Price Comparison: A Reliable Approach to Identifying Shill Bidding in Online Auctions?

[1]Fei Dong, [2]Sol M. Shatz, [3]Haiping Xu, and [4]Dibyen Majumdar

[1,2]Computer Science Department

University of Illinois at Chicago, Chicago, IL 60607, USA

{fdong, shatz}@cs.uic.edu

[3]Computer and Information Science Department

University of Massachusetts Dartmouth, North Dartmouth, MA 02747, USA

hxu@umassd.edu

[4]Mathematics, Statistics, and Computer Science Department

University of Illinois at Chicago, Chicago, IL 60607, USA

dibyen@uic.edu

Abstract

Shill bidding has become a serious issue for innocent bidders with the growing popularity of online auctions. In this paper, we study the relationship between final prices of online auctions and shill activities. We conduct experiments on real auction data from eBay to examine the hypotheses that state how the difference between final auction price and expected auction price implies shill bidding. In the experiments, a neural network based approach is used to learn the expected auction price. In particular, we trained the Large Memory Storage and Retrieval (LAMSTAR) Neural Network based on features extracted from item descriptions, listings and other auction properties. The likelihood of shill bidding is determined by a previously proposed shill certification technique based on Dempster-Shafer Theory. By employing chi-square test of independence and logistic regression, the experimental results indicate that a higher-than-expected final auction price might be used as direct evidence to distinguish likely shill-infected auctions from trustworthy auctions, allowing for more focused evaluation of shill-suspected auctions. As such, this work contributes to providing a feasible way to identify suspicious auctions that may contain shill biddings. It may also help to develop trustworthy auction houses with shill detection services that can protect honest bidders and benefit the auction markets in both the short-term and long term.

Keywords: Online auction fraud; shill bidding; shill detection; price prediction; final auction price; empirical study.

1. Introduction

Online markets have become increasingly prevalent and extended to a wide variety of formats such as auctions, group sales, private sales and direct sales. A rapidly growing number of people consider online shopping as an alternative to in-store shopping due to convenience and bargain prices. When it comes to shopping for difficult-to-get items such as antiques, collectibles, limited editions or hot consumer electronics during presales (e.g., iPad presale), many consumers turn to online auction websites. EBay is listed as the No. 3 e-commerce website worldwide, which attracted about 2.7% of world Internet users daily on average in 2010 according to Alexa.com, an Internet traffic statistics website.

Although the number of sellers and buyers attracted to online auctions is increasing, due to the nature of Internet applications, such as a high degree of anonymity, incomplete legal constraints and lower barriers to entry and exit, this contemporary business medium faces an important challenge – auction fraud. In fact, both sellers and buyers can participate in auction fraud for their own benefit. Among all types of auction frauds, shill bidding is very popular among sellers and their collusive bidders. Shill bidding occurs when a bidder, who is not interested in purchasing an item, places bids in the auction for the purpose of inflating the final auction price. Shill bidders can be the sellers themselves using a "fake" bidder ID or their helpers who collude with the sellers. As a type of auction fraud in e-commerce, shill bidding has its deep roots in traditional on-site auctions. However, since online users are anonymous and participants cannot physically interact or see each other, shill bidding is categorized as one of the most difficult types of online auction fraud to detect. While sellers might realize that shill bidding is immoral, most sellers may not realize that shilling behavior is in fact a crime, which is governed by Title 18 Section 1343 of United States Code (Brill et al. 1998).

Although the ultimate victims of shilling behavior are innocent bidders, the victims often do not realize that shilling behavior has taken place because shill bidding can be very subtle. Many consumer guidance websites and newspapers have provided auction fraud detection tips, such as checking if a seller and a suspicious bidder are from the same geographic region, and searching for a shill suspect's bidding history to determine if a seller and the shill suspect have a partnership. However, even if a suspected shill and a seller are located far from each other, they still could be working as shill partners using modern Internet communication applications to easily communicate. In addition, whereas there may be a large number of historical records for long-time sellers, it is difficult and time-consuming for a bidder to discover the partnership between sellers and their collusive bidders. Shill bidding is so tricky that such tips are hardly useful for consumers to protect their own interests.

In order to alleviate the effects of shill bidding and provide bidders convenient assessment rules, in this research we study the relationship between shill bidding and final selling price in online auctions and also examine rules that may help detect shill bidding. Note that in this paper, the terms *actual auction price* and *final auction price* are used interchangeably to mean the final bid in first price auctions, and the *expected auction price* is the predicted auction price, which can be learned by the observation of historical auction data from the same category of items. We explore answers to the following questions:

- If there is a significant price difference between actual auction price and expected auction price, is this difference related to shill bidding? Does this difference provide a clue for shill bidding?

- What if an auction's final auction price is equal to or lower than the expected auction price? Does this imply unlikely shill bidding?

In this paper, we extend the methodology and results of a preliminary study (Dong et al. 2010). In the previous work, we collected auction data from eBay, and calculated the probability of auctions involving (or not involving) shills when their actual auction prices are higher (or not higher) than the expected prices. We found that the statistical results are quite suggestive; however, due to the preliminary nature of the earlier study, the results were not justified by any formal statistical tests. Therefore, in this paper, we formally examine the association of final auction price and shill bidding using chi-square statistic, and carefully inspect their relationship by adopting a logistic regression model. Our empirical study closely follows the guidelines for empirical research in software engineering defined in (Kitchenham et al. 2002). As our major contributions, we studied the relationship between final auction prices and shill bidding activities, and found that shill bidding has a positive impact on final auction prices. We also reported data sampling criteria, hypotheses formation, parameter significance test and goodness fit of models. Based on the formal statistical models proposed in this paper, we provide both the auction houses and online bidders a time-efficient yet reliable way to assess the likelihood of shill bidding in online auctions.

This paper is organized as follows. Section 2 reviews related work to this research. Section 3 outlines the hypotheses that are to be examined. Section 4 describes the experimental setups. The model development and hypotheses tests are reported in Section 5. Section 6 discusses the implications and potential threats to the validity of this research. Finally, Section 7 concludes this paper.

