctioning easily accessible to everyone from novice to advanced users. This approach removes the physical and logistical limitations of traditional auctions (e.g., difficulty of travelling to look for the required goods, time, presence, and small target audience) [21]. Therefore, online auctions are becoming extremely popular for sellers and buyers since 2002 (Site 1). In the first quarter of 2017, ebay alone has reached 169 million active users after surpassing 167 million active users at the end of 2016 (Site 3). Moreover, the number of online auction items hosted by eBay has increased from 266 million to more than 1 billion between September 2014 and October 2016 (Site 4).

However, despite the overwhelming benefits of online auctions, they are attractive to fraudsters. Auction fraud is one of the fastest growing forms of Internet-based crime. The U.S. Federal Bureau of Investigation's Internet Crime Complaint Center (IC3) reveals that approximately 42 thousand complaints of auction related fraud were received from June 1, 2014 to December 31, 2014 (Site 5). Auction related complaints are still ranking at the top of the complaint list with approximately 116,292 complaints in 2015 found in the latest report of IC3 (Site 6). As participants are anonymous, both sellers and bidders can participate in fraudulent activities for their own benefit. For example, the seller may misrepresent or not deliver an item. Likewise, a bidder can refuse to pay for the item, or have his/her bid forged. Furthermore, the auctioneer could block bids or influence the auction in a manner that maximizes its revenue.

Moreover, there are certain types of bidding behaviour that can be used to influence the auction in an undesirable manner. For instance, Trevathan and Read [25] identifies *Bid shielding*, which is a strategy performed by one or more bidders to preserve a low bid in an online auction. *Bid sniping* is a strategy employed by a bidder in which the majority of bids arrive in the final moments of the auction [25]. Trevathan and Read [25] also describe *Bid siphoning*, which is the situation where an outsider contacts bidders and offers to sell them an identical item, thus keeping bidders away from the legitimate seller's auction [25]. Moreover, they discuss *Shill Bidding* which is the topic of this paper.

Shill bidding is one of the most severe types of auction fraud where a seller enters bids on his/her own auction or engages friends to place bids on his/her behalf. The goal of shill bidding is to artificially raise the item's price so that legitimate bidders must pay more in order to win [25, 26]. Shill bidding is the hardest type of fraud to detect because it is generally performed by the seller of an auction colluding with one or more bidders in the auction. Alternatively, a seller might create multiple fake bidding accounts which they use to submit fake bids. The long term effect of shill bidding is that it hurts the online auction community because it causes buyers to pay more for their products [15]. Authorities have recognised shill bidding as harmful, making it illegal. As a result, several successful prosecutions have occurred where the perpetrators were convicted of shill bidding [25]. For example, three men were charged for participating in a group of shill bidding in hundreds of art auctions on eBay in March 2011 (Site 7). The men created more than 40 user IDs on eBay with false information and hosted more than 1,100 auctions from late 1998 until May 2000 (Site 7). This enabled them to place shill bids on more than half of those auctions (Site 7). Moreover, Trade Me, a New Zealand based auction site, fined \$42000 to a car trading company for allegedly engaging in shill bidding (Site 2).

Shill bidding detection is very difficult because online auction houses allow users to create multiple accounts quickly and easily [25]. They also allow for a high level of anonymity, many only requiring an email address to create an account. The relative ease of creating accounts together with the level of anonymity, make online auctions particularly susceptible to shill bidding. Additionally, it is hard to prove that someone is indeed guilty of shill bidding. Although online auctioneers claim to monitor their auctions for signs of shill bidding, they do not disclose how such monitoring occurs. Therefore, it is even harder to determine how to prevent shill bidders from engaging in such shill bidding activities. Thus, there is now growing momentum in the research community to investigate shill bidding behaviours; and propose detection and prevention mechanisms to combat this crime before the victim incurs any monetary loss.

Researchers from economics, ecommerce and computer science are working on methods for combating shill bidding [3], [12], [23]. Almost all the research proposals for shill bidding detection/prevention are conducted after an auction ends. Nonetheless, there are some limitations of this approach. Firstly, it becomes too late to react to shill bidders (i.e., the fraud occurs and the victim has already been cheated due to shill bidding). Secondly, if the auction system can successfully detect the presence of shill bidding as early as possible after it happens, the auction house can take the initiative (e.g., suspend the auction) so that shill bidding can be discouraged. In this instance, the honest bidders can be protected from losing money.

This paper proposes an algorithm for detecting shill bidders in a running auction. Our proposed algorithm splits an auction into four time periods and examines a bidder's bidding behaviours during each period. The algorithm calculates a *Live Shill Score (LSS)* for each bidder during each of these periods. The auctioneer then notifies legitimate bidders and/or takes appropriate actions against suspected shill bidders based on the *LSS*. However, the bidding behaviour of a legitimate bidder may match with shill bidding patterns. Therefore, some innocent bidders may be classified as shill bidders. We have introduced a *Post-Filtering Process (PPF)* in our *LSS* algorithm for exonerating the innocent bidders who are misclassified as shill bidders. We have implemented the run-time shill bidding detection algorithm and

also undertaken some preliminary tests on simulated data and commercial auction data (collected from eBay). Our experimental results show that our proposed *LSS* algorithm can potentially detect shill bidders during an auction. Furthermore, our experimental results also illustrate that a bidder who passes *PPF* can be exonerated.

This paper is organized as follows: Section 2 gives a background on run-time shill detection approaches, describes the problem motivation for our research, and analyses bidding behaviours of shill bidders; Section 3 presents our proposed algorithm for run-time detection of shill bidding; Section 4 discusses *LSS* and *PPF* algorithm in details and also explains the calculation procedure of *LSS* for each bidder during an auction. Section 5 shows our experimental setup and preliminary results. Finally Section 6 provides concluding remarks and avenues for future work.

2 Background Study and Problem Motivation

Shill bidding detection research proposals can be classified in various ways depending on their objectives (i.e., concurrent [30], statistical [26], machine learning [7], etc.). We are focusing on two distinct approaches:

- Offline (or static) - The shill bidding detection/prevention mechanism runs after an auction has terminated; or
- Real-time (or dynamic) - The shill bidding detection/prevention mechanism operates while an auction is in-progress.

The main distinction between these two types of shill bidding detection/prevention mechanisms is that real-time systems have the ability to take disciplinary actions during an auction when suspected shill bidders has been detected [15]. On the other hand, an offline mechanism will let the fraud occur, and then produce evidence that shill bidding has occurred [15]. Shill detection techniques provide a deterrent to future participants of an auction about bidders or sellers potentially involved in shill bidding [15].

Among the offline approaches, Trevathan and Read [26] present a *Shill Score* (SS) reputation system to determine the degree of shill bidding in online ascending auctions that terminate at a pre-determined time. They consider a series of auctions by a specific seller for identifying bidding patterns and calculate a *Shill Score* for each bidder [26]. The Shill Score is used to indicate the likelihood of the user being a potential shill bidder [26]. Bidders can observe other bidders' Shill Scores and decide whether they want to participate in the particular seller's auction [26]. However, their proposed shill bidding detection algorithm does not work in real-time. Later, Trevathan and Read [27] improved their previous work by proposing a method to detect colluding shill bidding where multiple bidders participate in shill bidding activities under a group. The authors introduce a score called *Collusion Score* which can detect a colluding group that involves in shill bidding activities [27]. The *Collusion Score* combines all the *Shill Score* ratings and gives each bidder a score based on the likelihood that they are engaging in collusive shill bidding activity [27]. However, their proposal works only when an auction ends.

However, there are very few proposals for real-time shill detection techniques. For example, Xu et al. [31] develop a multi-agent trust management system for detecting shill bidders in real-time based on some shill patterns. The authors define an *Agent-based Trust Management* (ATM) module for online auction systems. Their proposed ATM module is composed of three types of agents: the monitoring agent, the analysis agent and the security agent [31]. The monitoring agent matches bidding behaviours with shill patterns to detect shill bidders and informs the security agent about the shill suspect for further verification whether it is an actual shill or not [31]. The analysis agent of the ATM module analyses the state-based information and history information of each bidder in the online auction system [31]. The security agent keeps track of each transaction, detects unusual activities and makes decisions on shill detection [31]. When the shill bidders are detected, the authors consider increasing the suspected bidder's Shill Score by one [31]. The security agent is responsible for taking two actions: either (i) cancelling the auction; or (ii) warning the bidder, according to the fraud activities of the suspected bidder [31]. However there are some limitations in their proposed approach. First of all, a bidder cannot be blamed only on a few shill patterns. Moreover, the authors did not address the bidding patterns which they have been selected for calculating the shill score of each bidder. Besides, the authors use feedback rating for verifying suspected bidders which is not the most suitable approach because a shill bidder can involve his/her friends or relatives to give fake positive ratings on his/her auctioned product.

Later, Xu et al. [32] extends their previous work [31] into identifying shill bidding patterns in concurrent auctions. The authors introduce a *Dynamic Auction Model* (DAM) that uses real-time model checking techniques for detecting shill bidders based on *Linear Temporal Logic* (LTL) [32]. The authors assign a monitoring agent for each registered bidder which keeps track of each live auction the bidder is participating in [32]. Their proposed DAM models the basic functionality of an online auction which is initially created using an auction base model [32]. When the monitoring agent starts to perform verifications, the DAM is then updated dynamically [32]. Xu et al. [32] select a list of LTL formulas from an LTL formula library for verifying shill bidding patterns. Finally, the monitoring agent uses a *Simple Promela Interpreter* (SPIN) model checker to verify whether LTL formulas are violated or not; and take appropriate actions [32]. Nevertheless, the problem of this approach is that monitoring an auction after every single submitted bid may make the detection process computationally expensive, which is not efficient when there are a large number of participants. Furthermore, the authors did not present LTL formulas for all proposed shill bidding strategies. Therefore, it is difficult to reproduce and validate the work presented in this paper.

In contrast, Ford et al. [8] present a framework of *Real-Time Self-Adaptive Classifier* (RT-SAC) that identifies suspicious bidders in real-time using neural network method. The authors first retrieve real auction data from eBay and store it as historical data for generating training datasets [8]. The authors proposed a clustering module which parses the bidding history of each auction [8]. Then, the parsed bidding histories are used to calculate and normalize values for a collection of well-defined shill bidding attributes [8]. In the auction dataset, each bidder has its own values for such attributes (e.g. a bidder's feedback rating, number of bids) [8]. When all of the bidders are defined using these attributes, a hierarchical clustering algorithm is applied to the dataset for creating sets of grouped bidders [8]. Ford et al. [8] manually label the clusters of bidders as either *normal* or *suspicious* based on the characteristics of each cluster. When labelling of all clusters has done, the cluster labels are applied to their bidders [8]. The resulting labelled bidders create a training dataset which is used to initialize the RT-SAC [8]. After initializing the RT-SAC, it is applied to classify and update incrementally to adapt to new auction data in real time [8]. For instance, if a bidder from the real-time auction data is classified as *normal*, it is sent to the RT-SAC, and the classification result is directly used for incremental training [8]. On the other hand, if a bidder is classified as *suspicious*, it is sent to an external shill bidding verifier for further investigation which uses additional evidence to check each of the detected suspicious bidders [8]. After completing the external verification, the RT-SAC makes a decision based on the verification results by labelling the suspicious bidder either as *normal* or *suspicious* [8]. Moreover, the authors mention that their proposed RT-SAC is able to correct any false-positive classifications (i.e. normal bidders classified as suspicious) by the external verifier [8]. However, manually labelling the large training data set of users is a time consuming task.

Mamun and Sadaoui [16], and Mamun [17] propose a three-layer software architecture named *ShillFree1 auction system* for protecting auction systems from shill bidding in online auctions. Their proposed *ShillFree1 auction system* composed of: (i) the application layer (i.e. *Graphical User Interface* (GUI)) which is used by sellers and bidders for interacting with the online auction house; (ii) the data layer which keeps track of all the information and history of users and auctions (e.g., auction starting and closing time, bidding time, bidding amount); (iii) the business layer which monitors, detects and takes necessary actions against shill bidding in real-time, [16], [17]. The authors also consider an IP tracker in the business layer that determines a user who uses multiple identities with the same IP address, [16], [17]. The authors claim that their proposed mechanism keeps the auction system secured from auction fraud and also maintain trust among users as well as the online auction system [16], [17]. Although the authors [16], [17] mention the IP tracker in their proposed architecture, they do not address how to implement the IP tracker for testing their proposed shill detection system. Moreover, the authors did not address when someone fakes an IP address. For instance, *Dynamic Host Configuration Protocol* (DHCP) assigns a new IP address each time a device joins a network. Therefore, a user can just disconnect and reconnect a device with shill bidding using multiple accounts. Furthermore, they do not test the effectiveness of their proposal in any meaningful manner.

Sadaoui et al. [19] propose a generic framework that covers real-time monitoring of multiple live auctions. This framework observes the progressing auctions to be able to take action in time and prevent shill bidders from succeeding [19]. The monitoring is performed at different times (i.e., the real-time events) depending on the auction duration (in terms of hours or days). The authors [19] claim that real-time monitoring processes a smaller number of bids than offline processing; therefore the system can react quicker. Furthermore, the system can respond to potential shill bidders while an auction is in progress and can shut down suspect auctions before the innocent bidders become the victim of shill bidding [19]. The authors divide the real-time monitoring functionality into three components: (i) detecting shill bidding; (ii) reacting to shill bidding; and (iii) updating bidders' clusters [19]. The first task examines in real-time the bidding activities of bidders in running auctions using detection mechanisms [19]. The second one determines how to react to suspicious activities by taking appropriate real-time actions against suspect auctions [19]. Finally, the participants' Shill Scores and their clusters are updated dynamically after every successful or unsuccessful auction ending [19]. The participants' clusters are publicly available so that they serve as a deterrent for bidders to commit shill bidding and are the basis for reputation-based models [19]. However, the proposed shill detection task is performed after every single submitted bid and every bid is evaluated against the same set of bidding patterns, which is very time consuming.

Later, Sadaoui and Wang [20] improve their previous work [19] by proposing a framework with a dynamic agent architecture where multiple monitoring agents can be created and detected according to the status of their corresponding auctions. Their proposed dynamic agent architecture composed of three layers: (i) the user interface layer which is used to interact with the proposed auction house; (ii) the application layer which is responsible for dynamic creation and deletion of monitoring agents, extraction and storage of auction and user data, inspection of bidding activities, reaction to fraud, and updating users' fraud scores; and (iii) the data layer which stores all the details of users, running and past auctions [20]. Sadaoui and Wang [20] mention that their proposed shill bidding detection system monitors every live auction systematically at different times. These time points depend on the auction duration and *In-Auction Fraud* (IAF) types. These claim that the bidders' fraud scores are updated dynamically after every time an auction is completed or terminated [20]. Moreover, they also validate their proposed IAF monitoring service through commercial auction data [20]. In both papers [19], [20], the authors take into account eight bidding behaviours for detecting shill bidding in real-time. However, the authors did not justify the reason for selecting eight bidding behaviours among 17 proposed suspicious bidding patterns.

2.1 Problem Motivation

Given the limitations of previous proposals on shill bidding detection, it is essential to revise the existing shill detection methods and provide innovative ideas for shill bidding detection in a running auction rather than detect it afterwards. Moreover, it is necessary to take appropriate actions to shill bidders according to the fraud activities when it is detected. The research questions we are seeking to address are as follows:

1. How will our proposed shill bidding detection method be able to identify potential shill bidder in a running auction?
2. What are the typical shill bidding patterns that can be used in determining whether shill bidding is happening in a live auction?
3. What are the consequences for taking disciplinary actions against a shill bidder after detecting his/her involvement in a particular live auction?