2. Related Work

A substantial amount of work has been done in the study of auction data. Heijst et al. (2007) combined text mining and boosting algorithms to predict final auction prices. Ghani and Simmons (2004) compared a regression model, a neural network and a decision tree, and they achieved the best result using the neural network when treating the price prediction problem as a series of binary classification problems. Lim et al. (2008) employed grey system theory to predict auction closing prices in a simulated auction environment. Different from the above approaches, in this paper, we use a neural network approach, but for the special purpose of predicting the expected auction prices. In particular, the expected auction price is learned from the Large Memory Storage and Retrieval (LAMSTAR) Neural Network (Graupe 2007), where price prediction is based on features extracted from item descriptions, listings and bid properties.

The topics of shill detection and verification have also attracted significant research attention. Jenamani et al. (2006) derived an equilibrium bidding strategy called shill counteracting bidding strategy (SCBS) to help honest bidders counteract shills in English auctions. Patel et al. (2007) proposed a real-time shill monitor for agent-based online auction systems using role-based access control mechanisms. Xu et al. (2007) introduced a formal model checking approach for detecting shilling behaviors, especially competitive shilling behaviors. Trevathan and Read (2009) identified various shill bidding patterns and designed an algorithm to detect the patterns in online auctions. Dong et al. (2010) proposed a decision support system based on Dempster-Shafer (D-S) theory (Shafer 1976) to compute the likelihood of shill bidding activities. Kauffman and Wood (2005) discovered that the existence of shill bids in an online auction could drive up the final selling price of the auction. In contrast to the previous work, we aim to study whether and how ordinary bidders can infer potential shill bidding activities by simply investigating the differences between actual auction prices and expected auction prices.

The body of empirical research on online auctions is growing. Roth and Ockenfels (2002) had observed the prevalence of "bid sniping," which is a strategy that helps bidders counteract shill bidding by avoiding bidding until the final moment of an auction. Lucking-Reiley (2000) and Bajari and Hortacsu (2003), based on empirical investigations, concluded that a lower starting price can bring more bidders into an auction and may increase the final price of an auction. Kauffman and Wood (2003) studied the factors that can be used to predict reserve price shilling. They found book value and starting bids are indicative of reserve price shilling. Unlike previous research, we propose to examine the relationship between shilling behavior and final auction prices using logistic regression.

3. Hypotheses Formulation

In this section, we first present different opinions of how shill bidding might relate to final prices in online auctions. Based on these opinions, we propose some hypotheses regarding the relationship between shill bidding and final auction prices.

3.1 Conflicting Opinions

Generally, two conflicting opinions can be observed. On one hand, there is the possibility of a so-called "Shiller's curse," which presumes that when there are shill bids in an auction, the auction's final price will be lower than it otherwise would be – as if the shills were cursed. On the other hand, there is a view that shill bids might be interpreted as signals that the item is highly valued, thus inflating the final auction price.

Shiller's curse. "Shiller's curse" is a term that was proposed by Wang et al. (2002). The idea is that when buyers realize or suspect the existence of shill bidding, they may shield their bid and quit the auction or wait for the seller to sell the item for a lower price in the next round of the auction. In accordance to this logic, we may expect that under such a situation the final auction price would be suppressed. This is because bidders become conservative when considering shill bidders' existence and may even quit the bid war, so the auction price is not bid up. This scenario is practical since most online auction websites allow sellers to relist their items several times until the items are sold. Relisting occurs for many reasons, for example no bidders having placed bids in the auction, the final price of the auction not reaching the reserve price, or a seller engaged in shilling and accidentally "won" the auction. Kosmopoulou and De Silva (2007) examined the effect of shill bidding in online auctions on the seller's payoff and on the price. Through a series of experiments they found that auction prices decreased as bidders anticipated the behavior of sellers and adjusted their bidding strategies.

Signaling Effect. There are also researchers who believe that shill bidding will inflate the final price of an auction. Wang et al. (2001) showed that private-value English auctions with shill bidding could result in a higher expected seller profit than other types of auctions, violating the classical revenue equivalence theory. Kauffman and Wood (2005) examined the effects of shill bidding on final bid price in rare coin auctions and showed that some bidders might view shill bids as signals that an item is worth more, thus they might pay more than other bidders who cannot see such signals. Roth and Ockenfels (2002) noted that shill bids could be considered as valuation indicators for an item. The earlier a shilling signal appears for an item, the higher the final price will be. Thus, bidding high early in the auction could possibly invoke a bidding war.

Actually, it is possible, and we feel likely, that both viewpoints are valid, depending upon the context. Over the short haul, when shill bidding is not a widely known practice to general bidders, the signal effect of shill bidding may play a key role in determining the final auction price of individual auctions. However, over the long haul, if shill bidding is not counteracted by auction houses in an effective way and bidders become widely aware of the practice of shill bidding, then the shiller's curse may become an important factor in auction prices.

3. 2 Hypotheses

Although shill bidding is a growing issue in online auctions, it is not yet a widely understood practice. Thus, we believe that shill bidding does likely increase the final price of many auctions. This motivates our first analysis using the theory of hypothesis testing – we provide two hypotheses, a null hypothesis H_0 (which we will show is not supported by the analysis) and H_1 (which essentially claims that H_0 is not true).

□***H_o*: *The difference between actual and expected auction price in online auctions IS NOT related to shill bidding.***

□***H_1*: *The difference between actual and expected auction price in online auctions IS related to shill bidding.***

Since a shill bidder's primary goal is to drive up the final price, it seems reasonable that the final auction price should be significantly higher than it would have been if no shill bids were placed in the auction. In addition, Roth and Ockenfels (2002) claimed that bids might act as signals to other bidders about the quality of the item, as well as the quality of service and trustworthiness of the seller. As such, shill bidding may lead to a high final price of the auctioned item. Kauffman and Wood (2005) also found that shill bidding acted as a signal for other bidders to place higher bids and thus increased the auction's winning bid. In other words, if shill bidding occurs in an auction, the auction will very likely end with a higher final auction price. Based on these views, we state the following hypothesis:

□ ***H_2* :** *A higher-than-expected final auction price is possibly due to shill bidding.*
If the statistical test of H_0 versus H_2 is rejected then the data will validate H_2.

As we discussed before, by the logic of Shiller's Curse, sophisticated bidders will reduce their bids to a greater extent if they observe a suspicious bidder, so the auction's final price will be decreased. If this is true, it would support the following hypothesis:

□ ***H_3*:** *A lower-than-expected final auction price is possibly due to shill bidding.*

If the statistical test of H_0 versus H_3 is rejected then the data will validate H_3.

Overall, we have four hypotheses, each based upon previous research results.

4. Experiments

We now turn our attention to the development of experiments to investigate the relationship between shilling behavior and final selling price in online auctions.

4.1 Experiment Design

First, we built and trained a neural network to predict auction prices (to be introduced in Section 4.3). When the neural network based price predictor achieved good performance, we employed the price predictor to predict the final prices of new auctions that were not used in the training and testing phases. Since the price predictor can achieve a relatively high accuracy, we consider the predicted prices as "expected" prices. We randomly selected 192 auctions that are not used for training and testing the price predictor. A predicted price is computed for each of the auctions and compared to the actual price. The relationship between actual price and expected price is recorded. A shill score is also computed for each of the 192 auctions. The shill score of an auction is defined as the highest belief of a bidder being a shill among those who participate in the auction. The shill score is a number between 0 and 1 (inclusive) that indicates the likelihood of an auction involving shills. The shill score and belief of shill are defined in Section 4.4. In brief, the predicted price is compared with the actual final auction price, and the relationship is recorded together with shill information. We performed chi-square test and tests based on logistic regression on the data to test the hypotheses formulated in Section 3.2. In the following sections, we describe the methods for obtaining expected auction prices and deciding shills.

4.2 Data Collection and Sample Choice

The popular auction website, eBay.com, offers a broad range of auctioned items and provides detailed information of each auction as well as some limited information on sellers and bidders. There are several ways to collect data from an auction website such as eBay.com. Users can query auction data from an auction website using web services. Such web services are usually easy to use but the types of data exposed to the customers are limited. Another way to collect data from an auction website is to program a customized web data crawler, which gives a user more flexibility. For our study, we designed a customized software agent to collect auction data from eBay.com. Given the data collecting criteria, the agent is able to retrieve specified data for completed auctions, and store the data locally. The data collecting agent follows three steps to collect the auction data. First, a central server obtains the URLs of those completed auctions from the auction listing pages in accordance with collecting criteria such as the category of the item, the name of the item, etc. Second, the server establishes and maintains a global queue that can be accessed by crawlers to track the URLs. Third, with the URLs, the crawlers sequentially scan HTML tags to extract and store the desired data.

The data we gathered is under the category of Nintendo WII game console systems. There are two reasons why we chose Nintendo WII game console system for our empirical study. First, Nintendo WII game console systems were popular among shoppers when the data was collected, and thus auctions of WII game systems attracted ample bidders and bids for this study. Second, the popularity and the price range made WII game systems good targets for shilling. Note that if an item is less popular, or the price range of the item is too low, such as several dollars, the item is less likely to be a target of shill bidders. Although the broad categorization of the auction data is WII game console system, the items bundled with the game systems vary from auction to auction. For example, one such auction may include a WII game console and a new WII FIT, which is an accessory of WII game system; while in another auction, the auctioned item could be a bundle consisting of a Nintendo WII System Console, Steering Wheel and 13

Games. This is why the cost of Nintendo WII game console systems covers a reasonable price range and makes Nintendo WII a quite suitable subject for our study.

We gathered a dataset of 1792 distinct auctions. The data was separated into two disjoint groups randomly. One group containing 1600 shill-free auctions was used to train and test the price prediction model that is to be introduced in Section 4.3. Another group containing 192 auctions is used to fit the logistic regression model, to be introduced in Section 5.2. Descriptive statistics for the data are given in Table 1.

Table 1
Descriptive statistics for actual auction price

	MIN.	1st Quartile	MEDIAN	MEAN	MAX.	3rd Quartile	Std. Deviation
Actual Auction Price	$135.8	$271.6	$308.7	$303.2	$410.0	$349.2	$61.3804

For each auction, we collected data that was provided by the seller, including information about the seller, details of the item (name, specifications, description, photos, etc.) and attributes about the auction (length, starting bid, reserve price, shipping charges, etc.). The data is processed to extract attributes and create new features that are then used to predict the final prices. The data features, classified in four different groups, are listed in Table 2.

Table 2
Data features in four different groups

<table>
<tr><th>Group Name</th><th>Item</th><th>Seller</th><th>Bid Details</th><th>Category Specific</th></tr>
<tr><td rowspan="8">Features</td><td rowspan="2">Condition (new, used, refurbished, unspecified)</td><td rowspan="2">Reputation percentage (%)</td><td>Number of bids</td><td rowspan="2">Bundle (Yes/No)</td></tr>
<tr><td>Number of bidders</td></tr>
<tr><td>Number of pictures</td><td>Has a web-store (Yes/No)</td><td>Starting bid ($)</td><td>Game (number)</td></tr>
<tr><td rowspan="3">Shipping cost ($)</td><td rowspan="3">Reputation score</td><td>Month</td><td rowspan="3">Wheel (Yes/No)</td></tr>
<tr><td>Day</td></tr>
<tr><td>Starting time</td></tr>
<tr><td rowspan="2">Description</td><td rowspan="2">Top-Rated (Yes/No)</td><td>End Time</td><td rowspan="2">Fit (Yes/No)</td></tr>
<tr><td>Duration</td></tr>
</table>

4.3 Price Prediction

We built a price predictor based on a neural network, specifically the Large Scale Memory Storage and Retrieval (LAMSTAR) (Graupe and Kordylewski 1998). The LAMSTAR, which combines Self Organization Map (SOM) and statistical decision tools, has been successfully applied to diagnosis, prediction and detection applications (Graupe 2007). The trained network for predicting final auction prices is shown in Fig. 1.