2.2 Analysis on Shill Bidding Behaviour

Shill bidding detection mechanisms operate by looking at a series of behavioural characteristics [7], [8]. These characteristics effectively describe a bidder's bidding behaviour during an auction. Many shill bidding patterns have been identified in past research studies. However, some of these behaviours are contradictory, or are not really supported by evidence. We compiled the following shill-bidding patterns from [1], [2], [4]-[6], [8]-[10], [13], [16]-[20], [22], [26], [28], [29], [31]-[33]:

1. Bidder feedback rating [4], [9], [33] - A bidder's feedback rating can be used to determine/establish trustworthiness.
2. Early bidding [8], [10], [32] - A shill bidder usually places bids very early in an auction to stimulate bidding.
3. Avoid late bidding [4], [26], [29] - A shill bidder stops placing bids before an auction ends in order to reduce his/her chance of winning the auction.
4. Small bid increment [9], [26], [29] - A shill bidder outbids with small bid increments in order to reduce the risk of not being outbid.
5. Rapid outbid time [8], [26], [29] - A shill bidder is more likely to quickly outbid a legitimate bid.
6. Average increment difference [8] - A shill bidder's average bid increment may decrease throughout the auction.
7. Average time difference user bids [8] - A shill bidder's average time to outbid may decrease throughout an auction.
8. Bid frequency [4], [26], [29] - A shill bidder generally has a high number of bids compared to other legitimate bidders.
9. Auction starting price [4], [5], [8] - This is the impact of the auction starting price on a bidder's decision to participate in the auction. A low starting price is more conducive to shill bidding as legitimate bidders are more likely to participate in the auction. Alternatively, a high starting price may deter legitimate bidders from participating in the auction.
10. Affinity to the sellers [9], [26], [29] - A shill bidder usually participates in auctions run by a particular seller (as the shill is working for the seller).
11. Successive bidding [16], [19], [20] - A shill bidder continually bids to outbid his/her own bid even he/she is the top bidder in the auction to increase the final price.
12. Reserve price shilling [19], [20] - A shill bidder will list a low reserve price in order to attract bidders. The shill bidder will then attempt to inflate the auction price to the true value of the item.
13. Low win ratio [4], [26], [29] - A shill bidder will have a low number of wins compared to the number of auctions he/she has participated in.

14. Concurrent bid activity [2], [5], [18] - In the event where there are two concurrent auctions underway, a legitimate bidder will bid in the auction with the lower price. However, a shill bidder will not follow this behaviour.
15. Seller feedback rating [8], [9], [33] - Most bidders are more likely to bid in an auction if the seller has significant positive feedback in the past auctions.
16. IP tracking [16], [17] - A shill bidder places bids multiple times on the same item using the same computer/device.

Table 1 summarizes all the shill bidding behaviours for past studies. Each behaviour is identified by a unique Behaviour ID. We have also provided a tally of how many research papers support a particular behaviour.

Table 1: Bidding behaviours of shill bidding

Behaviour ID	Bidding Behaviours	Total research papers in support
1	Bidder feedback rating	5
2	Early bidding	10
3	Avoid late bidding	11
4	Small bid increment	14
5	Rapid outbid time	10
6	Average increment difference	1
7	Average time difference user bids	1
8	Bid frequency	16
9	Auction starting price	3
10	Affinity to a seller	11
11	Successive bidding	4
12	Reserve price shilling	2
13	Low win ratio	11
14	Concurrent bid activity	6
15	Seller feedback rating	3
16	IP tracking	2

Figure 1 illustrates an analysis of bidding behaviours of shill bidders. The Behaviour IDs and paper counts from Table 1 are used. There appear to be seven (7) bidding behaviours referred to in the literature that have gained the most popularity. Based on this endorsement, this paper will focus on these popular behaviours only.

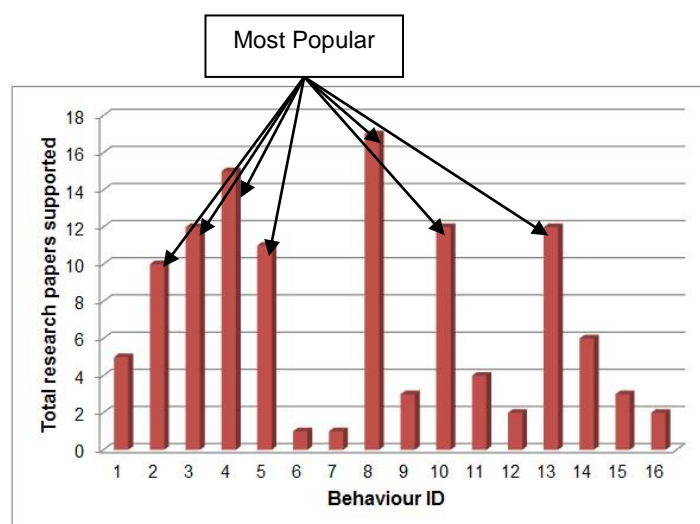


Figure 1: Analysis of bidding behaviours of shill bidders

3 Run-Time Shill Bidding Detection

Our shill detection and prevention methodology considers bidding and auction properties while an auction is running. We propose a *Live Shill Score (LSS)* as a way to represent a bidder's behaviour during an auction. The *LSS* is calculated for each bidder during an auction based upon the selection of bidding behaviours/characteristics defined as

the most popular in Section 2.2. The *LSS* can be used to potentially identify shill bidders, enforce disciplinary actions during an auction, and acts as a deterrent to a potential shill bidder.

3.1 Assumptions

To narrow the scope of what the run-time *LSS* algorithm is trying to achieve, we consider a live auction and a number of past auctions hosted by a seller. Though we focus on shill bidding detection in a running auction in this paper, each bidder's bid history in past auctions is used to provide further evidence of potential shill bidding behaviour during the *LSS* calculation of the bidder in the running auction.

Furthermore, each bidder's bidding patterns throughout the running auction has been undertaken by considering the assumptions of our *LSS* algorithm. In addition, we have also tested the effect of each bidder's past bidding behaviours in a running auction by introducing *Post-Filtering Process* (PFP) in the *LSS* algorithm. As such comprehensive testing of the approach has not been undertaken by previous proposals; our shill detection approach in a running auction advances the state of the art over existing proposals. The assumptions we consider in this paper will allow us to develop the shill detection model without being embroiled in extraneous detail. In future work, we will relax some of these assumptions as the basic model is introduced.

In this paper, we do not address concurrent auctions [30]. Moreover, we are not investigating reserve price shilling [11]. We are focusing on an ascending-price, open bid online auction format with a pre-determined closing time where the winner pays a price equal to the highest bid (i.e., similar to eBay).

3.2 Monitoring Stages of a Running Auction

Bids submitted at certain times throughout the auction can be considered more conducive to shill bidding. For example, it becomes more risky to submit shill bids closer to the end of an auction as there is an increased likelihood that the shill bidder will win the auction if he/she is not outbid in time. Therefore, our proposed *LSS* algorithm breaks an auction into a series of stages depending on the time elapsed for reacting to shill bidder and other legitimate bidders in a running auction. This strategy was first suggested in Xu et al.'s real-time shill detection proposal [32]. The different auction stages are as follows:

1. Early Stage - This is the first quarter of the auction duration (up to 25% of the auction duration). A shill bidder usually places bids early in an auction to encourage bidding.
2. Middle Stage - This is the time period between 25.1% and 80% of the auction duration.
3. Late Stage - The next 15% (between 80.1% and 95%) of the auction is the late stage. A shill bidder will try to avoid placing bids during this period to reduce his/her chance of winning the auction.

In our run-time shill detection methodology, we consider the above stages for monitoring shill bidding patterns in a live auction. However, we differ from Xu et al. [32] by introducing an additional stage, referred to as the Final stage:

1. Final Stage - The last 5% of the auction duration is used for verifying the detected bidding patterns from the early, middle and late stages. The overall *LSS* value of each bidder will be calculated during this stage.

The reason for the final stage is that Xu et al. [32] did not clearly address when reactions would be taken against the potential shill bidders after the end of the late stage. Therefore, the final stage is required for taking appropriate actions against potential shill bidders after the late stage and for exonerating bidders with good behaviour. Section 5.4 will discuss the rationale for selecting 25-80-95% cuts.

3.3 The Run-Time Shill Bidding Detection Process

Figure 2 illustrates the process for detecting shill bidding in run-time. Time flows from left to right.

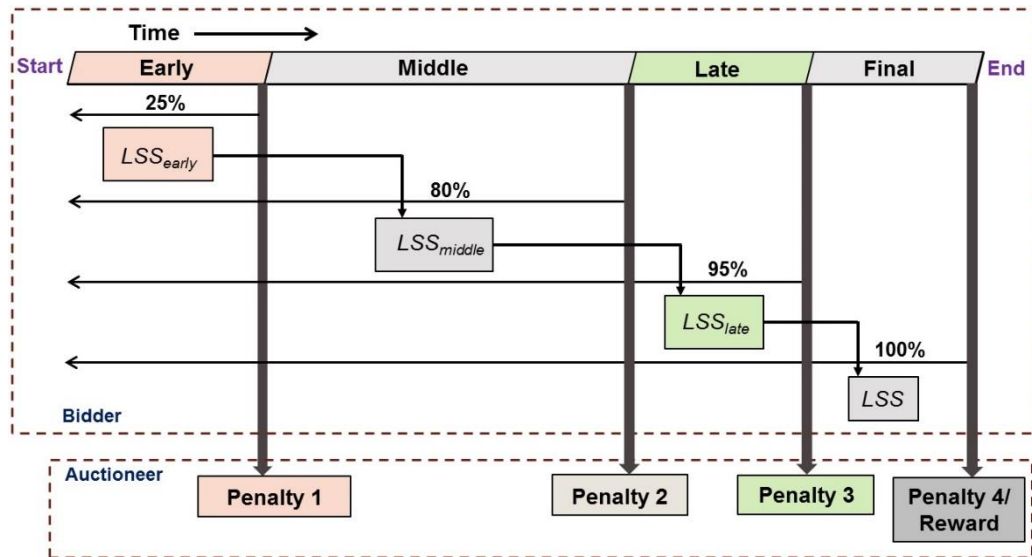


Figure 2: Run-time shill bidding detection process

An auction is broken up into the four aforementioned stages (i.e., early, middle, late, final). A bidder's behaviour is examined against other bids which are placed by the bidder during each auction stage. An LSS for each bidder is calculated during each respective auction stage as follows:

1. LSS_{early} - Calculated after completing 25% of the auction. If the bidder who has a LSS_{early} higher than a predetermined threshold, the auctioneer will generate Penalty 1 to him/her.
2. LSS_{middle} - Calculated after completing 80% of the auction. If the bidder who has a LSS_{middle} higher than a predetermined threshold, the auctioneer will generate Penalty 2 to him/her.
3. LSS_{late} - Calculated after completing 95% of the auction. If the bidder who has a LSS_{late} higher than a predetermined threshold, Penalty 3 will be generated by the auctioneer.
4. LSS - LSS is calculated based on the bidding behaviours of all bidders throughout the entire auction. Once the overall LSS is calculated, a Post-Filtering Process (PFP) is used based on an evaluation of how the bidder behaved during each auction stage. The result of the PFP either generates Penalty 4 to the potential shill bidder or exonerates legitimate bidders (see in Section 4.2).

The higher the bidder's LSS , the increased likelihood the bidder is engaging in price inflating behaviour.

3.4 Reactions to Shill Bidders

After every stage of the auction, the algorithm will determine whether any action is required based on the bidding behaviour so far in the auction. Apart from Xu et al. [32], to our knowledge there is no literature discussing potential penalties or the impact of these penalties on bidders' bidding behaviour. We propose the following penalties for our algorithm:

1. Penalty 1 (At the end of the early stage) - Send a warning message to the suspected bidders. This penalty encourages honest bidders to continue bidding without causing anxiety. A potential shill bidder would also realise that the auction is being monitored, thereby forcing him/her to rethink his/her behaviour.
2. Penalty 2 (At the end of the middle stage) - Pause the auction temporarily and notify all participating bidders about the suspicious activity. This penalty has a stronger psychological impact on auction participants. The penalty sends a message that the auction is potentially at risk of shill bidding. Therefore, everyone is on notice to improve their behaviour. At this point, honest bidders can decide whether they want to remain involved with the auction or not. The shill bidder also knows that his/her behaviour has been detected which should serve as a deterrent for further shill bidding.

3. Penalty 3 (At the end of the late stage) - Postpone the auction and also ask suspicious bidders and/or the seller to show cause for why the auction should continue. This penalty informs honest bidders this is a monitored and safe bidding system. The suspicious bidders and/or seller would realise that they need to improve their behaviour. Otherwise the postponed auction will not be reopened. The penalty reduces the potential for monetary loss.
4. Penalty 4/ Reward (At the end of the final stage) - Cancel the auction or exonerate innocent bidders. The *LSS* for a bidder showing stereotypical shill bidding behaviour is reinforced/increased. In contrast, bidders whose behaviour appears to be normal are exonerated (i.e., the winner of the auction, late bidders, bid snipers).

3.5 Shill Bidding Behaviour

Section 2.2 (Figure 1) identified the seven most popular shill bidding characteristics based on an analysis of the literature. As *early bidding* and *avoid late bidding* indicate the same bidding characteristics (refer to Section 2.2), we consider six bidding patterns in our *LSS* algorithm. Table 2 shows the six bidding behaviours for calculating the *LSS* value for each bidder in a running auction.

Table 2: Bidding behaviour for run-time detection of shill bidding

Rating	Bidding Behaviours
α	Affinity to a seller
β	Bid frequency
γ	Win/lose factor
δ	Rapid outbid time
ϵ	Small bid increment
ζ	Early bidding

Bidding behaviours the *LSS* examines are as follows:

1. Bid frequency (β rating) - β rating indicates the number of bids submitted by a bidder so far up until the end of the particular auction stage.
2. Rapid outbid time (δ rating) - δ rating refers to the time difference between when a bidder places a new highest bid and the previous highest bid submitted by another bidder during a particular auction stage.
3. Small bid increment (ϵ rating) - ϵ rating refers to a value that a bidder outbids the current highest bidder during a particular auction stage.
4. Early bidding (ζ rating) - ζ rating denotes the time difference between the starting time of an auction and a bidder's first bid in the auction.

For calculating the final *LSS* for each bidder, we are considering the following additional bidding behaviours:

1. Affinity to a seller (α rating) - A shill bidder usually participates in auctions run by one particular seller. A bidder with a high α rating is suspicious to be a shill bidder. In this paper, we will not consider this behaviour for calculating *LSS* of each bidder as we assume only one auction (refer to Section 3.1). However, this behaviour has been used in our post-filtering process for exonerating innocent bidders in simulated and commercial auction data (see in Section 5.1 and 5.2).
2. Win/lose factor (γ rating) - A shill bidder avoids winning an auction, as the auction will have to be repeated. A bidder with a low γ rating indicates suspicious behaviour. It can be mentioned that a winner of an auction is not a shill bidder. However, this paper is only considering one auction at a time. Therefore, the γ rating of a bidder is 0 when the bidder wins the auction or 1 otherwise.

Table 3 shows the bidding behaviours considered at different stages (i.e., early, middle, late and final stage) for calculating the *LSS* of each bidder in an auction.

Table 3: Bidding behaviours with different stages of an auction duration

Stages	Bidding Behaviours
Early Stage	Bid frequency (β rating)
Middle Stage	Rapid outbid time (δ rating)
Late Stage	Small bid increment (ϵ rating)
	Early bidding (ζ rating)
Final Stage	Bid frequency (β rating)
	Win/Lose factor (γ rating)
	Rapid outbid time (δ rating)
	Small bid increment (ϵ rating)
	Early bidding (ζ rating)

4 Calculating the Live Shill Score

This section describes our proposed *LSS* algorithm to detect a shill bidder by looking for common shill bidding behaviours. Our *LSS* algorithm considers the case with only one shill bidder. The goal of this algorithm is to determine which bidder is most likely to be the shill out of a group of I bidders. A bidder is examined over single auction held by a seller for the behaviour outlined in Section 3.5. Each characteristic of shill behaviour is given a rating which is combined to form the bidder's *LSS*. The *LSS* gives a bidder a value between 0 and 10. The closer the *LSS* is to 10, the corresponding bidder will be forwarded to the post-filtering process. If the bidder shows the shill bidding patterns in the post-filtering process, then the bidder is a potential shill bidder in the auction. Otherwise, the bidder will be exonerated from shill bidding.