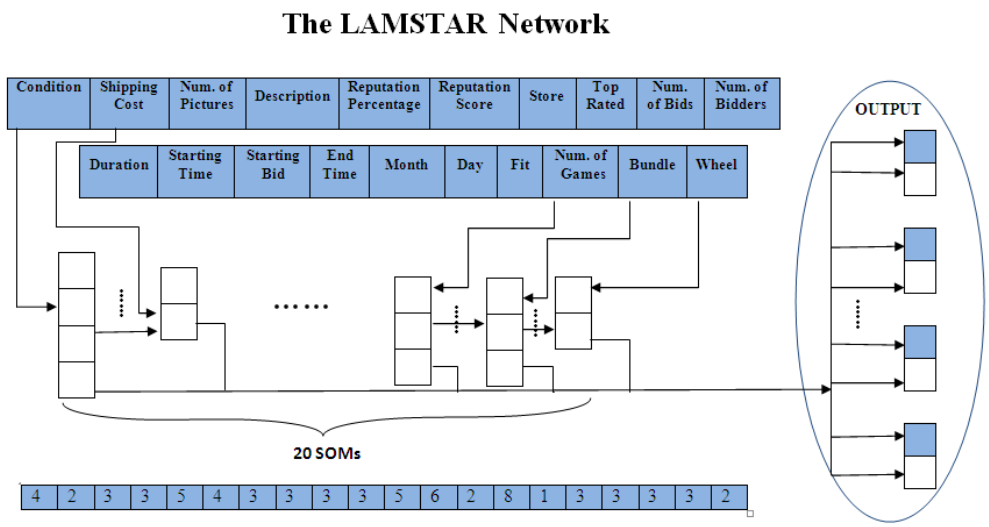

Fig. 1. The network for predicting final auction prices

The grids on the top of the figure are the subwords, representing the features in Table 2. The features are preprocessed and then provided as inputs to the neural network. For every subword there is an associated SOM module (in the middle of the figure) that is used to store and retrieve information in the training process. For each subword, a winning neuron in the associated SOM module is determined based on the similarity between the input and a storage-weight vector (stored information).

In the middle of the figure, the arrows between SOM modules encode the correlations between them. The link-weights between different SOM modules and the link-weights from the SOM modules to the output decision layer are continuously trained during normal operational runs. The output decision layer is depicted as the grids on the right side of the figure (pointed to by arrows and also circled by an oval shape). The link-weights are adjusted on a reward/punishment principle. Specifically, for the weights of links to the output layer, if the output of the particular output neuron is desired, the link weight for that neuron is rewarded by a small non-zero increment, while if the output is not desired, the link weight is punished by a small non-zero decrement. The link-weights between SOMs are trained in a similar way.

The network is designed with multiple output layers, and each layer consists of two neurons, so each layer represents a binary classifier: whether the final auction price is within a certain $\$X$ range or not. A

$50 range (i.e., X=50) was determined by the Wii gaming system training data. We observed during training that a $50 range produces the best prediction precision. So, the price predictor is trained to predict if an auction's final price is within a $50 range or not, such as ($185, $235], rather than a specific numeric value. Since the minimum price in the collected data set is around $135, and the maximum price is $410, six output layers (categories) are automatically decided by the data, namely ($135, $185], ($185, $235] , ($235, $285], ($285, $335], ($335, $385], and ($385, $435]. Table 3 shows the frequency of occurrence for expected auction prices in each category.

Table 3
Frequencies of expected auction prices

	($135, $185]	($185, $235]	($235, $285]	($285, $335]	($335, $385]	($385, $435]
Frequency	19	25	37	33	65	19

Because the actual final price is a specific number, while the expected price is defined using a range, to compare these prices, the actual price is compared to the average price of the expected range. For example, if the predicted price range is ($185, $235], its average $210 is considered as the predicted price. When an actual final auction price p_1 is compared with a predicted price p_2 of a price range (p_2-25, p_2+25], p_1 is considered as higher or lower than p_2 if p_1-p_2 > 25 or p_1-p_2 <= -25, respectively, and p_1 is considered as equivalent to p_2 if p_1 falls into the price range (p_2-25, p_2+25].

A total of 1600 auctions were randomly selected from auctions that were tested to be shill free, based on the shill analysis to be introduced in Section 4.4. After training the neural network on 1000 auctions and testing it on another 600 auctions, the neural network achieved a precision of 95.5%. This means that, given an auction, the price predictor can determine with a small chance of error if the final auction price will fall in a price span of $50. In this experimental study, we only focus on predicting the final price of one specific category of items. We leave the work of predicting the final price of general items as future work.

4.4 Shill Analysis

Because this paper aims to present an empirical study of how the difference between final auction price and expected auction price can provide clues for shill bidding, shill analysis results are needed to feed the logistic regression model so as to examine the hypotheses H_0, H_1, H_2, and H_3. Dong et al. (2010) introduced a shill certification method based on the mathematical theory of evidence, Dempster-Shafer (D-S) theory. Six bid-level properties and two auction-level properties are quantified to compute the *belief of shill* for every bidder in an auction. The bid-level properties include the time of a bidder's last bid in an auction, the bidder's concurrent bidding activities, the bidder's reputation score, the bidder's average bid increment, the bidder's winning ratio, and a bidder's affinity for the seller. The auction-level properties include the number of bids and the starting price of the auction.

The D-S theory considers a universe of discourse Θ (also called frame of discernment) that consists of a finite set of mutually exclusive atomic states in a problem domain (Shafer 1976). For example, in the auction shill detection domain, the frame of discernment for a bidder is Θ = {shill, ~shill}. The power set 2^{Θ}, which is the set of all possible subsets of Θ including the empty set, can be denoted as 2^{Θ} = {Ø, {*shill*}, {~*shill*}, Θ }. The D-S theory assigns a belief mass to each subset of the power set by function *m*: $2^{\Theta} \rightarrow [0,1]$. The function is called basic mass assignment (BMA) if it satisfies the following two equations:

$$\sum_{A \in 2^{\Theta}} m(A) = 1 \qquad (1)$$

$$m(\varnothing) = 0 \qquad (2)$$

Given a certain piece of evidence, *m*(*A*) represents one's belief exactly on state *A*, not any subset of *A*. The empty set Ø represents a contradiction, which cannot be true in any state. Therefore, the BMA for Ø is assigned 0. The basic mass assignment *m*(Θ) can be interpreted as the total ignorance of the problem domain, where one feels uncertain about the truth because every state is present. For the shill detection problem, Eqs. (1) and (2) imply that *m*(*shill*) +*m*(~*shill*) +*m*(Θ) = 1.

To obtain the overall belief on state *A*, one must take the sum of beliefs on all subsets of *A*. As defined in Eq. (3), a belief function is defined as the mass sum of all *B*s, which are subsets of *A*. Thus, in D-S theory, a degree of belief is represented as a belief function rather than a Bayesian probability function, and mass values are assigned to sets of elements rather than singletons.

$$bel(A) = \sum_{B \subset A} m(B) \qquad (3)$$

Based on Dempster's rule of combination, the formula for computing the belief of shill is as follows:

$$belief(shill_i) = m(shill_i) \qquad (4)$$

$$m(shill_i) = m_1(shill_i) \oplus m_2(shill_i) \oplus \cdots \oplus m_n(shill_i) \qquad (5)$$

The shill certification approach is demonstrated to be effective and accurate. In the experiments, we compute the shill score for every bidder in an auction. The shill score for an auction is defined as the highest shill score among the bidders. The auction is suspected to involve shill bidding when the shill score for the auction is higher than 0.9; while the auction is considered to be free of shills when the shills score for the auction is less than 0.5. For more details on this approach, refer to (Dong et al. 2010).

5. Hypothesis Tests and Results

In this section, we present the model development procedures and results of the hypotheses tests. As stated earlier, chi-square test of independence and logistic regression analysis are carried out on the collected auction data that directly relate to the research questions.

5.1. Testing of Hypothesis H_0 versus H_1

To test that whether shill bidding is associated with the difference between final auction prices and expected auction prices, the collected auction data are analyzed using chi-square statistics. The chi-square ($\chi 2$) test is a non-parametric statistical method that can be used to test the significance of association between two variables. The $\chi 2$ value is computed from a contingency table, which records the counts or frequencies of different combinations of attribute values. As a rule of thumb, statistics references (Mason et al. 1998) recommend the use of chi-square test only if (1) all cells in the contingency table have expected values greater than 1, and (2) at least 80% of the cells in the contingency table have expected values greater than 5. The $\chi 2$ statistics is approximately distributed with a chi-square distribution, which is often tabulated in statistics texts and available in statistical software packages.

Given an $r \times c$ contingency table, the $\chi 2$ is computed as follows:

$$\chi^2 = \sum \frac{(E_{ij} - O_{ij})^2}{E_{ij}} \qquad (6)$$

$$E_{ij} = \frac{R_i * C_j}{N} \qquad (7)$$

where $\sum$ is taken over all cells of the table, E_{ij} is the expected frequency for the cell in the i^{th} row and the j^{th} column, O_{ij} is the observed frequency in the cell, R_i is the total number of subjects in the i^{th} row, C_j is the total number of subjects in the j^{th} column, and N is the total number of subjects in the whole table. The degree of freedom is equal to $(r-1)(c-1)$ where r is the number of rows, and c is the number of columns.

The numbers of auctions that are detected to involve shill bidding and not involve shill bidding for different relationships between the actual auction price and the expected auction price are provided in Table 4. Recall that auctions' prices are predicted in ranges. When comparing expected price with actual price, the expected price takes the value of midpoint of the expected range. A total of 192 WII online auctions are sampled for this study.

Table 4
The distribution of shills and non-shills

	Shill	Non-shill	∑row
Actual Price -Expected Price > 25	44	22	66
| Actual Price - Expected Price | ≤ 25	1	61	62
Actual Price - Expected Price < -25	4	60	64
∑column	49	143	192

Since the Hypotheses H_0 and H_1 are non-directional, the first row and third row of Table 4 are merged in order to apply a non-directional chi-square test. Table 5 shows a contingency table for testing Hypotheses H_0 and H_1. The number in parenthesis is the expected number for that cell, calculated using Eq. (7).

Table 5
The contingency table for shill bidding and auction prices

	Shill	Non-shill	∑row
| Actual Price - Expected Price | > 25	48 (33.18)	82 (96.82)	130
| Actual Price - Expected Price | ≤ 25	1(15.82)	61(46.18)	62
∑column	49	143	192

Applying Eq. (6) to the statistical data in Table 5, we get $\chi 2$= 27.54 and degree of freedom = 1. The result is significant at the 99% confidence level. In other words, the result indicates that there is a strong relationship between final auction prices and shill bidding. Therefore, Hypothesis H_0 is rejected in favor of the alternative hypothesis H_1. Hence we may conclude that the difference between actual and expected auction price in online auctions is related to shill bidding.

5.2 The Logistic Regression Model

To better study the relationship between final auction price and shill bidding, an empirical model based on logistic regression is developed. The reason why we choose logistic regression (aka logit model) is that the outcome variable in logistic regression is binary or dichotomous, such as shill or non-shill. In addition, logistic regression allows a discrete outcome to be predicted from values of one or more variables that may be continuous, discrete and binary, or a mixture of any of these (Hosmer and Lemeshow 2000). This model will be used to test Hypothesis H_0 versus H_2 and Hypothesis H_0 versus H_3.