The remainder of this section describes how to quantify the bidding characteristics of a shill bidder to deduce a shill score. However, to do this some preliminary notation is required:

Let $L = \{1, \dots, I\}$ be the set of bidder numbers;

$$|L| = I$$

Let $M = \{1, \dots, m\}$ be the set of auction numbers;

$$|M| = m$$

For $j \in M$, let $N_j = \{1, \dots, n_j\}$ be the bid numbers (e.g., 1st bid, 2nd bid, etc) in auction j ;

$$|N_j| = n_j$$

For $i \in L$, let $N_j^i = \{k \mid k \in N_j, \text{ bidder } i \text{ makes the } k\text{th bid in auction } j\}$ be the bid numbers (e.g., 1st bid, 2nd bid, etc) in auction j ;

$$|N_j^i| = n_j^i$$

For $i \in L$, let $W^i = \{j \mid j \in M, \text{ bidder } i \text{ wins auction } j\}$;

$$|W^i| = w^i$$

For $j \in M$, $k \in N_j$, let $t_{j,k}$ be the time of the k th bid in auction j , let

$$T_j = \{t_j \mid j \in M, k \in N_j\}$$

be the set of bid times for auction j .

For $i \in L$, $j \in M$, $k \in N_j$, let $t_{j,k}^i$ be the time of the k th bid in the auction j by the bidder i . Let

$$T_j^i = \{t_j^i \mid i \in L, j \in M, k \in N_j\}$$

be the set of bid times for bidder i in auction j ;

$$|T_j^i| = n_j^i$$

For $j \in M$, $k \in N_j$, let $p_{j,k}$ be the value of the k th bid in auction j , let

$$P_j = \{p_{j,k} \mid j \in M, k \in N_j\}$$

be the set of bid values in auction j ;

$$|P_j| = n_j$$

For $i \in L$, $j \in M$, $k \in N_j$, let $p_{j,k}^i$ be the value of the k th bid in the auction j by the bidder i . Let

$$P_j^i = \{p_{j,k}^i \mid i \in L, j \in M, k \in N_j\}$$

be the set of bid values for bidder i in auction j ;

$$|P_j^i| = n_j^i$$

4.1 Calculation of the Bidder's Rating

We first need to calculate β , δ , ε , ζ , and γ ratings in a particular stage of auction duration. Then we can calculate the value LSS of each bidder in every stage of the auction duration (see Figure 3).

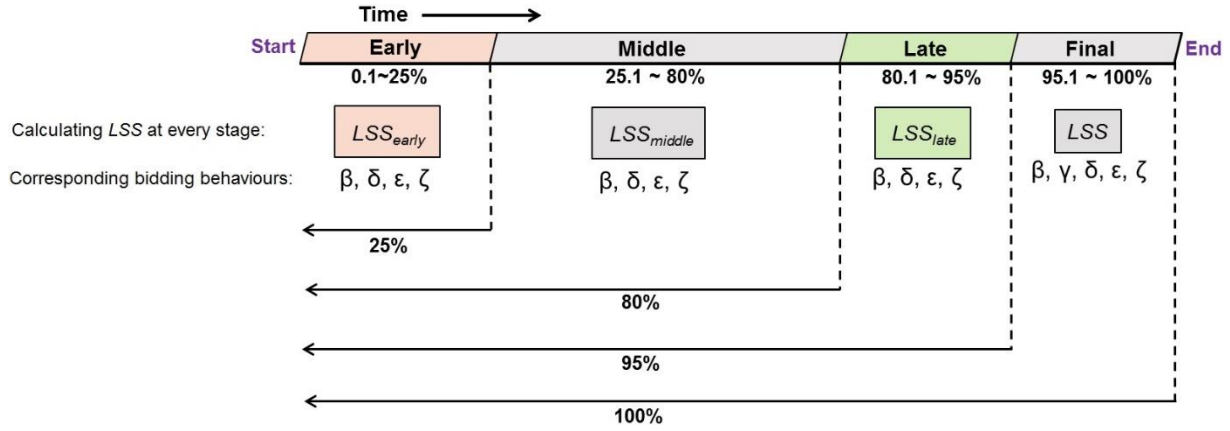


Figure 3: Calculation process of a bidder's rating and LSS

Now the calculations of β , δ , ε , ζ , and γ ratings in the different stages of auction duration (i.e., early, middle, late, final stages) are shown as below:

- β rating: An aggressive shill bidder tends to place a high number of bids as they are continually trying to outbid other legitimate bidders. This might involve the shill bidding after every legitimate bid [26]. In the worst case scenario, a shill bidder must place $\left\lceil \frac{n_j}{2} \right\rceil$ bids, where n_j denotes the total number of bids for an auction j . If a shill bidder submits any more bids than $\left\lceil \frac{n_j}{2} \right\rceil$ bids, he/she will win the auction j [26]. This bound holds regardless of how many bidders there are in the auction [26].

The β rating indicates the average percentage of bids that bidder i has submitted throughout in an auction he/she has participated in. The percentage for an individual auction is calculated in terms of the worst case bound of $\left\lceil \frac{n_j}{2} \right\rceil$ bids [26]. The β rating for each bidder i in an auction j can be calculated as follows:

Let β_j^i be bidder i 's β rating for auction j .

Step 1: If the auction is in the early, middle, or late stage; then proceed as below. Otherwise go to Step 2.

- Calculate the number of bids, $n_{j_early/middle/late}^i$, that have been submitted by all bidders in the corresponding stage (i.e., early, middle, late stage) of auction j .
- Calculate the worst case number of bids, $w_{j_early/middle/late}$, for auction j :

$$w_{j_early/middle/late} = \left\lceil \frac{n_{j_early/middle/late}}{2} \right\rceil$$

- Calculate the number of bids, $n_{j_early/middle/late}^i$, bidder i has submitted in the corresponding stage (i.e., early, middle, late stage) of auction j .
- Calculate the percentage of bids made by bidder i in terms of the worst case bound $w_{j_early/middle/late}$:

$$\beta_{j_early/middle/late}^i = \frac{n_{j_early/middle/late}^i}{w_{j_early/middle/late}}, \text{ where } 0 \leq \beta_{j_early/middle/late}^i \leq 1$$

- If the auction is in progress, repeat from 1 to 4 of Step 1. Otherwise go to Step 2.

Step 2: If the auction is at the end of the final stage, then proceed as below:

1. If bidder i wins the auction j , then the β rating of the bidder i in auction j , β_j^i , will be zero. Otherwise perform the remaining steps.
2. Calculate the number of bids, n_j , that have been submitted by all bidders in auction j .
3. Calculate the worst case number of bids, w_j , for auction j :

$$w_j = \left\lfloor \frac{n_j}{2} \right\rfloor$$

4. Calculate the number of bids, n_j^i , bidder i has submitted in the correspond stage (i.e., early, middle, late stage) of auction j .
5. Calculate the percentage of bids made by bidder i in terms of the worst case bound w_j :

$$\beta_j^i = \frac{n_j^i}{w_j}, \text{ where } 0 \leq \beta_j^i \leq 1$$

So, the final β rating of any bidder i participated in auction j is as follows:

$$\beta_j^i = \begin{cases} 0, & \text{if bidder } i \text{ wins the auction } j \\ \frac{n_j^i}{w_j}, & \text{Otherwise} \end{cases}, \text{ where } 0 \leq \beta_j^i \leq 1 \quad (1)$$

In general, the β rating will be high for an aggressive skill bidder.

- δ rating: A skill bidder wants to give other honest bidders as much time as possible to consider a skill bid [26]. Therefore, a skill bidder generally submits a new bid within a small time period of a rival bid [26]. This bidding pattern can be measured by observing inter bid times for all bidders. The average inter bid time is found for bidder i participated in auction j . The bidders who wait longer between bids have a lower average inter bid time score [26]. The δ rating for each bidder i in an auction j can be calculated as follows:

Let δ_j^i be bidder i 's δ rating for auction j .

Step 1: If the auction is in the early, middle, or late stage; then proceed to the following steps. Otherwise go to Step 2.

1. Calculate the inter bidding time for each bid, k , submitted by bidder i in auction j :

$$\Delta t_{j_{\text{early/middle/late}},k}^i = \begin{cases} 0, & k=1 \\ t_{j_{\text{early/middle/late}},k}^i - t_{j_{\text{early/middle/late}},k-1}^i, & k > 1, k \in N_{j_{\text{early/middle/late}}} \end{cases}$$

where $t_{j_{\text{early/middle/late}},k}^i$ represents the time when bidder i placed a current highest bid during a particular auction stage and $t_{j_{\text{early/middle/late}},k-1}^i$ denotes the time of a previous bid placed by a rival bidder.

2. Calculate the average inter bidding time for bidder i :

$$\bar{\Delta t}_{j_{\text{early/middle/late}}}^i = \frac{1}{n_{j_{\text{early/middle/late}}}^i} \sum_{k \in N_{j_{\text{early/middle/late}}}^i} \Delta t_{j_{\text{early/middle/late}},k}^i$$

3. The maximum and minimum average inter bidding time for auction j are:

$$\begin{aligned} \bar{\Delta t}_{j_{\text{early/middle/late}}}^{\max} &= \max_{i \in L} \{ \bar{\Delta t}_{j_{\text{early/middle/late}}}^i \} \\ \bar{\Delta t}_{j_{\text{early/middle/late}}}^{\min} &= \min_{i \in L} \{ \bar{\Delta t}_{j_{\text{early/middle/late}}}^i \} \end{aligned}$$

Then, perform the normalisation for finding bidder i 's $\delta_{j_{\text{early/middle/late}}}^i$ for auction j .

$$\delta_{j_{\text{early/middle/late}}}^i = \frac{\bar{\Delta t}_{j_{\text{early/middle/late}}}^i - \bar{\Delta t}_{j_{\text{early/middle/late}}}^{\min}}{\bar{\Delta t}_{j_{\text{early/middle/late}}}^{\max} - \bar{\Delta t}_{j_{\text{early/middle/late}}}^{\min}}$$

Now, calculate the normalised inter bidding time for bidder i over the auction j participated in:

$$\delta_{\text{early/middle/late}}^i = 1 - \delta_{\text{early/middle/late}}^i, \text{ where } 0 \leq \delta_{\text{early/middle/late}}^i \leq 1$$

4. If the auction is in progress, repeat from 1 to 3 of Step 1. Otherwise go to Step 2.

Step 2: If the auction is at the end of the final stage, then proceed to the steps as below:

1. If bidder i wins the auction j , then the δ rating of the bidder i in auction j , δ_j^i , will be zero. Otherwise perform the remaining steps.
2. Calculate the inter bidding time for each bid, k , submitted by bidder i in auction j :

$$\Delta t_{j,k}^i = \begin{cases} 0, & k=1 \\ t_{j,k}^i - t_{j,k-1}^i, & k > 1, k \in N_j \end{cases}$$

where $t_{j,k-1}^i$ is the time of a previous bid submitted by a rival bidder.

3. Calculate the average inter bidding time for bidder i : N_j

$$\bar{\Delta t}_j^i = \frac{1}{n_j} \sum_{k \in N_j} \Delta t_j^i$$

4. The maximum and minimum average inter bidding time for auction j are:

$$\begin{aligned} \bar{\Delta t}_j^{\max} &= \max_{i \in L} \{\bar{\Delta t}_j^i\} \\ \bar{\Delta t}_j^{\min} &= \min_{i \in L} \{\bar{\Delta t}_j^i\} \end{aligned}$$

Then, perform the normalisation for finding bidder i 's δ_j^i for auction j .

$$\delta_j^i = \frac{\bar{\Delta t}_j^i - \bar{\Delta t}_j^{\min}}{\bar{\Delta t}_j^{\max} - \bar{\Delta t}_j^{\min}}$$

Finally, calculate the normalised inter bidding time for bidder i over the auction j participated in:

$$\delta_j^i = 1 - \delta_j^i, \text{ where } 0 \leq \delta_j^i \leq 1$$

So, the final δ rating of any bidder i participated in auction j is as follows:

$$\delta_j^i = \begin{cases} 0, & \text{if bidder } i \text{ wins the auction } j \\ 1 - \delta_j^i, & \text{Otherwise} \end{cases}, \text{ where } 0 \leq \delta_j^i \leq 1 \quad (2)$$

In general, the δ rating will be higher for a skill bidder in comparison to a legitimate bidder.

- ϵ rating - A skill bidder tends to only bid the minimal amount to stay ahead of the current highest bid. The reason is that, if a skill bidder outbids by a large amount, it will increase the risk of winning the auction by the skill bidder (as a skill bidder avoids winning the auction). This behaviour can be measured by observing inter bid increments for all bidders [26]. The average inter bid increment is found for bidder i who participated in auction j . Bidders who submit smaller bid increments have a lower average inter bid increment score [26]. The ϵ rating for each bidder i in an auction j can be calculated as follows:

Let ϵ_j^i be bidder i 's ϵ rating for auction j .

Step 1: If the auction is in the early, middle, or late stage; then proceed to the following steps. Otherwise go to Step 2.

1. Calculate the inter bid increment for each bid, k , submitted by bidder i in auction j :

$$\Delta p_{\text{early/middle/late}}^i = \begin{cases} 0, & k=1 \\ p_{\text{early/middle/late}}^i - p_{\text{early/middle/late}}^i, & k > 1, k \in N_{\text{early/middle/late}} \end{cases}$$

where $p_{j, \text{early/middle/late}, k}^i$ denotes a current high bid placed by bidder i during a particular stage and $p_{j, \text{early/middle/late}, k-1}^i$ is the previous bid of $p_{j, \text{early/middle/late}, k}^i$ placed by any participating bidder in a particular stage of the auction j .

2. Calculate the average inter bid increment for bidder i :

$$\bar{\Delta p}_{j, \text{early/middle/late}}^i = \frac{1}{n_j^i} \sum_{k \in N_j^i} \Delta p_{j, \text{early/middle/late}, k}^i$$

3. The maximum and minimum average inter bid increment for auction j are:

$$\begin{aligned} \bar{\Delta p}_{j, \text{early/middle/late}}^{\max} &= \max_{i \in L} \{ \bar{\Delta p}_{j, \text{early/middle/late}}^i \} \\ \bar{\Delta p}_{j, \text{early/middle/late}}^{\min} &= \min_{i \in L} \{ \bar{\Delta p}_{j, \text{early/middle/late}}^i \} \end{aligned}$$

Then, perform the normalisation for finding bidder i 's $\varepsilon_{j, \text{early/middle/late}}^i$ for auction j .

$$\hat{\varepsilon}_{j, \text{early/middle/late}}^i = \frac{\bar{\Delta p}_{j, \text{early/middle/late}}^i - \bar{\Delta p}_{j, \text{early/middle/late}}^{\min}}{\bar{\Delta p}_{j, \text{early/middle/late}}^{\max} - \bar{\Delta p}_{j, \text{early/middle/late}}^{\min}}$$

Now, calculate the normalised inter bid increment for bidder i over the auction j participated in:

$$\varepsilon_{j, \text{early/middle/late}}^i = 1 - \hat{\varepsilon}_{j, \text{early/middle/late}}^i, \text{ where } 0 \leq \varepsilon_{j, \text{early/middle/late}}^i \leq 1$$

4. If the auction is in progress, repeat from 1 to 3 of Step 1. Otherwise go to Step 2.

Step 2: If the auction is at the end of the final stage, then proceed to the steps as below:

1. If bidder i wins the auction j , then the ε rating of the bidder i in auction j , ε_j^i , will be zero. Otherwise perform the remaining steps.
2. Calculate the inter bid increment for each bid, k , submitted by bidder i in auction j :

$$\Delta p_{j, k}^i = \begin{cases} 0, & k=1 \\ p_{j, k}^i - p_{j, k-1}^i, & k > 1, k \in N_j^i \end{cases}$$

where $p_{j, k-1}^i$ is the price of a previous bid submitted by a rival bidder.

3. Calculate the average inter bid increment for bidder i :

$$\bar{\Delta p}_j^i = \frac{1}{n_j^i} \sum_{k \in N_j^i} \Delta p_j^i$$

4. The maximum and minimum average inter bid increment for auction j are:

$$\begin{aligned} \bar{\Delta p}_j^{\max} &= \max_{i \in L} \{ \bar{\Delta p}_j^i \} \\ \bar{\Delta p}_j^{\min} &= \min_{i \in L} \{ \bar{\Delta p}_j^i \} \end{aligned}$$

Then, perform the normalisation for finding bidder i 's ε_j^i for auction j .

$$\hat{\varepsilon}_j^i = \frac{\bar{\Delta p}_j^i - \bar{\Delta p}_j^{\min}}{\bar{\Delta p}_j^{\max} - \bar{\Delta p}_j^{\min}}$$

Finally, calculate the normalised inter bid increment for bidder i over the auction j participated in:

$$\varepsilon_j^i = 1 - \hat{\varepsilon}_j^i, \text{ where } 0 \leq \varepsilon_j^i \leq 1$$

So, the final ϵ rating of any bidder i participated in auction j is as follows:

$$\epsilon_j^i = \begin{cases} 0, & \text{if bidder } i \text{ wins the auction } j \\ 1 - \epsilon_j^i, & \text{Otherwise} \end{cases}, \text{ where } 0 \leq \epsilon_j^i \leq 1 \quad (3)$$

In general, the ϵ rating will be higher for a skill bidder in comparison to a legitimate bidder.