The general form of the logit model is shown as in Eqs. (8) and (9), in which *f(z)* represents the expected value of the outcome variable given the independent variables, x_i .

$$f(z) = \frac{e^z}{1+e^z} \qquad (8)$$

$$z = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \qquad (9)$$

where α is the model intercept, β's are the regression parameters for the independent variables, and n is the number of predictive variables in the model.

The logit transformation is defined in terms of *f(z)* as in Eq. (10).

$$g(z) = \ln[\frac{f(z)}{1 - f(z)}] = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \qquad (10)$$

The transformation is important because *g(z)* has many of the desirable properties of a linear regression model. The function *g(z)* is linear in its parameters that may be continuous, and may range from -∞ to +∞, depending on the range of x_i. The conditional distribution of the outcome variable follows a binomial distribution with the probability given by *f(z)*. For more details, refer to (Hosmer and Lemeshow 2000).

5.3 The Conceptual Model

We will build a model to predict whether an auction involves shill bidding using its actual and expected auction price. Our goal is to show that online auction users can easily determine if an auction is trustworthy (i.e., no presence of shill bidding) with the help of an auction price predictor. Recall that the price prediction model described in Section 4.3 is trained by four groups of features such as item, seller, bid and category specific features, and each feature-group includes several detailed features such as the seller's reputation, date and time when an auction begins, starting bids, etc. Since the price prediction model has already factored in many crucial features except shill bidding, we focus on studying the relationship of shill bidding and online auction price, specifically, how shill bidding might be detected by the difference between actual auction price and expected auction price. This explains why only auction prices are considered in our model.

First, we define two indicator variables as in Eqs. (11) and (12).

$$x_1 = \begin{cases} 1, \text{if actual price - expected price} > 25 \\ 0, \text{otherwise} \end{cases} \qquad (11)$$

$$x_2 = \begin{cases} 1, \text{if actual price - expected price} < \text{-25} \\ 0, \text{otherwise} \end{cases} \qquad (12)$$

The reason why x_1 *and* x_2 are treated as categorical variables is that the relationship of higher or lower is more straightforward and meaningful for bidders to identify shilling behaviors than numbers. For example, it may be too subtle for bidders to understand the difference between \$30 higher and \$55 higher. However, when bidders are provided a range of the lowest price and the highest price, they can easily know if the actual price is out of the reasonable price range. For the data collected from auction listings of the WII game systems, we consider a \$25 difference as significant because each predicted price range spans \$50, and an actual price is in the expected range only if it is \$25 (or less) higher or lower than the expected price.

The logit model (Hosmer and Lemeshow 2000, Geweke et al. 1994) is defined as the following:

$$prob[ShillBidding = 1] = f(x_i) = \frac{e^{\alpha+\beta_1 x_1+\beta_2 x_2}}{1+e^{\alpha+\beta_1 x_1+\beta_2 x_2}} \tag{13}$$

Eq. (13) implies the relationship between the probability of shill bidding and the difference of actual price and expected price as shown in Eq. (14).

$$prob[ShillBidding = 1] = \begin{cases} \frac{e^{\alpha}}{1+e^{\alpha}}, \text{if } |\text{ actual price - expected price }| \le 25 \\ \frac{e^{\alpha+\beta_1}}{1+e^{\alpha+\beta_1}}, \text{if actual price - expected price} > 25 \\ \frac{e^{\alpha+\beta_2}}{1+e^{\alpha+\beta_2}}, \text{if actual price - expected price} < \text{-}25 \end{cases} \tag{14}$$

5.4 Results

We used the statistical package *R* (http://www.r-project.org/) to perform logistic regression. The model parameters were estimated using the *maximum-likelihood* method. Table 6 provides the estimation results.

Table 6
Logit model parameter estimation results

Model Variables	Estimated Coefficient	Std. Error	Pr(>|Z|)
β_1	4.804	1.041	3.95e-06***
β_2	1.403	1.132	0.215
α (Intercept)	-4.111	1.008	4.53e-05***
Note: ***=significant at p<0.0001; Log Likelihood = -62.09163 (df=3);			

As we can see from Table 6, β_1 is statistically significant at the 99% confidence level, however, β_2 is not statistically significant. This means that in our data higher-than-expected auction price has a significantly positive relationship to shill bidding while the lower-than-expected auction price does not indicate likely shill bidding. Therefore, in testing H_0 versus H_3 we fail to reject H_0, so we cannot conclude that shilling suppresses expected price. However, we found evidence to reject H_0 in favor of H_2. Table 7 shows the predicted probabilities of shill bidding and non-shill bidding for different relationships of actual auction price and expected auction price.

Table 7
Predicted probability of shill bidding

x_1	x_2	Relationship	Prob. of Shill Bidding	Prob. of non-shill bidding
0	1	**Actual Price - Expected Price< -25**	0.0625	0.9375
0	0	**| Actual Price - Expected Price | ≤ 25**	0.0161	0.9839
1	0	**Actual Price -Expected Price > 25**	0.6667	0.3333

The probability of shill bidding is 66.67% when the actual price of an auction is significantly higher than the expected auction price. This figure is quite suggestive. Honest bidders could possibly reduce their risks of being victimized to a great extend if they quit auctions with actual price significantly higher than the expected price. Meanwhile, when the auction price is less-than or equal-to the expected auction price, the likelihood of shill bidding is much lower but the likelihood of non-shill bidding is quite high. Thus, the lower-than-expected or equal-to-expected auction prices are also valuable signals, indicating an auction is likely to be trusted.

5.5 Model Fit

Now we consider the effectiveness of the model to predict the outcome. This is referred to as *goodness of fit*. There is no generally accepted goodness of fit measure, such as R-square for linear regression, for binary outcome models such as logit (Kennedy, 1998). However, chi-square and deviance are often adopted by researchers to evaluate the goodness of fit for logistic regression. The test statistic examines whether the model with exploratory variables fits significantly better than the null model. For our model (χ^2=93.92, p<0.0001), the chi-square of 93.92 with two degrees of freedom and an associated p-value of less than 0.0001 demonstrate that the model as a whole fits significantly better than a model without the exploratory variable. Thus, we have convincing evidence that x_1 is a significant variable in predicting shill bidding. Note this chi-square test is different from the chi-square test of independence presented in Section 5.1. For details of the chi-square and deviance test, refer to (Hosmer and Lemeshow 2000).

6. Implications and Threats to Validity

6.1 Implications

Our findings have several practical implications. First, shill bidding increases the final auction price, and thus increases the commission fees that sellers have to pay to an auction house. So, in the short run, shill bidding is profitable for both sellers and online auction hosts. However, shill bidding can be detrimental to online auctions in the long run. If shill bidding becomes rampant or even if it is just suspected as being widespread, honest bidders will tend to bid at a price much lower than their evaluation after they factor in shill bidding. In this case, sellers will not be able to sell items at the proper market prices and have to reduce the prices. As a result, auction houses will earn less commission fees. Therefore, auction houses need to actively and aggressively detect online auction fraudsters. Our research results offer a time-efficient way to access the likelihood of shill bidding in online auctions. Since auction houses have access to related historical auction data, it is possible for auction houses to utilize the data to obtain an accurate predicted auction price for each auction. By comparing the final auction price to the predicted auction price, an auction can be classified into one of the two groups: auctions that have higher-than-expected prices and auctions that have lower-than or equal-to expected prices. An auction in the first group might be then subjected to some appropriate actions such as suspension of the auction and further investigation if needed.

For bidders in online auction markets, this paper examines how a final auction price might reflect shill bidding and provides simple rules to discover signals of shill bidding. A better understanding of the relationship between final auction prices and shill bidding can help honest bidders protect themselves from being cheated and thus reduce the chance of monetary loss incurred by fraudulent bidding behavior. Perhaps auction houses or third parties should provide price estimation or prediction as a service to assist honest bidders in diagnosing auctions for shill bidding. When a bidder finds that the final price of an auction is higher than the expected price, bidders can decide whether they want to continue participating in the auction.

The difference between the final auction price and the expected auction price is in fact reliable in identifying auctions involving shilling behavior because there is a very slight chance for a shill to counteract this indicator. If shills want to counteract this indicator, they can either devise some ways to impair the accuracy of price prediction or lower the final auction prices intentionally. However, neither way is practical to be put into action. To impair the accuracy of the price prediction, shills need to pollute the training data but they can neither know which auctions will be included in the training data nor constantly place shilling bids without being caught. Reducing the final auction price is against the purpose of shill bidding so shill bidders probably will not take actions to achieve this. Therefore, it appears that the difference between final auction price and expected auction price is a reliable indicator of shill bidding.

6.2 Threats to validity

During the experiments and results interpretation, we held certain assumptions that could lead to validity threats. In this section we list the major issues that could pose some threats to the validity of our arguments and findings in the paper. The main threats to validity in this study relate to auction price prediction, shill detection, and data collection, each detailed below.

Expected Price The statistic analysis process was carried out with the assumption that the expected price range of an auction is known precisely when an auction ends. As introduced in Section 4.3, the expected auction price is the same as the auction price predicted by a trained neural network. Hence, the accuracy of the auction price predictor may be a potential threat to validity of the statistic analysis results. In other word, if an auction's price cannot be predicted in high accuracy, the hypotheses test results in this paper may be invalid. To limit this threat, the LARMSTAR neural network based price predictor is carefully trained and tested using real auction data. It can predict the final auction price in a range with high precision. However, to generalize this approach to a broader category, more research efforts are still needed to refine the accuracy of auction prices prediction.

Detection of Shill The data used for logistic regression analysis we presented in Section 5.2 depends critically on the accuracy of identifying shill bidders in an auction, which is reported in Section 4.4. An ineffective and low-precision shill detection technique could affect the hypotheses test results and in turn affect the effectiveness of the rules derived from the hypotheses tests. Nonetheless, the Dempster-Shafer theory of evidence based shill detection approach proposed by Dong et al. (2010) has been compared with manual investigation results and demonstrated to be effective and accurate. Thus, this threat should be limited.

Data Collection Including only auction data from WII gaming systems that were listed in eBay platform limits the possibility to generalize the application of this approach to other categorization of auction items. This also introduces the risk of missing information that could be helpful in improving auction prices prediction. A number of other factors, such as the influence of different goods and the reputations of bidders, can be considered in order to extend the approach. However, as most auctions have similar listing structures, and the gaming system auctions feature plenty of accessories and console combinations, the lessons and experiences that are learned from gaming system auctions can be applied to other categories of items. We will leave generalization of this approach to other categories of items as future work.

7. Conclusions and Future Work

Due to the popularity of online auctions and the seriousness of shill bidding, preventing and detecting shill bidders has become an increasingly urgent task. It is important that effective and efficient systems be built to protect honest bidders and the online auction markets. This paper studied whether auction users can infer shill-bidding behavior from the difference between actual auction price and expected auction price. By employing chi-square test of independence and a logistic regression model, we examined the contrary predictions made by extensions to existing auction theory in an attempt to explain how bidders can infer shill bidding from the difference of final selling price. The results show that the relationship of final auction price and expected auction price could be considered as a reliable indicator of shill bidding. We also found that a higher-than-expected final auction price suggests possible shill bidding. We believe that the rules derived from the hypotheses tests in this paper are helpful and applicable for both auction houses and auction users. Honest bidders can protect themselves from being cheated and reduce the risk of monetary loss in online auctions. In addition, auction houses can adopt the rules to complementing existing shill detection techniques and enhancing the confidence of auction users towards auction houses.

In our future work, we will study auction data in different categories of items, refine our price prediction techniques and further validate the signals that can help identify shill bidding activities. We also plan to study how the proposed method can be implemented in actual auction environments so that auction houses can provide trustworthy services to their customers.