- ζ rating - If a skill bidder places a bid late in an auction, he/she risks winning the auction by not being outbid in time [26]. Therefore, a skill bidder submits bids early in an auction to avoid winning any auction. The ζ rating represents how early in an auction bidder i placed bidding [26]. This behaviour can be measured by the difference between the auction's expiration time and the time that bidder i first submitted a bid. This result is then normalised against all other participants in the auction. An individual's ζ rating is calculated as follows:

Let ζ_j^i be bidder i 's ζ rating for auction j .

Step 1: If the auction is in the early, middle, or late stage; then proceed as below. Otherwise go to Step 2.

1. Calculate the difference between auction j 's starting time, t_j , and the time, $t_{j, \text{early/middle/late}, k_0}^i$ of the first bid placed by bidder i in auction j .

$$\Delta t_{j, \text{early/middle/late}}^i = t_{j, \text{early/middle/late}, k_0}^i - t_j$$

2. The maximum and minimum average final bidding time difference for auction j are:

$$\begin{aligned} \Delta t_{j, \text{early/middle/late}}^{\max} &= \max_{i \in L} \{ \Delta t_{j, \text{early/middle/late}}^i \} \\ \Delta t_{j, \text{early/middle/late}}^{\min} &= \min_{i \in L} \{ \Delta t_{j, \text{early/middle/late}}^i \} \end{aligned}$$

Then, perform the normalisation for finding bidder i 's $\zeta_{j, \text{early/middle/late}}^i$ for auction j .

$$\zeta_{j, \text{early/middle/late}}^i = \frac{\Delta t_{j, \text{early/middle/late}}^i - \Delta t_{j, \text{early/middle/late}}^{\min}}{\Delta t_{j, \text{early/middle/late}}^{\max} - \Delta t_{j, \text{early/middle/late}}^{\min}}, \text{ where } 0 \leq \zeta_{j, \text{early/middle/late}}^i \leq 1$$

3. If the auction in progress, repeat 1 and 2 of Step 1. Otherwise go to Step 2.

Step 2: If the auction is at the end of the final stage, then proceed as below:

1. If bidder i wins the auction j , then the ζ rating of the bidder i in auction j , ζ_j^i , will be zero. Otherwise perform the remaining steps.
2. Calculate the difference between auction j 's starting time, t_j , and the time, t_{j, k_0}^i of the first bid placed by bidder i in auction j .

$$\Delta t_j^i = t_{j, k_0}^i - t_j$$

3. The maximum and minimum average final bidding time difference for auction j are:

$$\begin{aligned} \Delta t_j^{\max} &= \max_{i \in L} \{ \Delta t_j^i \} \\ \Delta t_j^{\min} &= \min_{i \in L} \{ \Delta t_j^i \} \end{aligned}$$

Then, perform the normalisation for finding bidder i 's ζ_j^i for auction j .

$$\zeta_j^i = \frac{\Delta t_j^i - \Delta t_j^{\min}}{\Delta t_j^{\max} - \Delta t_j^{\min}}$$

Finally, calculate the normalised bidding time difference for bidder i over the auction j participated in:

$$\zeta_j^i = 1 - \zeta_{j, \text{early/middle/late}}^i, \text{ where } 0 \leq \zeta_j^i \leq 1$$

So, the final ζ rating of any bidder i participated in auction j is as follows:

$$\zeta_j^i = \begin{cases} 0, & \text{if bidder } i \text{ wins the auction } j \\ 1 - \zeta_j^i, & \text{Otherwise} \end{cases}, \text{ where } 0 \leq \zeta_j^i \leq 1 \quad (4)$$

In general, the ζ rating will be higher for a skill bidder in comparison to a legitimate bidder.

In the final stage of the auction duration, an additional rating (γ rating) is calculated including the above four ratings for measuring the final LSS of each bidder participated in the auction.

- γ rating: A skill bidder avoids to win an auction, as the auction will have to be repeated. This behaviour is defined as γ rating. A bidder with a low γ rating indicates suspicious behaviour. The γ rating for each bidder i can be calculated as follows:

$$\gamma^i = 1 - \left(\frac{w^i}{m^i} \right) \quad (5)$$

where w^i denotes the number of auctions won by the bidder i and m^i represents the total number of auctions participated by the bidder i . In this paper, as we are assuming only one auction, the γ rating of each participating bidder i in the auction j will be as follows:

$$\gamma_j^i = \begin{cases} 0, & \text{if bidder } i \text{ wins the auction } j \\ 1, & \text{Otherwise} \end{cases} \quad (6)$$

4.2 LSS Calculation

After calculating β , δ , ϵ , ζ , and γ ratings in a particular stage of auction duration, then we can calculate the value LSS in every stage (see Figure 4).

In the early stage, LSS_{early} is calculated as follows:

$$LSS_{early} = \frac{\theta_1 \beta_{early} + \theta_2 \delta_{early} + \theta_3 \epsilon_{early} + \theta_4 \zeta_{early}}{\theta_1 + \theta_2 + \theta_3 + \theta_4} \times 10 \quad (7)$$

At the middle stage, LSS_{middle} is calculated as follows:

$$LSS_{middle} = \frac{\theta_1 \beta_{middle} + \theta_2 \delta_{middle} + \theta_3 \epsilon_{middle} + \theta_4 \zeta_{middle}}{\theta_1 + \theta_2 + \theta_3 + \theta_4} \times 10 \quad (8)$$

At the late stage, LSS_{late} is calculated as follows:

$$LSS_{late} = \frac{\theta_1 \beta_{late} + \theta_2 \delta_{late} + \theta_3 \epsilon_{late} + \theta_4 \zeta_{late}}{\theta_1 + \theta_2 + \theta_3 + \theta_4} \times 10 \quad (9)$$

At the final stage, LSS is calculated as follows:

$$LSS = \frac{\theta_1 \beta + \theta_2 \delta + \theta_3 \epsilon + \theta_4 \zeta + \theta_5 \gamma}{\theta_1 + \theta_2 + \theta_3 + \theta_4 + \theta_5} \times 10 \quad (10)$$

where θ is the weighted coefficient associated with each rating in Equations (6), (7), (8) and (9). The weights used in this paper are: $\theta_1 = 2$, $\theta_2 = 2$, $\theta_3 = 2$, $\theta_4 = 2$, $\theta_5 = 5$. These weightings are obtained by experimenting with simulated auction data and using the following justifications:

The γ rating is given the highest weighting because it is the strongest sign that the winner of the auction is not a skill bidder. Next, the β rating measures the number of bids submitted in terms of $\left[\frac{n_i}{2} \right]$ [26]. This β rating will be high for an aggressive skill bidder or low for one-time bidder. It may happen that a one-time bidder submits bid most of the auctions either early or middle stage of the auction duration. As a result, the β rating can present mixed results depending on the type of skill behaviour employed. For this reason, the β rating has given a lower weighting to consider both type of bidders.

The δ and ϵ weightings also receive a lower rating like β weighting because of the effect of one-time bidders. That is, if a bidder only bids once, placing the bid quickly after the current highest bid and by minimal amount required, then the bidder will have high δ and ϵ ratings. As the bidder does not bid again, his/her δ and ϵ ratings will always remain

high and not average down if the bidder submits slower and larger bids later. Therefore, weightings of δ and ϵ highly results in many one-time bidders scoring high overall even though such bidders are clearly innocent.

The ζ weightings are also given two as it is not more influential than other bidding behaviour ratings. Instead all the bidding characteristics weightings must be examined as a group to determine if the bidder's bidding behaviour fits the profile of a shill bidder.

The LSS values in all stages of the auction duration are between 0 and 10. The bidder with the highest LSS is the most likely to be a potential shill bidder.

Algorithm 1 describes our proposed LSS algorithm.
Algorithm 1 LSS Algorithm

INPUT: A set of auction data in a running auction j

OUTPUT: LSS of each bidder i in the auction j , Penalties for reacting to the suspicious bidder, and Reward (Pardon) for exonerating innocent bidders.

```

Begin
Initially, split a live auction  $j$  into four time periods.
Set penalty threshold value  $\rho_{\text{early}}, \rho_{\text{middle}}, \rho_{\text{late}}, \rho_{\text{final}}$ 
if the auction  $j$  is in progress then
    if (current time == end of early stage) then
        for each bidder  $i$  in the auction  $j$  do
            Calculate  $LSS(i, j_{\text{early}})$ 
        end for
        if ( $LSS(i, j_{\text{early}}) > \rho_{\text{early}}$ ) then
            Generate Penalty 1
        end if
    else
        if (current time == end of middle stage) then
            for each bidder  $i$  in the auction  $j$  do
                Calculate  $LSS(i, j_{\text{middle}})$ 
            end for
            if ( $LSS(i, j_{\text{middle}}) > \rho_{\text{middle}}$ ) then
                Generate Penalty 2
            end if
        else
            if (current time == end of late stage) then
                for each bidder  $i$  in the auction  $j$  do
                    Calculate  $LSS(i, j_{\text{late}})$ 
                end for
                if ( $LSS(i, j_{\text{late}}) > \rho_{\text{late}}$ ) then
                    Generate Penalty 3
                end if
            end if
        else
            for each bidder  $i$  in the auction  $j$  do
                Calculate the resultant  $LSS(i, j)$ 
            end for
            if ( $LSS(i, j) < \rho_{\text{final}}$ ) then
                Exonerate the bidder  $i$  and generate Reward(Pardon)
            else
                if (Post-Filtering-Process(bidder  $i$ ) == True) then
                    Exonerate the bidder  $i$ 
                else
                    Generate Penalty 4
                end if
            end if
        end if
    end if
    Store the resultant  $LSS(i, j)$  into the database
End

```

There is a possibility that a legitimate bidder's behaviour may match with the bidding pattern of a shill bidder. Therefore, some innocent bidders might receive a high LSS which will classify them wrongly as a potential shill bidder. To avoid

this misclassification, Algorithm 2 presents a *Post-Filtering Process* (PFP) to exonerate innocent bidders after the auction ends. A bidder will be exonerated when any of the following circumstances occur:

- (a) The bidder has won the auction;
- (b) The bidder has a low α rating, or has a high overall LSS but no earlier LSSs;
- (c) The bidder's LSS_{early} is higher than the other LSSs;
- (d) The bidder has low average ratings compared to a shill bidder.

Algorithm 2 *Post-Filtering Process* (PFP)

INPUT: Auction j , Bidder i and the bid history of the bidder in past auctions

OUTPUT: *True* or *False*

```

Begin
if bidder  $i$  won the auction  $j$  then
    return True
else if bidder  $i$  is a one-time bidder in auction  $j$  then
    if (bidder  $i$ 's  $\alpha$  rating is low) || (bidder  $i$ 's overall LSS is high but no intermediate LSSs) then
        return True
    else if bidder  $i$ 's  $LSS_{early}$  is higher than other LSSs in auction  $j$  then
        return True
    else if bidder  $i$ 's mean values of average  $\beta, \delta, \epsilon, \zeta$  and  $\gamma$  ratings are lower than a predetermined threshold then
        return True
    Else
        return False
end if
End

```

The rationale for the *PFP* will be discussed in Section 5.

4.3 Flowchart of LSS Algorithm

Figure 4 represents the flowchart of the proposed LSS algorithm. A bidder's LSS represents the bidder's behaviour during an auction. We calculate the LSS value for each bidder in different stages (i.e., early, middle, late, final) of the online auction based upon the selection of bidding behaviours. The bidder, who receives LSS value higher than a certain threshold, will get appropriate penalties according to the different stages.

We have considered different threshold values (i.e., $\rho_{t_{early}}$, $\rho_{t_{middle}}$, $\rho_{t_{late}}$, $\rho_{t_{final}}$) in different stages (i.e., early, middle, late, and final) for determining a potential shill bidder. The threshold values used in this paper are: $\rho_{t_{early}} = 8$, $\rho_{t_{middle}} = 7$, $\rho_{t_{late}} = 7$, and $\rho_{t_{final}} = 6$. These weightings were obtained by experimenting with simulated and commercial auction data. The justification of these threshold values are as follows:

In the early stage, $\rho_{t_{early}}$ has the highest value because few bidders place bids in the beginning of an auction. Therefore, the bidders who place bids in the early stage may receive a high $\rho_{t_{early}}$. If we consider the value of $\rho_{t_{early}}$ low, then *Penalty 1* might be generated by the auctioneer for innocent bidders.

In the middle and late stage, both values of $\rho_{t_{middle}}$ and $\rho_{t_{late}}$ are considered the second highest. The reason is that, the shill bidder usually places a high number of bids in the middle stage and less bids during the late stage. So, if we consider the values of $\rho_{t_{middle}}$ and $\rho_{t_{late}}$ high as $\rho_{t_{early}}$, then a potential shill bidder might be successful.

The value of $\rho_{t_{final}}$ has received the lowest weight because when we analyse the experimental results for commercial auction data, we found that the overall LSS for a potential shill bidder is 6 or above. However, the overall LSS for innocent bidders is less than 6. Therefore, we consider the value of $\rho_{t_{final}}$ is 6 in our experiment results.

In the case of the final stage, a *Post-Filtering Process* (PFP) has been included to our LSS algorithm for exonerating innocent bidders (see in Section 4.2). A bidder who passes the *PFP* will be exonerated. Otherwise, LSS algorithm will cancel the auction.

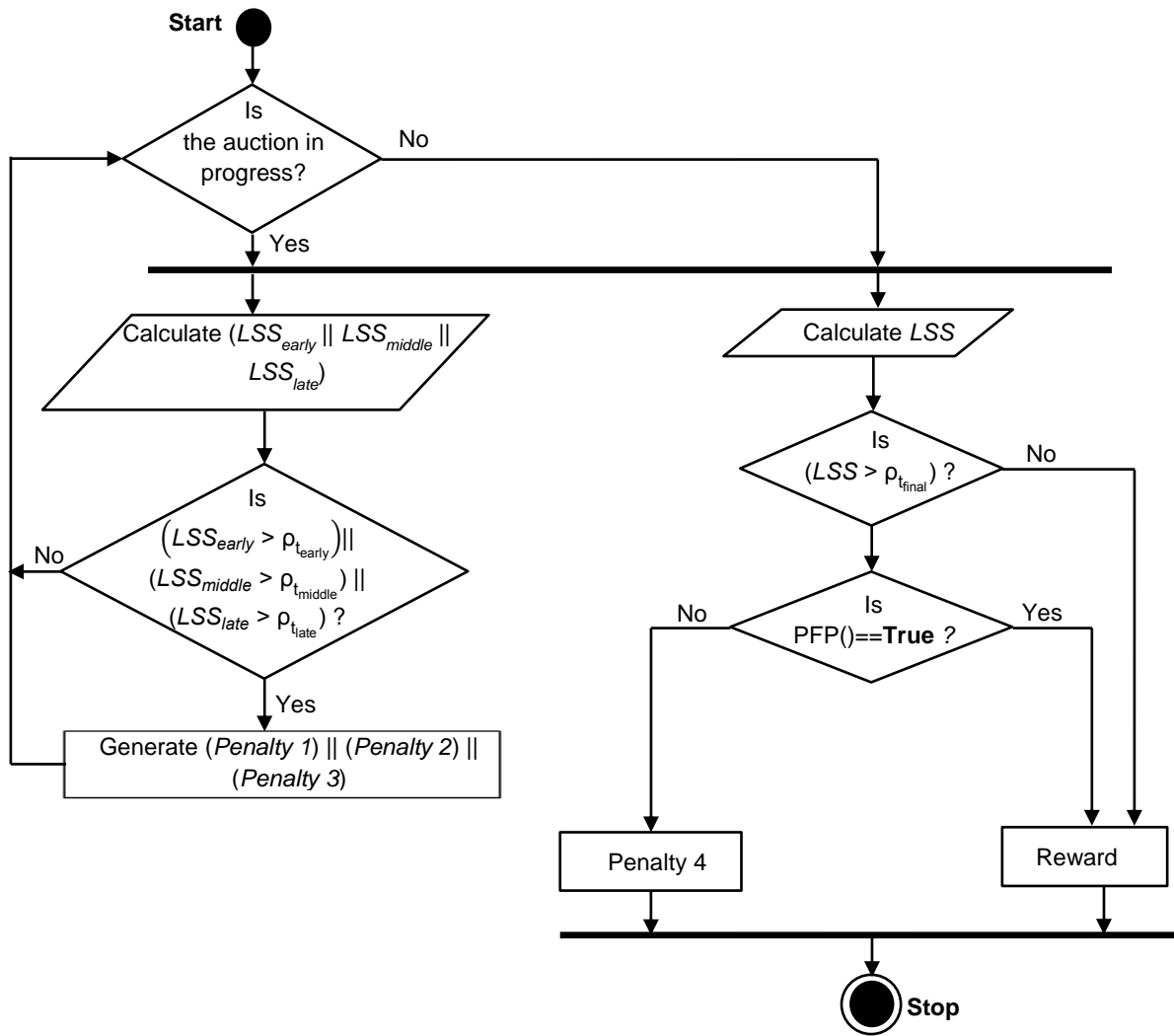


Figure 4: LSS algorithm flowchart

5 Experimental Results

This section describes how the LSS algorithm performed on synthetic and real auction data. For our experimental setup, we obtained two types of data: (i) the first was a series of simulated auction trials which were acquired from Trevathan and Read [24]; and (ii) the second was from commercial online auctions which were collected from Jank and Shmueli (Site 8). In addition, we developed *uAuction* [14] for conducting various types of testing on shill bidding detection using simulated and commercial auction data. Furthermore, we compare our proposed method with Trevathan and Read's Shill Score (SS) reputation system [26].

5.1 Simulated Auctions

Trevathan and Read [24] conducted a series of simulated auction trials to obtain auction data. The data contained thirty-nine auctions where each of these auctions was for a different item. All auctions were considered to be for one seller [24].

The auction proceedings involved twenty-six bidders. Each bidder was given a random amount of fake money at the beginning of the trials. A bidder was free to bid his/her true valuation in any auction provided that collectively the amount bid did not exceed the initial amount provided. When a bidder won an auction, the balance of his/her account was reduced by the value of the winning bid. A bidder's goal was to win while also trying to save his/her money.

The shill bidder's goal was to force a bidder into spending as much of his/her money as possible. Bidders were not informed that shill bidding was occurring. Furthermore, the shill bidder had no knowledge of how much money bidders had.

There were two types of tests performed: (1) with one shill bidder, and; (2) without shill bidding. The *LSS* algorithm was run on all of the tests to determine its effectiveness for determining shill bidding and the likelihood of it incriminating innocent bidders.

5.1.1 Auctions with Shill Bidding

The first test involved ten auctions and sixteen bidders, including one shill bidder (i.e., *Shelly*). We calculated the *LSS* for each bidder across the different stages of each auction. Table 4 shows the *LSS* for each bidder in the different stages of Auction ID: 9. The shill bidder (i.e., *Shelly*) is clearly identified as the bidder that has engaged in the most price inflating behaviour. The shill bidder is the only bidder who scores consistently on the behavioural characteristics that optimally inflate the price for the seller.

Table 4: *LSS* for each bidder in the different stages of Auction ID: 9 (with one shill bidder)

Bidder ID	Average β	Average δ	Average ϵ	Average ζ	γ	<i>LSS</i> _{early}	<i>LSS</i> _{middle}	<i>LSS</i> _{late}	<i>LSS</i>
Shelly	0.79	0.71	0.74	1.00	1.00	5.00	9.30	9.00	9.40
Marianne	0.39	0.74	1.00	0.84	1.00	6.10	7.90	7.80	8.70
Wayne	0.31	0.56	0.71	0.15	1.00	0.00	5.70	5.20	7.70
Soraya	0.29	0.43	0.73	0.23	1.00	0.00	5.00	5.60	7.70
Buzzcook	0.02	0.25	0.25	0.01	1.00	0.00	0.00	0.00	7.00
Townsville	0.06	0.25	0.14	0.00	1.00	0.00	0.00	0.00	6.60
Ness	0.43	0.56	0.25	0.48	0.00	3.20	7.10	6.90	0.00

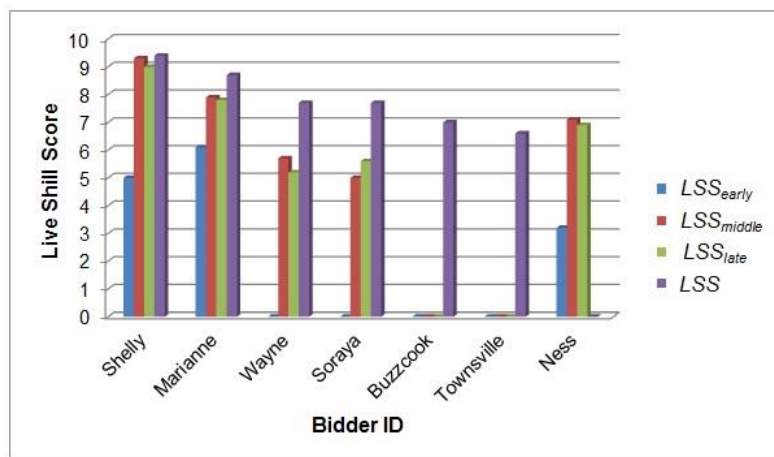


Figure 5: *LSS* for each bidder in Auction ID: 9 (with one shill bidder)

Figure 5 gives a graphical representation of the *LSS*s for each bidder in the Auction ID: 9. Figure 5 shows that *Shelly* has the highest *LSS* in the middle, late and final stages of the auction duration. This indicates that *Shelly* follows shill bidding patterns to artificially drive up the final price.

However, at some stages during an auction, a legitimate bidder's behaviour can emulate that of a shill bidder. This raises the possibility of classifying innocent bidding behaviour as shill bidding. Some innocent bidders place bids quickly after the current highest bid and by the minimal amount required till their true valuation. This behaviour increases the δ and ϵ ratings for the innocent bidders which will ultimately raise their *LSS*s in the different stages of the auction duration. For example, *Buzzcook* is an innocent bidder who gets a reasonably high *LSS*. To avoid the misclassification of innocent bidders as potential shill bidders, the *PFP* scrutinises the *LSS* calculations after the auction ends. To do so, the *PFP* compares the bidding patterns of innocent bidders with that of shill bidder and then exonerates innocent bidders.

Using the auction data we developed the *PFP* based on the following rationale:

The auction winner (i.e., *Ness*) can be exonerated according to the *PFP*. The reason is a shill bidder avoids winning any auction. Moreover, as bidders having a high overall *LSS*, but no intermediate *LSS*s (i.e., *LSS*_{early}, *LSS*_{middle}, *LSS*_{late}) can be exonerated as this indicates the bidder is a one-time or late bidder (e.g., *Townsville* and *Buzzcook*). In case of *Townsville*, he submitted bids at the end of the auction. So, his overall *LSS* is high with no earlier *LSS*s (see Figure 5). On the other hand, Figure 5 shows the *LSS* patterns of *Buzzcook* which is almost similar

to that of *Townsville*. As *Buzzcook* is a one-time bidder who submitted only one bid at the end of the auction, he can be exonerated.

Next we investigate the rest of the bidders who have *LSSs* higher than a predetermined threshold value (i.e., *Marianne*, *Wayne*, *Soraya*). The threshold value is determined using this example data with a known shill bidder. If other bidders match this threshold, this indicates a high likelihood that the bidder is behaving in price-inflating behaviour. We compare the *LSS* scores of innocent bidders with those of a shill bidder in the different stages of the past auctions. Figure 6 shows bidders with a high *LSS* who participated in other auctions including Auction ID: 9. *Shelly's* *LSSs* (see Figure 6(a)) do not match with those of *Marianne*, *Wayne*, and *Soraya* (see Figure 6(b), 6(c), and 6(d)).

Furthermore, *Shelly's* average ratings (see Figure 7(a)) show that the bidding patterns are different than the bidding patterns of *Marianne*, *Wayne*, and *Soraya* (see Figure 7(b), 7(c), and 7(d)). This indicates that *Marianne*, *Wayne* and *Soraya* are innocent bidders.

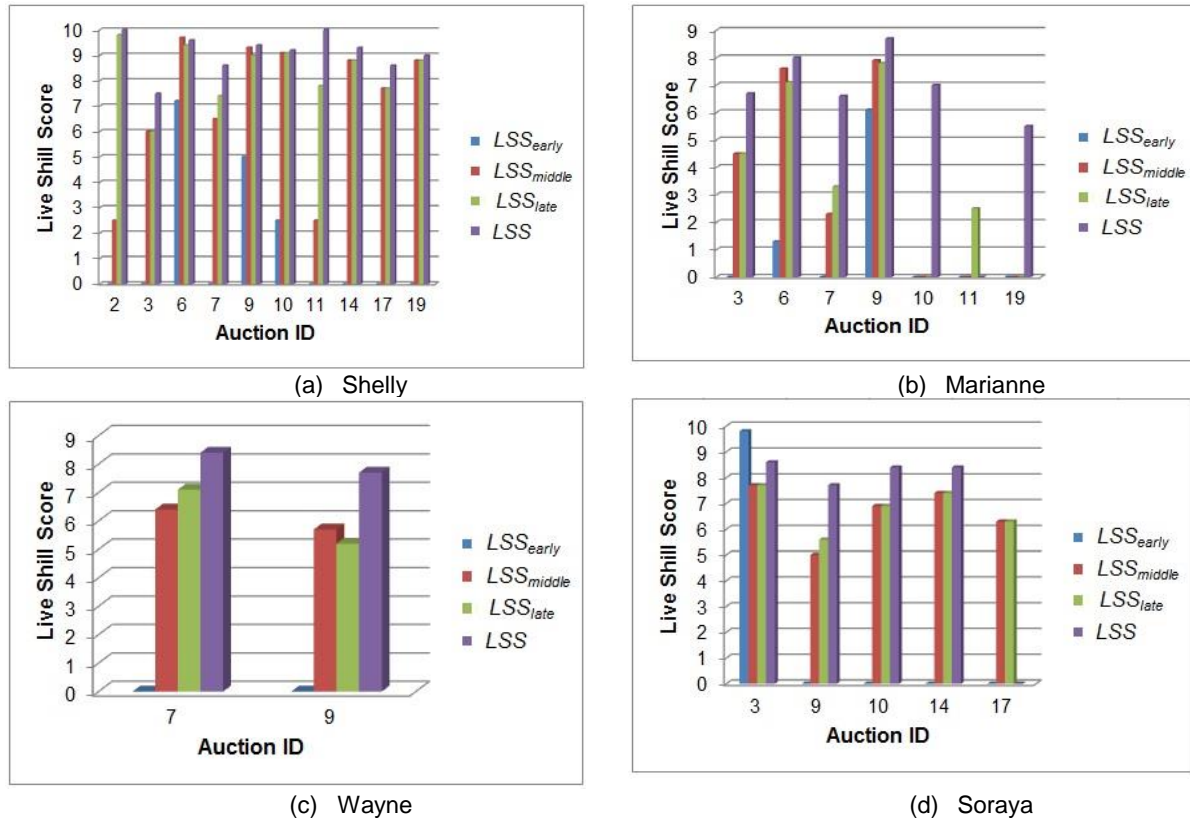


Figure 6: Bidders with high *LSS* who participated in other auction including Auction ID: 9

The *PFP* algorithm is used to further investigate a bidder's legitimacy. Table 5 shows the mean (average of average) of β , δ , ϵ , ζ , and γ ratings of each bidder in other auctions including Auction ID: 9. It also presents the bidding patterns of each bidder in Auction ID: 9. As the shill bidder's (i.e., *Shelly*) ratings are always high, so each of the rating is considered as 1 for a shill bidder. For example, *Shelly* got the bidding patterns of 11111. A bidder with a bidding rating is less than the corresponding rating of a shill bidder considered as 0, or 1 otherwise. Furthermore, *Wayne* receives the bidding patterns of 00101. The reason is the β , δ and ζ ratings of *Wayne* are less than those of *Shelly* but the ϵ rating of *Wayne* is higher than *Shelly's* ϵ rating, and *Wayne's* γ rating is equal to *Shelly's* γ rating.

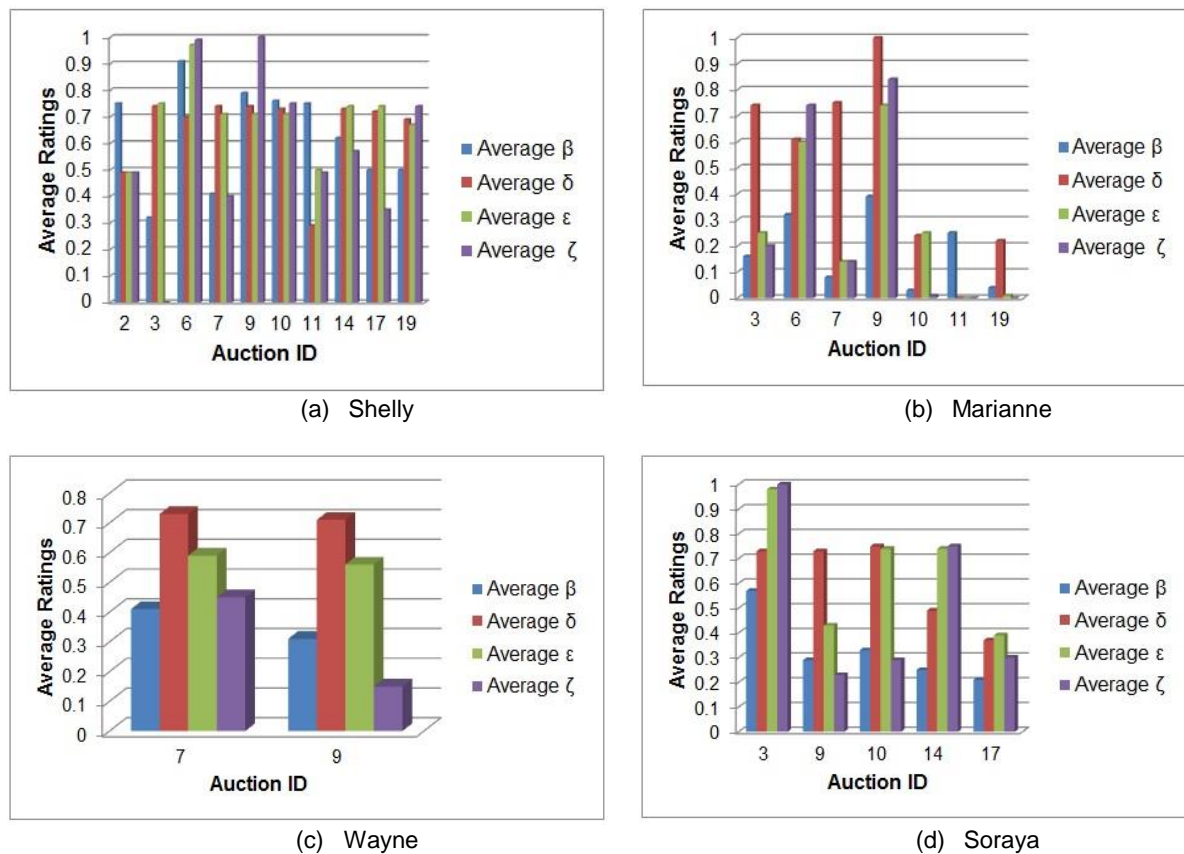


Figure 7: Average ratings of each bidder who participated in other auctions including Auction ID: 9

Table 5: Mean of ratings and bidding patterns for each bidder in past auctions including Auction ID: 9

Bidder ID	Mean β	Mean δ	Mean ϵ	Mean ζ	Mean γ	Bidding patterns	Status
Shelly	0.63	0.66	0.70	0.58	1.00	11111	Shill
Marianne	0.18	0.51	0.28	0.28	0.86	00000	Innocent
Wayne	0.36	0.58	0.72	0.30	1.00	00101	Innocent
Soraya	0.33	0.61	0.66	0.51	0.80	00000	Innocent
Buzzcook	0.01	0.13	0.13	0.01	0.50	00000	Innocent
Townsville	0.17	0.56	0.56	0.42	1.00	00001	Innocent
Ness	0.46	0.48	0.51	0.57	0.80	00000	Innocent

In addition, Table 5 shows the status of each participating bidder where a bidder having the bidding pattern of 11111 could be a potential shill bidder or any 0 in the bidding patterns of 11111 could be an innocent bidder. The reason is a bidder who has the bidding pattern of 11111 means that bidder's mean of β , δ , ϵ , ζ , and γ ratings are the highest among the other bidders. This is the strong sign of being a potential shill bidder.