Acknowledgment

This material is based upon work supported by the U.S. National Science Foundation under grant number CNS-0715648 and CNS-0715657. We thank the editors and the reviewers for their very helpful comments that improved the presentation of this work. We also thank Dr. D. Graupe for his advice on designing and improving the precision of the LAMSTAR neural network.

References

Bajari, P. and Hortaçsu, A. Winner's Curse, Reserve Prices and Endogenous Entry: Empirical Insights from eBay Auctions. *Rand Journal of Economics,* 2, 2003, 329-355.

Brill, A. E., Baldwin, F. N. and Munro, R. J. *Cybercrime and Security*. Oceana Publications, 1998, 29.

Dong, F., Shatz, S. and Xu, H. Combating Online in-auction Fraud: Clues, Techniques and Challenges. *Computer Science Review*, 3, 4, 2009, 245-258.

Dong, F., Shatz, S. and Xu, H. Inference of Online Auction Shills Using Dempster-Shafer Theory. Proceedings *of the 6th International Conference on Information Technology: New Generations (ITNG 2009)*, April 2009, 908-914.

Dong, F., Shatz, S. and Xu, H. An Empirical Study of Online Auctions with Unexpected Prices. *Proceedings of 22nd Intl. Conf. on Software Engineering & Knowledge Engineering*, San Francisco Bay, 2010, 286-291.

Dong, F., Shatz, S. and Xu, H. Reasoning Under Uncertainty for Shill Detection in Online Auctions Using Dempster-Shafer Theory. *International Journal of Software Engineering and Knowledge Engineering (IJSEKE)*, 20, 7, 2010, 943-973.

Geweke, J., Keane, M. and Runckle, D. Alternative Computational Approaches to Inference in the Multinomial Probit Model. *Review of Economics and Statistics*, 76, 4, Nov.1994, 609-632.

Ghani, R. and Simmons, H. Predicting the End-price of Online Auctions. *Proceeding of the International Workshop on Data Mining and Adaptive Modelling Methods for Economics and Management*, Pisa, Italy, 2004.

Graupe, D. and Kordylewski, H. A Large-Memory Storage and Retrieval Neural Network for Adaptive Retrieval and Diagnosis. *International Journal of Software Engineering and Knowledge*

Engineering, 6, 1, March 1998, 115-138.

Graupe, D. *Principles of Artificial Neural Networks*, 2nd Edition. World Scientific Publishing Co. Pte. Ltd., Singapore, 2007.

Heijst, D. V., Potharst, R. and Wezel, M. V. A Support System for Predicting eBay End Prices. *Decision Support Systems*, 44, 4, 2007, 970-982.

Hosmer, D. W. and Lemeshow, S. *Applied Logistic Regression*, 2nd edition. Wiley-Interscience Publication, 2000.

Jenamani, M., Zhong, Y. and Bhargava, B. K. Cheating in Online Auction - Towards Explaining the Popularity of English Auction. *Electronic Commerce Research and Applications*, 6, 1, 2007, 53-62.

Kauffman, R. J. and Wood, C. A. *Running up the bid: detecting, predicting and preventing reserve price shilling in online auctions.* Technical report, University of Notre Dame, 2003.

Kauffman, R. J. and Wood, C.A. The Effects of Shilling on Final Bid Prices in Online Auctions. *Electronic Commerce Research and Applications*, 4, 1, 2005, 21-34.

Kitchenham, B. A., Pfleeger, S. L., Pickard, L. M., Jones, P. W., Hoaglin, D. C., El Emam, K. and Rosenberg, J. Preliminary Guidelines for Empirical Research in Software Engineering. *IEEE Trans. on Software Engineering*, 2002, 28, 8, 721-734

Kosmopoulou, G. and De Silva, D. G. The Effect of Shill Bidding upon Prices: Experimental Evidence. *International Journal of Industrial Organization,* 25, 2007, 291-313.

Lim, D., Anthony, P., and Ho, C. Agent for Predicting Online Auction Closing Price in a Simulated Auction Environment. *Proceedings of Pacific Rim International Conference on Artificial Intelligence - PRICAI*, 2008, 223-234.

Lucking-Reiley, D. Auctions on the Internet: What's being auctioned, and how? *Journal of Industrial Economics*, 48, 3, 2000, 227–252.

Mason, R. D., Lind, D. A. and Marchal, W. G. *STATISTICS: An Introduction*, 5th edition. Duxbury Press, 1998.

Patel, R., Xu, H. and Goel, A. Real-Time Trust Management in Agent Based Online Auction Systems. *Proceedings of the 19th International Conference on Software Engineering and Knowledge Engineering*, July 9-11, 2007, Boston, USA, 244-250.

Roth, A. E. and Ockenfels, A. Last-Minute Bidding and the Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon Auctions on the Internet. *American Economic Review* 92, 4, September 2002, 1093-1103.

Russell, S. J. and Norvig, P. *Artificial Intelligence: A Modern Approach*, 3rd Edition. Prentice Hall, 2009.

SAS Institute. *SAS user's manual: the logistic procedure*, 1999, SAS Institute, Cary, North Carolina.

Shafer, G. *A Mathematical Theory of Evidence*. Princeton University Press, 1976.

Trevathan, J and Read, W. Detecting Shill Bidding in Online English Auctions. *Handbook of Research on Social and Organizational Liabilities in Information Security*, M. Gupta and R. Sharman (eds.), Information Science Reference, 2009, pp. 446-470.

Wang, W. L., Hidvègi, Z. and Whinston, A.B. *Shill Bidding in English Auctions*. Technical report, Emory University, 2001, Available at http://oz.stern.nyu.edu/seminar/fa01/1108.pdf

Wang, W. L., Hidvègi, Z. and Whinston, A.B. Shill Bidding in Multi-Round Online Auctions. *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*, January 2002.

Xu, H. and Cheng, Y. T. Model Checking Bidding Behaviors in Internet Concurrent Auctions. *International Journal of Computer Systems Science & Engineering*, 22, 4, 2007, 179-191.