A shill bidder tends to bid exclusively in auctions held by one particular seller [18]. As stated in Section 3.5, this behaviour is represented as *Affinity to a seller* or ' α rating'. A bidder having α rating of 1 indicates suspicious or 0 indicates otherwise. Suppose m auctions held by a seller. To calculate a bidder's α rating, we count the number of the seller's auctions, m^i , bidder i has participated in. Then, α rating of bidder i can be calculated as follows:

$$\alpha^i = (m^i - w^i) / m; \quad 0 \leq \alpha^i \leq 1 \quad (11)$$

where w^i denotes the number of auctions bidder i won. In general, α rating will be high for a shill bidder.

Figure 6(a) and 7(a) show that *Shelly* had participated in all 10 auctions and did not win any of these auctions. Moreover, she received the highest LSSs in all 10 auctions. Therefore, *Shelly*'s α rating is 1 which represents shill bidding behaviour. However, Figure 6(b) and 7(b) show that *Marianne* had participated in 7 auctions among 10 auctions and won an auction (i.e., Auction ID: 11). In case of *Marianne*, the α rating is 0.6. That means *Marianne* participated more than 50% of auctions held by the seller. However, if we look at Figure 6(b), we can see that *Marianne* had received high LSS in the Auction ID: 10 and 19 with no earlier LSSs. This indicates that *Marianne* had participated in those two

auctions as a late bidder and tried to win. *Wayne* and *Soraya* have α ratings of 0.2 and 0.4 respectively. This indicates that they only participated in 20%-40% of the total auctions. As such, they can be considered to be legitimate.

5.1.2 Auctions without Shill Bidding

The second test also involved ten auctions held by a seller. There were 18 bidders. However, unlike the previous test no intentional shill bidding behaviour has been engaged in. Table 6 shows the *LSS* scores of each bidder in Auction ID: 8 and Figure 8 represents Table 4 graphically. The test's purpose is to gauge regular bidding behaviour. Figure 8 shows that *Jc112425* and *Ness* have the highest *LSS*s.

Table 6: *LSS* for each bidder in the different stages of Auction ID: 8 (without shill bidder)

Bidder ID	Average β	Average δ	Average ϵ	Average ζ	γ	<i>LSS</i> _{early}	<i>LSS</i> _{middle}	<i>LSS</i> _{late}	<i>LSS</i>
Jc112425	0.49	0.75	0.63	1.00	1.00	5.00	7.80	7.80	8.80
Ness	0.61	0.74	0.53	0.75	1.00	2.50	7.90	7.90	8.80
Marie	0.12	0.68	0.66	0.11	1.00	0.00	5.10	5.10	7.20
Buzzcook	0.12	0.74	0.69	0.01	1.00	0.00	5.10	5.10	7.10
Brenda	0.12	0.74	0.37	0.00	1.00	0.00	4.10	4.10	6.40
Marianne	0.27	0.38	0.44	0.14	1.00	0.00	4.10	4.10	6.30
Bear	0.17	0.25	0.09	0.26	0.00	0.00	3.80	3.80	0.00

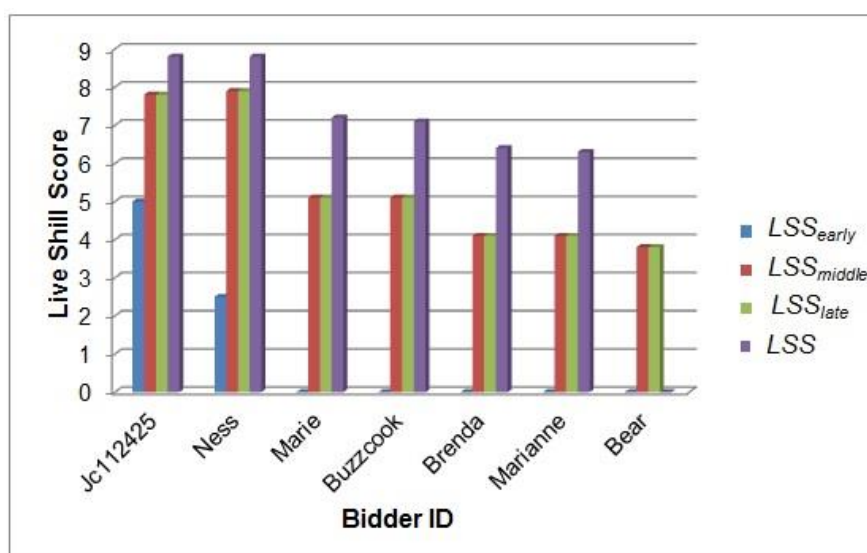


Figure 8: *LSS* for each bidder in Auction ID: 8 (without shill bidder)

Using the *PFP*, the auction winner (e. g., *Bear*) can be eliminated. We can perform *PFP* process for investigating bidding behaviours of all participating bidders in Auction ID: 8 except *Bear*. Figure 9 shows the *LSS*s of all bidders except the winner of the auction (e. g., *Bear*). It can be seen that *LSS* patterns of each bidder do not match that of a shill bidder like *Shelly* in Figure 6(a). Similarly, Figure 10 represents the average ratings of each bidder in Auction ID: 8 except the winner (e. g., *Bear*). It also shows that the bidding patterns of each bidder do not match with the bidding patterns of *Shelly* (see Figure 7(a)).

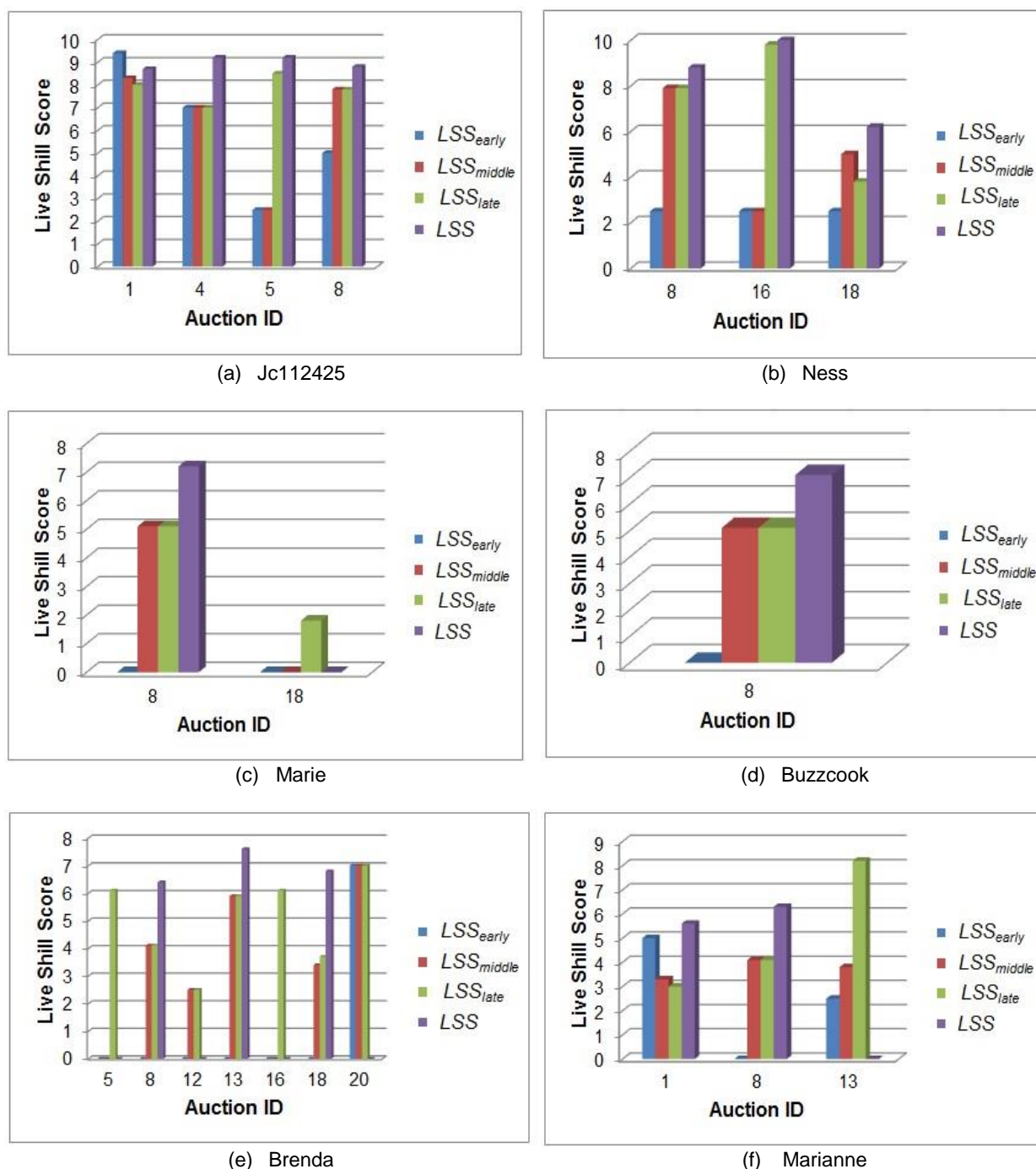


Figure 9: LSS of each bidder who participated in other auction including Auction ID: 8

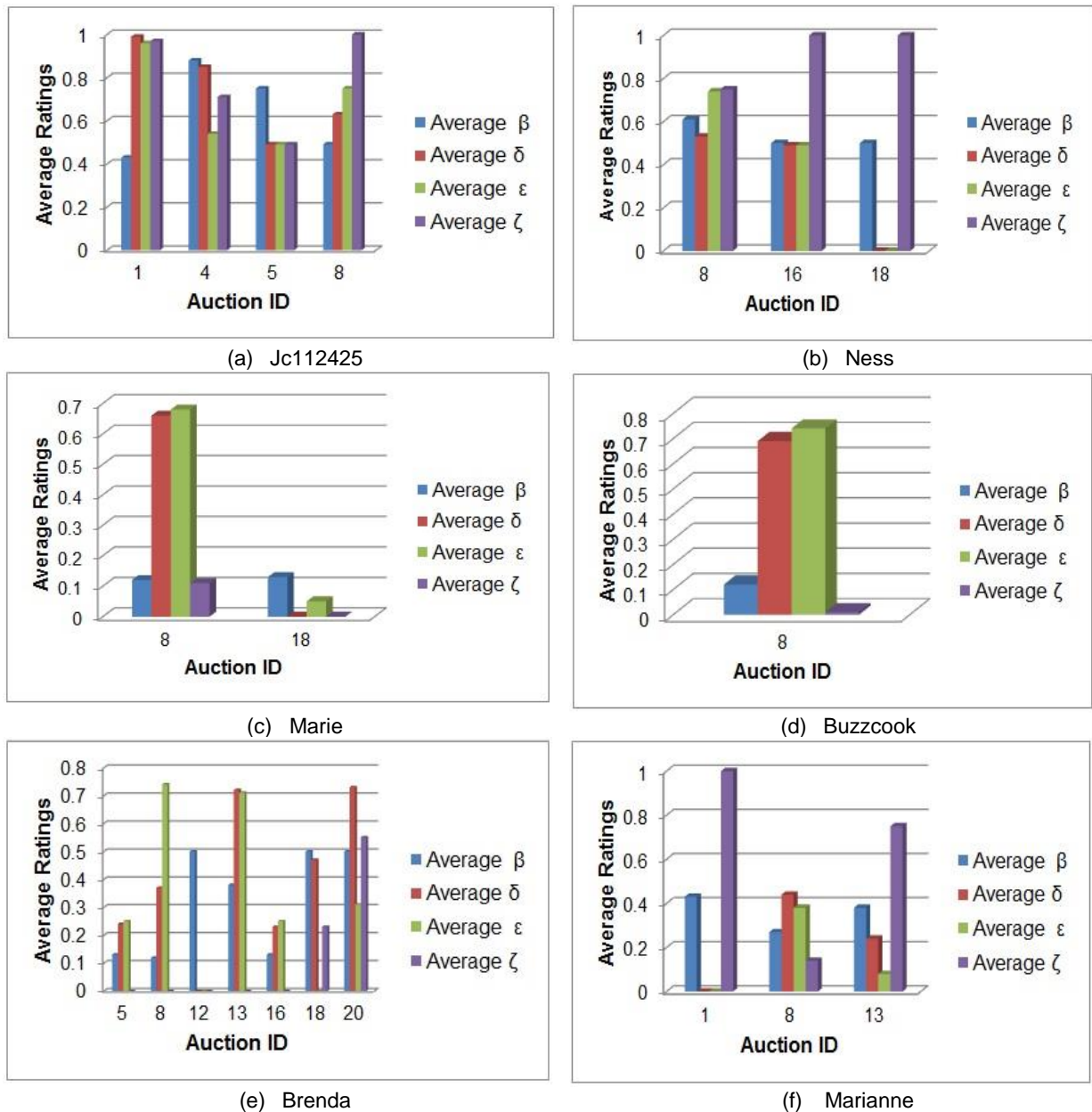


Figure 10: Average ratings of each bidder who participated in other auctions including Auction ID: 8

Table 7: Mean of ratings and bidding patterns for each bidder in past auctions including Auction ID: 8

Bidder ID	Mean β	Mean δ	Mean ϵ	Mean ζ	Mean γ	Bidding patterns	Status
Jc112425	0.54	0.61	0.63	0.79	1.00	00011	Innocent
Ness	0.54	0.34	0.41	0.92	1.00	00011	Innocent
Marie	0.13	0.33	0.37	0.10	0.50	00000	Innocent
Buzzcook	0.12	0.69	0.74	0.01	1.00	01101	Innocent
Brenda	0.32	0.24	0.32	0.11	0.43	00000	Innocent
Marianne	0.36	0.23	0.15	0.63	0.67	00010	Innocent
Bear	0.33	0.43	0.53	0.34	0.67	00000	Innocent

Table 7 presents mean (average of average) of ratings and bidding patterns for all bidders in Auction ID: 8. The mean ratings of all bidders in Auction ID: 8 do not match the bidding patterns of *Shelly* (see Table 5).

Moreover, when we consider the α ratings of each bidder, it can be seen that all the participating bidders' submitted bids in less than 50% of the auctions held by the same seller. For example, *Brenda* submitted bids in 7 auctions and she won 4 auctions among the 10 auctions (see Figure 9(e)). So, her α rating is 0.30 which is low. This indicates *Brenda* is a legitimate bidder

5.2 Commercial Auctions

We consider an auction listing of Palm Pilot PDAs auctioned on eBay (Site 8). This listing consists of three high demand items on eBay such as Cartier wristwatches, Palm Pilot M515 PDAs, and Xbox game consoles. We have selected the PDA because it was popular to a large number of bidders and bids. The Palm Pilot PDA auction data was collected over a period of two months during 2003.

Table 8: Statistical information of Palm PDAs

Features	Palm PDAs
Total no of auctions	149
Total no of bidders	1,029
Total no of sellers	71
Total no of bids	3,166
Average of total bidders	7
Average of total bids	21
Average winning price	\$229.04
1-time bidders	75.2%
2-time bidders	14.3%

Table 8 provides statistical information about the Palm PDAs dataset. In Table 8, *1-time bidder* refers to the bidder who participated in only one auction, and *2-time bidder* means that the bidder participated in only two auctions.

Table 9 shows the *LSS* scores of each bidder in the *Auction ID: 3020532816* posted by *syschannel*. Figure 11 shows the graphical representation of Table 9. Figure 11 shows that *szukaih* has received the highest *LSSs* across the different stages of *Auction ID: 3020532816* among 20 bidders. Using the *PPF*, the auction winner (i.e., *grafonalamo*) can be eliminated. However, the rest of the bidders who have a *LSS* less than a predetermined threshold (i.e., p_{final}) can be exonerated.

Table 9: *LSSs* for Auction ID: 3020532816

Bidder ID	Average β	Average δ	Average ϵ	Average ζ	γ	<i>LSS</i> _{early}	<i>LSS</i> _{middle}	<i>LSS</i> _{late}	<i>LSS</i>
szukaih	0.54	0.99	0.88	1.00	1.00	9.80	7.60	5.60	6.50
zebedin	0.12	0.50	0.45	0.07	1.00	0.00	0.00	4.20	5.60
hawkswimmers	0.11	0.50	0.42	0.04	1.00	0.00	0.00	3.90	5.30
msh39	0.07	0.95	0.70	0.82	1.00	3.50	3.80	3.80	5.30
kc10	0.07	1.00	0.28	0.70	1.00	2.10	3.10	3.40	5.10
gaylanm	0.10	0.71	0.57	0.20	1.00	0.00	3.60	3.30	5.00
tfalcrazd	0.06	0.48	0.41	0.08	1.00	0.00	0.00	3.30	5.00
ward42556	0.03	0.74	0.69	0.20	1.00	0.00	3.10	3.10	4.90
loc820	0.02	0.25	0.23	0.00	1.00	0.00	0.00	0.00	4.80
mongo6104	0.05	0.68	0.48	0.14	1.00	0.00	2.00	3.00	4.80
fzuluaga	0.03	0.72	0.64	0.13	1.00	0.00	2.80	2.90	4.70
meritcc	0.00	0.25	0.00	0.00	1.00	0.00	0.00	0.00	4.60
biged091371	0.01	0.25	0.22	0.00	1.00	0.00	0.00	0.00	4.50
dacsmilles	0.01	0.22	0.24	0.00	1.00	0.00	0.00	0.00	4.40
ev530i	0.02	0.47	0.36	0.05	1.00	0.00	0.00	2.40	4.40
be4real0	0.00	0.49	0.30	0.02	1.00	0.00	0.00	2.40	4.40
samtemple	0.03	0.74	0.21	0.14	1.00	0.00	1.80	2.40	4.30
gsrescuedog	0.02	0.49	0.24	0.08	1.00	0.00	0.00	2.30	4.30
adprice14	0.02	0.47	0.00	0.03	1.00	0.00	0.00	1.40	3.70
depietsch	0.03	0.00	0.21	0.28	1.00	0.00	0.80	1.30	3.50
grafonalamo	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

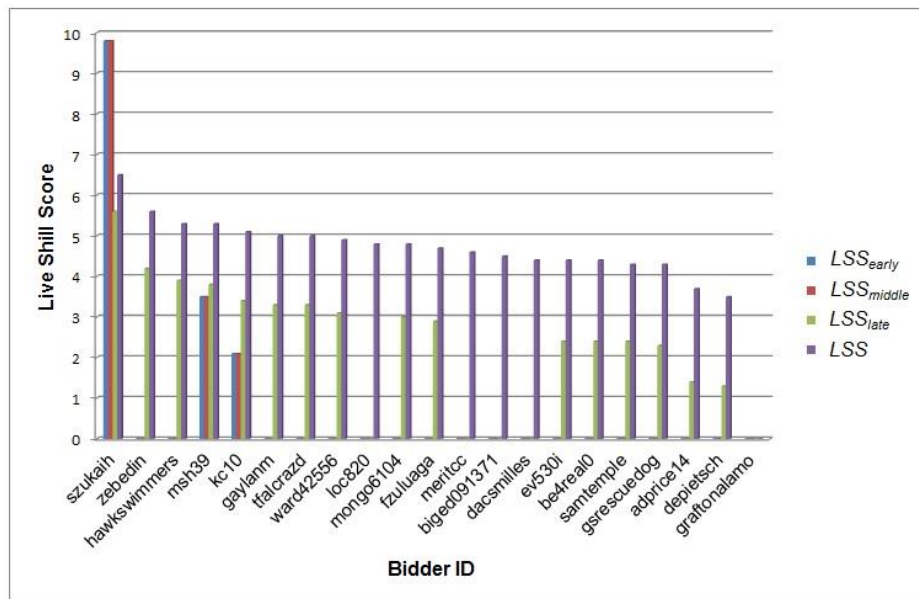


Figure 11: LSS for each bidder in Auction ID: 3020532816

When we check the LSS scores of *szukaih* in other auctions, it can be found that the bidder submitted bids only in this auction (i.e., Auction ID: 3020532816). This may indicate that the bidder *szukaih* is an innocent bidder. This means that the auction (i.e., Auction ID: 3020532816) does not show any shill bidding activities.

However, we found that the bidder *kc10* has the highest α rating among all the bidders in this dataset. The reason is that *kc10* participated in 21 auctions out of the 40 auctions held by the same seller (i.e., *syschannel*) which is almost 53% of the auctions run by the same seller (see Figure 12). Figure 13 shows the LSS s of *kc10* in all auctions where *kc10* participated in. It shows that *kc10*'s LSS s are increasing in all auctions (i.e., LSS_{middle} is higher than LSS_{early} , LSS_{late} is higher than LSS_{middle} , and LSS is higher than LSS_{late}). Furthermore, Figure 14 presents the average ratings of *kc10* participated in auctions held by the same seller (i.e., *syschannel*). Figure 14 also shows that the average β ratings of *kc10* are always low for all auctions as he/she is one-off bidder. However, the other ratings (i.e., δ , ϵ ratings) are high in all the auctions. More specially, the average ζ rating of *kc10* always remains high. This indicates that the bidder *kc10* submitted bids either in the early or middle stage of the auction duration. This may indicate suspicious bidding activity.

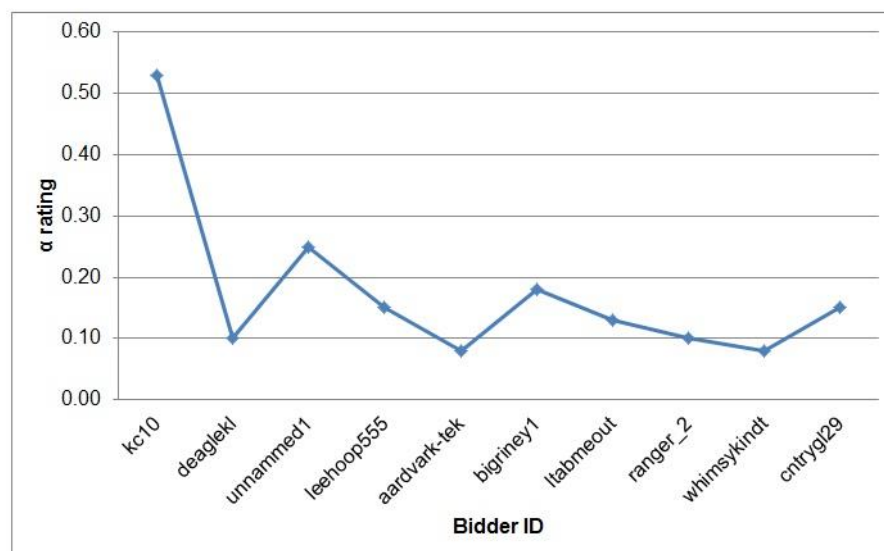


Figure 12: Top ten bidders' α rating who participated in the auctions held by *syschannel*

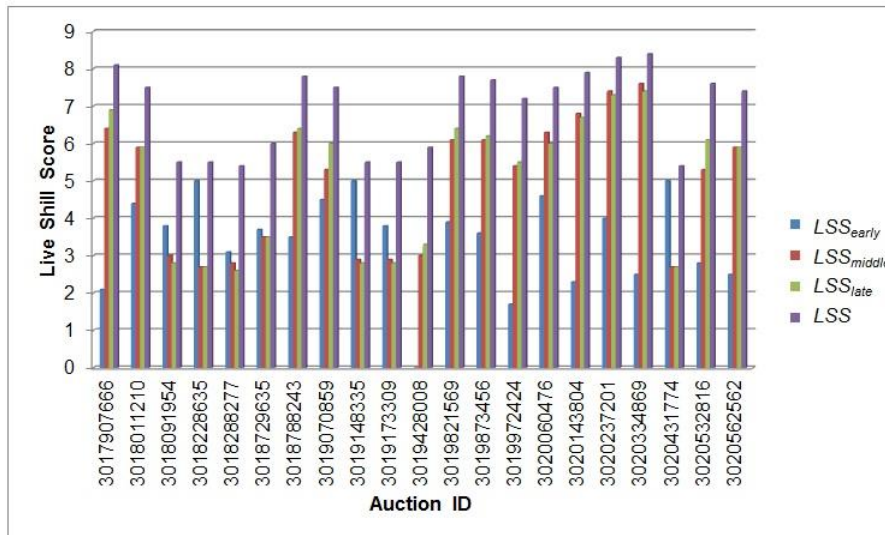


Figure 13: LSS scores for $kc10$ in all participated auctions

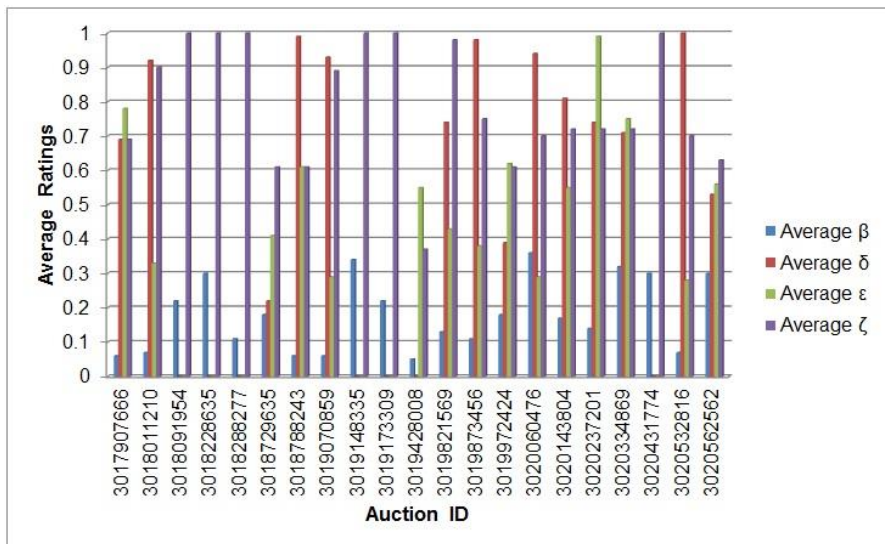


Figure 14: Average ratings for $kc10$ participated in all auctions

In addition, we have applied our LSS algorithm on dataset collected in 2016 from an Honours project involving a software tool that automates auction data collection [6]. The auction dataset was obtained from July 2016 to October 2016. Table 10 shows the LSS scores of each bidder in the Auction ID: 10825 posted by $m^{**}d$. Using the PFP , the auction winner (e. g., $m^{**}n$) can be eliminated. Except the winner of the auction, all bidders have the highest LSS s at the end of the final stage. When we further investigate using our PFP algorithm, we found that $b^{***}b$ has the highest α rating among all the bidders in this dataset. Because the bidder $b^{***}b$ participated in all auctions hosted by the same seller (e. g., $m^{**}d$) and never won any auction. This indicates a suspicious bidding behaviour. Whereas, all other bidders (e. g., $n^{***}t$, $1^{***}t$, $y^{***}y$, $h^{***}a$) submitted bids only in this auction (e. g., Auction ID: 10825) which indicates those bidders are innocent.

Table 10: LSS s for Auction ID: 10825

Bidder ID	Average β	Average δ	Average ϵ	Average ζ	γ	LSS_{early}	LSS_{middle}	LSS_{late}	LSS
$n^{***}t$	0.23	0.50	0.44	0.48	1.00	0.00	0.00	8.10	8.97
$b^{***}b$	0.12	0.49	0.46	0.48	1.00	0.00	0.00	7.60	8.43
$1^{***}t$	0.50	0.75	0.81	0.94	1.00	7.00	8.40	7.70	8.03
$y^{***}y$	0.19	0.62	0.70	0.11	1.00	0.00	5.90	5.30	6.89
$h^{***}a$	0.03	0.71	0.73	0.00	1.00	0.00	5.00	4.90	6.82
$m^{**}n$	0.16	0.49	0.34	0.03	0.00	0.00	5.30	4.80	0.00

5.3 Comparison with the State of the Art

We compared our run-time *LSS* algorithm with Xu et al.'s real-time shill detection proposal [32]. Xu et al. [32] implemented software bidding agent that is designed to simulate the properties of model-based bidding strategies. The authors used simulated data to demonstrate the efficiency and effectiveness of their proposed approach [32]. They used a flat scoring strategy in which a constant value of 2 is added to the accumulative shilling score of a monitored bidder each time a shill bidding behaviour is detected [32]. The authors [32] ran their real-time model checking program for three bidders (i.e., Bidder 1, Bidder 2, Bidder 3) and record the shilling scores of all three bidders over the auction time which is 48 hours (see Figure 15). They set the threshold value for warning as 15 and for cancelling an auction as 25 [32]. The yellow-filled circle in Figure 15 shows Bidder 1 and 2 receive a warning message around 18 and 35 hours after the auction starts, respectively [32]. In addition, Figure 15 also shows the auction will be cancelled around 2 hours before the auction ends which is denoted by the circle on the line with shilling score of 25 [32]. The authors [32] consider Bidder 3 is a normal bidder as his/her shilling score never exceeds 15.

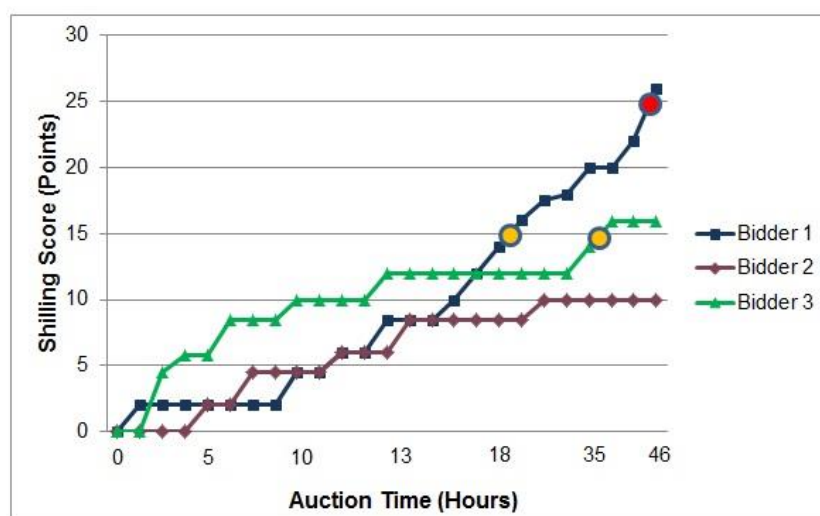
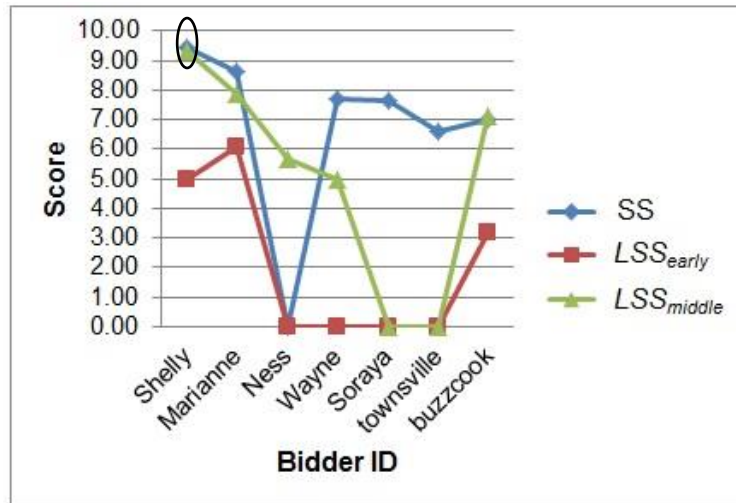


Figure 15: Comparison of shilling scores for three bidders [32]

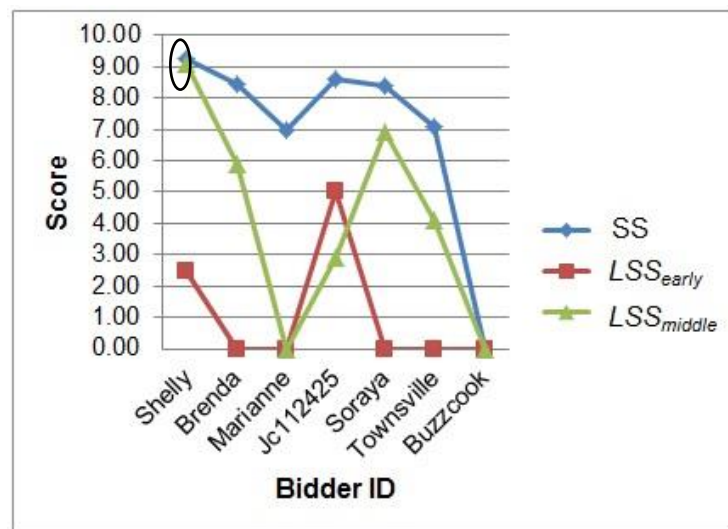
Compared to Xu et al. [32], our *LSS* algorithm provides more efficient and effective performance. Firstly, we tested our proposed run-time shill detection approach on simulated and commercial auction data. Secondly, we consider past bidding history when calculating *LSS* for each bidder. The reason is that it is unrealistic to point a bidder as a shill bidder only on his/her bidding behaviour in a running auction. For example, Xu et al. [32] identified a bidder (e.g., Bidder 1 in Figure 15) as a shill bidder based on his/her bidding behaviours in a live auction. However, it may happen that the bidder placed bids aggressively during the early and middle stages of a particular auction duration in order to win the auction and reached his/her true valuation during the final stage of the auction duration. This situation does not mean that the bidder is a shill bidder. Therefore, we consider each bidder's bidding history in past auctions for identifying shill bidding activities in a running auction which ensures the efficiency and effectiveness of our *LSS* algorithm.

Furthermore, we compared our *LSS* algorithm with *Shill Score* (SS) reputation system proposed by Trevathan and Read [26]. Although Trevathan and Read's approach [26] was able to detect a potential shill bidder, their *Shill Score* algorithm does not work in run-time. Therefore, we chose to run our *LSS* algorithm on the same auction data to see how it would execute in detecting a shill bidder in the running auction. Figure 16 presents the comparison of our *LSS* algorithm with Trevathan and Read approach [26]. The black-lined ellipses of Figure 16(a) and 16(b) tell us what the average difference is between the SS and the *LSS_{middle}* across the two auctions (e.g., Auction ID: 9 and 10).

Figure 16 provides a comparison of the various stages of the *LSS* algorithm against the *Shill Score* (SS). On average the *LSS_{middle}* is within 0.34 distance of the resultant SS. This indicates that the *LSS* is able to achieve roughly the same outcome as the SS, but is able to flag potential shill bidding at the 80% duration mark in most cases. This outcome is clearly quicker than the SS and has a similar degree of accuracy.



(a)



(b)

Figure 16: Comparison our LSS algorithm with Trevathan and Read method [26]: (a) Auction ID: 9, (b) Auction ID: 10

5.4 The Rationale for the Selection of 25-80-95% Cuts

As we mentioned earlier (see Section 3.2) that Xu et al. [32] introduced a strategy where a live auction duration divided into three stages, namely: (i) early stage (i.e., the first quarter of the auction duration); (ii) middle stage defined as $[0.25T, 0.9T]$; and (iii) final stage (i.e., the last 10% of the auction time). Initially we chose to use the same time periods that were proposed by Xu et al. [32] for the purposes of uniformity across the research area (except our model does introduce an additional stage (i.e., the final stage defined as the last 5% of the auction duration). We have altered different cuts of auction duration (e.g., 20-75-97%, 30-85-90%, etc.) and tested on simulated and commercial auction data. Thus far, there does not seem to be a large degree of difference in the final results. The reason is small changes to the time period thresholds (e.g., 25-80-95%, 20-75-97%, 30-85-90%, etc.) do not have a large impact on the results.

Besides, we have found minor differences between 25-80-95% and 20-75-97% cuts on the experiment results where we compared the various LSS s across the time periods of auction duration. Therefore, we are largely indifferent about which of these two cuts to use. We have selected 25-80-95% cuts for testing our LSS algorithm. However, 30-85-90% showed more significant difference and made our LSS algorithm less accurate. We suggest the rationale for this is that leaving the LSS_{middle} to 85% is too late for detecting a shill bidder as this allows the shill to impact much more of the auction duration than the earlier LSS_{middle} cuts. Furthermore, 90% is a way too early to calculate the LSS_{late} and does not produce the desired effect of exonerating late bidders. Finally, 85%-90% is too short an interval between the middle and late stages.

6 Conclusions and Future Work

This paper proposed a run-time detection method for combating shill bidding in online auctions. We identified the most common/popular shill bidding characteristics based on a review of significant shill bidding detection literature. Using these characteristics, we devised the *Live Shill Score* (LSS) algorithm. The LSS algorithm aims to not only detect potential shill bidders but also to react to the shill bidding behaviour as early as possible during an auction. Therefore, this approach potentially reduces the monetary loss of shill bidding victims. The LSS algorithm uses a *Post-Filtering Process* (PFP) to avoid misclassification of innocent bidders. We also discussed potential penalties and the impact these penalties have on the auction participants.

We implemented the LSS algorithm and undertook some experimentation to determine how effective the algorithm is in detecting shill bidders. The first tests involved simulated auction data (i.e., real users with fake auctions). These tests contained data for known shill bidding and also auctions with no shill bidding. The results show that the LSS was able to highlight a potential shill bidder during an auction. Furthermore, the PFP was then able to further exonerate innocent bidders. When the LSS was run on auctions without shill bidding, it was able to accurately determine that all bidders were legitimate.

Further testing involved using commercial auction data collected from eBay. It was not known whether this data contained any shill bidders. The results for the LSS algorithm were consistent with the simulated auction data. The algorithm was able to identify a highly suspicious bidder who participated in many auctions, but never won any of them. This particular bidder always bid in the early and middle stages of an auction. When the affinity for the seller characteristic is taken into account, this bidder was found to have only participated in auctions held exclusively by the same seller. Such behaviour is clearly suspicious. Furthermore, we compared our LSS algorithm with *Shill Score* (SS) reputation system proposed by Trevathan and Read [26] and found that LSS algorithm was able to identify potential shill bidding at the 80% duration mark of a running auction in most cases. This indicates that our LSS algorithm is 20% quicker than the SS.

The functioning of our algorithm is public. The idea is that if a shill bidder continues to engage in price inflating behaviour, then s/he will be detected by the LSS. If the shill attempts to alter his/her behaviour to avoid detection, then the shill bidder is deviating from the seven (7) most commonly accepted shill bidding strategies. As a result, the shill bidder will start acting more like a regular bidder and be less effective at inflating the price. Therefore, if *smart attacker* changes his/her behaviour in order to keep his/her score low, then s/he is essentially not having much impact on inflating the price. The security of an algorithm should not lie in its operation being secret. Instead the algorithm should be publicly available so that it can be scrutinised. Only once it is shown to be resilient, the community can have any confidence in the algorithm (i.e., Kerckhoff's Principle).

Future work involves undertaking more comprehensive testing of the LSS algorithm using a wider range of auction data. Additionally, we will incorporate the history of past auctions into the calculation of the LSS (i.e., affinity to a seller). It is hoped that this will serve to reduce the potential for innocent bidders to incur penalties during the early and middle stages of an auction. We would also like to further develop our auction system for detecting and preventing multiple shill bidders from performing price inflating attempts with collusive groups in running auctions. When extending our LSS algorithm to multiple auctions, setting all the rating to 0 is designed to exonerate the winning bidder - as a shill bidder's goal is to lose. There are only a certain number of times a shill bidder can lose an auction for a specific until it becomes unprofitable to engage in further shill bidding. This is because each time the seller lists an auction, the auctioneer charges a listing fee. If a shill bidder continually wins the auction, then the seller must absorb the listing fees and over time, the expected profit from shill bidding is negated. Further work also involves looking into the impact of imposing penalties on bidders and auctions.

Acknowledgments

This work is supported in part by Griffith University International Postgraduate Research Scholarship (GUIPRS) program.

Websites List

Site 1: Power Sellers Unite
<http://www.powersellersunite.com/auctionsitewatch.php>

Site 2: TradeMe
<http://www.trademe.co.nz/trust-safety/2013/5/23/autoco-shill-bidding-fine>

Site 3: eBay: Number of users 2017
<https://www.statista.com/statistics/242235/number-of-ebays-total-active-users/>

Site 4: Venture Beat

<https://venturebeat.com/2016/07/20/ebays-q2-2016-sees-2-2-billion-in-revenue-passes-1-billion-live-listings/>

Site 5: Internet Crime Complaint Center: Internet Crime Report in 2014

https://pdf.ic3.gov/2014_IC3Report.pdf

Site 6: Internet Crime Complaint Center: Internet Crime Report in 2015

https://pdf.ic3.gov/2015_IC3Report.pdf

Site 7: The New York Times

<http://www.nytimes.com/2001/03/09/business/3-men-are-charged-with-fraud-in-1100-art-auctions-on-ebay.html>

Site 8: Modeling Online Auctions

<http://www.modelingonlineauctions.com/dataset>

References

- [1] B. Bhargava, M. Jenamani and Y. Zhong, Counteracting shill bidding in online English auction, *International Journal of Cooperative Information Systems*, vol. 14, no. 2-3, pp. 245-263, 2005.
- [2] F. Dong, S. M. Shatz and H. Xu, Inference of online auction shills using dempster-shafer theory, in *Proceedings of International Conference on Information Technology: New Generations*, USA, 2009, pp. 908-914.
- [3] F. Dong, S. M. Shatz and H. Xu, Combating online in-auction fraud: Clues, techniques and challenges, *Computer Science Review*, vol. 3, no. 4, pp. 245-258, 2009.
- [4] F. Dong, S. M. Shatz and H. Xu, Reasoning under uncertainty for shill detection in online auctions using dempster-shafer theory, *International Journal of Software Engineering and Knowledge Engineering*, vol. 20, no. 7, pp. 943-973, 2010.
- [5] F. Dong, S. M. Shatz, H. Xu, and D. Majumdar, Price comparison: A reliable approach to identifying shill bidding in online auctions?, *Electronic Commerce Research and Applications*, vol. 11, no. 2, pp. 171-179, 2012.
- [6] N. Fisher, J. Trevathan and H. Gray, *Detecting Shill Bidding Utilising eBay's 30-Day Bid Summary*, Griffith University, Australia, Honours Project Report, 2016.
- [7] B. J. Ford, H. Xu and I. Valova, Identifying suspicious bidders utilizing hierarchical clustering and decision trees, in *Proceedings of the 12th International Conference on Artificial Intelligence (ICAI10)*, USA, 2010, pp. 195-201.
- [8] B. J. 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, 2012.
- [9] A. Goel, H. Xu and S. M. Shatz, A multi-state bayesian network for shill verification in online auctions, in *Proceedings of International Conference on Software Engineering and Knowledge Engineering (SEKE)*, USA, 2010, pp. 279-285.
- [10] R. J. Kauffman and C. A. Wood, Irregular bidding from opportunism: An exploration of shilling in online auctions, *Information Systems Research*, vol. 10, pp. 1-36, 2007.
- [11] R. J. Kauffman and C. A. Wood, Running up the bid: Detecting, predicting, and preventing reserve price shilling in online auctions, in the *Proceeding of the 5th International Conference on Electronic Commerce*, USA, 2003, pp. 259-265.
- [12] G. Kosmopoulou and D. G. De Silva, The effect of shill bidding upon prices: Experimental evidence, *International Journal of Industrial Organization*, vol. 25, no. 2, pp. 291-313, 2007.
- [13] B. Lei, H. Zhang, H. Chen, L. Liu, and D. Wang, A k-means clustering based algorithm for shill bidding recognition in online auction, in *Proceedings of the 24th Chinese Control and Decision Conference (CCDC)*, China, 2012, pp. 939-943.
- [14] N. Majadi, J. Trevathan and N. Bergmann, uAuction: Analysis, design and implementation of a secure online auction system, in *Proceeding of the 14th International Conference on Dependable, Autonomic and Secure Computing (DASC-2016)*, New Zealand, 2016, pp. 278-285.
- [15] N. Majadi, J. Trevathan, H. Gray, V. E. Castro, and N. Bergmann, Real-time detection of shill bidding in online auctions: A literature review, *Journal of Computer Science Review*, vol. 25, pp. 1-18, 2017.
- [16] K. Mamun and S. Sadaoui, Combating shill bidding in online auctions, in *Proceedings of International Conference on Information Society (i-Society)*, Canada, 2013, pp. 170-176.
- [17] K. Mamun, Combating shill bidding in real time: Prevention, detection and response, *Computer and Information Science*, vol. 8, no. 2, pp. 24-36, 2015.
- [18] R. Patel, H. Xu and A. Goel, Real-time trust management in agent based online auction systems, in *Proceedings of International Conference on Software Engineering and Knowledge Engineering (SEKE)*, USA, 2007, pp. 244-50.
- [19] S. Sadaoui, X. Wang and D. Qi, A real-time monitoring framework for online auctions frauds, *Current Approaches in Applied Artificial Intelligence, Lecture Notes in Computer Science*, vol. 9101, pp. 97-108, 2015.
- [20] S. Sadaoui and X. Wang, A dynamic stage-based fraud monitoring framework of multiple live auctions, *Applied Intelligence*, vol. 46, no. 1, pp. 1-17, 2016.
- [21] S. Stoner and N. Pervushina, *Reduced to Skin and Bones Revisited: An Updated Analysis of Tiger Seizures from 12 Tiger Range Countries (2000-2012)*, Kuala Lumpur: TRAFFIC, 2013.

- [22] H. S. Shah, N. R. Joshi, A. Sureka, and P. R. Wurman, Mining eBay: Bidding strategies and shill detection, in Proceedings of the 4th International Workshop on Mining Web Data for Discovering Usage Patterns and Profiles, Canada, 2002, pp. 17-34.
- [23] J. Sher, Optimal shill bidding in the VCG mechanism, *Economic Theory*, vol. 50, no. 2, pp. 341-387, 2012.
- [24] J. Trevathan and W. Read, RAS: A system for supporting research in online auctions, *ACM Crossroads*. vol. 12, no. 4, pp. 23-30, 2006.
- [25] J. Trevathan and W. Read, Undesirable and fraudulent behaviour in online auctions, in Proceedings of International Conference on Security and Cryptography, Portugal, 2006, pp. 450-458.
- [26] J. Trevathan and W. Read, Detecting shill bidding in online english auctions, in *Handbook of Research on Social and Organizational Liabilities in Information Security*, (M.Cupta and R. Sharman, Eds.). USA: IGI Press, 2005, pp. 446-470.
- [27] J. Trevathan and W. Read, Investigating shill bidding behaviour involving colluding bidders, *Journal of Computers*, vol. 2, no. 10, pp. 63-75, 2007.
- [28] S. Tsang, G. Dobbie and Y. S. Koh, Evaluating fraud detection algorithms using an auction data generator, in Proceedings of 12th International Conference on Data Mining Workshops, Belgium, 2012, pp. 332-339.
- [29] S. Tsang, Y. S. Koh, G. Dobbie and S. Alam, Detecting online auction shilling frauds using supervised learning, *Expert Systems with Applications*, vol. 41, no. 6, pp. 3027-3040, 2014.
- [30] H. Xu and Y. T. Cheng, Model checking bidding behaviors in internet concurrent auctions, *International Journal of Computer Systems Science & Engineering*, vol. 22, no. 4, pp. 179-191, 2007.
- [31] H. Xu, S.M. Shatz and C.K. Bates, A framework for agent-based trust management in online auctions, in Proceedings of the 5th International Conference on Information Technology: New Generations (ITNG), USA, 2008, pp. 149-155.
- [32] H. Xu, C. K. Bates and S. M. Shatz, Real-time model checking for shill detection in live online auctions, in Proceedings of International Conference on Software Engineering Research and Practice (SERP-2009), USA, 2009, pp. 134-140.
- [33] T. Yoshida and H. Ohwada, Shill bidder detection for online auctions, in Proceedings of the 11th Pacific Rim International Conference on Artificial Intelligence, Korea, 2010, pp. 351-358