Hindawi Publishing Corporation
Advances in Decision Sciences
Volume 2014, Article ID 891954, 14 pages
http://dx.doi.org/10.1155/2014/891954

Research Article

Uncertain Reasoning for Detection of Selling Stolen Goods in Online Auctions Using Contextual Information

Ladislav Beranek

Department of Applied Mathematics and Informatics, Faculty of Economics, University of South Bohemia, 37005 Ceske Budejovice, Czech Republic

Correspondence should be addressed to Ladislav Beranek; beranek@ef.jcu.cz

Received 13 June 2014; Accepted 28 October 2014; Published 25 November 2014

Academic Editor: Henry Schellhorn

Copyright © 2014 Ladislav Beranek. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This work describes the design of a decision support system for detection of fraudulent behavior of selling stolen goods in online auctions. In this system, each seller is associated with a type of certification, namely "proper seller," "suspect seller," and "selling stolen goods." The certification level is determined on the basis of a seller's behaviors and especially on the basis of contextual information whose origin is outside online auctions portals. In this paper, we focus on representing knowledge about sellers in online auctions, the influence of additional information available from other Internet source, and reasoning on bidders' trustworthiness under uncertainties using Dempster-Shafer theory of evidence. To demonstrate the practicability of our approach, we performed a case study using real auction data from Czech auction portal Aukro. The analysis results show that our approach can be used to detect selling stolen goods. By applying Dempster-Shafer theory to combine multiple sources of evidence for the detection of this fraudulent behavior, the proposed approach can reduce the number of false positive results in comparison to approaches using a single source of evidence.

1. Introduction

Online auction portals have become a convenient trading platform for a great number of sellers and buyers. Besides eBay [1], the Aukro auction portal [2], a Czech Republic division of the multinational Allegro group with a turnover of USD 250 million and 2.5 million users as of January 2012, is an example of such an Internet auction system. Online auctions allow their users to buy or sell products and services. A large number of users use Internet auctions as their main source of income. For example, the Aukro states [2] that, from 2.5 million users, 9,300 users are professional traders. On the other hand, the number of frauds in online auctions is also increasing [3]. The most common types of fraud are: intentionally incorrect description of goods, undelivered goods, irredeemable payments, sale of stolen goods, and others [4]. Fraudsters are attracted by low admission costs and high profit potential.

In order to reduce frauds, Internet auction portals use basically three methods:

(1) built-in mechanisms which allow estimation of the credibility of users: auction systems create user profiles (reputation) that are based on the evaluation which is performed by users after performed transaction. Any user can look at the profile of another user before she/he will do business with her/him; if this user (a bidder or a seller) has a bad reputation, no one will want to do transaction with her/him;

(2) legislation and incentives: legislation solves violations similar to trading in everyday life; for example, failure to deliver goods after payment may result in a charge of fraud and so forth; most online auction portals provide some compensation in the event of similar fraud [5]; another incentive is to reward users who abide by the rules;

(3) in addition, online auction portals have internal control systems to detect fraudulent behavior; these systems monitor continuously actual auctions; they are based on procedures which consist of the two

basic steps: (1) a set of attributes is proposed and their values are extracted from a transaction history to distinguish between normal traders and scammers and (2) a detection model based on these attributes is built using various machine learning techniques; based on the statistical attributes of data transaction, these systems attempt to distinguish normal transactions from fraudulent transactions; however, the effectiveness of these control systems is not high; fraudsters attempt to deceive without triggering identification of a fraudulent transaction; they pose as new users and change their address frequently; since fraudulent users change identity and essentially disappear before data mining methods acquire sufficient information to detect their fraudulent activities, the use of data mining methods to expose such clients cannot achieve effective outcomes; fraud detection is a complex problem and online auction portals are often unwilling to invest in complex systems or to perform some actions that would deter customers; however, successful online auction portals have to be functional, able to give certain guarantees, and deal with fraudulent users; therefore, approaches to fraud detection are continually being refined for use in practice; see, for example, [6–10] and others.

To address this problem of fraudulent behavior detection, this paper proposes a new approach that does not rely solely on the data mining techniques applied to data from auctions. The distinctive feature of the approach proposed here is to use additional sources of information (context information) on the Internet. After analyzing various sources of information on the Internet, we found out that information about the theft of various items was discussed by robbed people on various internet forums, discussions, and similar Internet sources. Subsequently, the same goods were often offered via online auctions on the auction website Aukro. This context information is considered in our model as additional specific (contextual) characteristic that can be used to improve detection results of fraudulent behavior, selling stolen goods.

To detect the fraudulent behavior selling stolen goods, we propose a four-step approach. We first collect evidence that supports a belief that a seller sells stolen goods and define functions representing analyzed fraudulent behavior. In the second and third steps, we evaluate the contribution of contextual information found eventually on Internet discussion groups, Internet forums, or other sources and use this contextual information to increase the value of our conviction that stolen goods are sold in analyzed online auction. In the last step, sellers are categorized into three categories: proper seller, suspect seller, and fraudulent seller. This assessment can be useful for both auction houses and bidders in deciding on the credibility of the sellers.

Each piece of evidence is determined to support either fraudulent behavior or normal behavior based on the analysis of the corresponding quantified properties of certain seller and contextual information from Internet. Because each proof involves uncertainty, it is necessary to use formal reasoning techniques. In this context, we use belief functions of Dempster-Shafer (DS) theory [11] to model the uncertainty associated with different evidence related to the supply of various properties. This allows us to explicitly represent uncertainty and to integrate knowledge from different sources of evidence to produce aggregate assessment, in particular. This work extends our previous proposed framework of using belief functions for detection of selling stolen goods [12]. The major extensions are as follows. First, auction-level properties and evidence are introduced, which complements the seller-level evidence for providing a macroexamination of auctions. Second, in the previous work, the focus was only on quantifying the degree of belief concerning whether a seller sells stolen goods, rather than considering both cases. Here, the first case means a seller sells stolen goods and the second one means a seller does not sell stolen goods. In this paper, each piece of evidence is determined to support either fraudulent behavior or normal bidding behavior based on the analysis of the corresponding quantified selling property. Third, and most importantly, we performed an extended number of experiments in order to evaluate statistically the values of respective parameters used in our model. Therefore, the scope of the decision support system is significantly improved.

The remainder of this paper is organized as follows. Section 2 summarizes some concepts and techniques related to the detection of fraud in online auctions including works using the belief function approach. Section 3 presents the basic principles of the belief theory, including a description of a situation when there might be some doubts about the reliability of information sources. Section 4 presents our model which represents potential fraudulent behavior (selling stolen goods) and the influence of contextual information when reasoning about the measure of possible fraudulent behavior. Section 5 presents the results of the experiments. Section 6 provides conclusions and suggestions for further work. The last Section 7 describes limitations of our study.

2. Related Work

Successful Internet auction portals must perform many different activities. In addition to the creation, implementation, and operation of their trading system (including users login, displaying data about items being sold, including initial price and duration of auction, and displaying of bids of buyers and other data), they must also solve problems with trust and trustfulness. This follows from the fact that transactions in the online auction predominantly take place in a situation where the users do not know each other. However, if users want to do business here, they must decide whether they will trust each other. Online auction portals, such as eBay [1] and Aukro [2], are successful primarily because they are able to create a trusted environment for users of online auctions [13–15]. Most of the mechanisms (reputation systems), which create a trusted environment, use a variety of attributes associated with users and their roles.

Although reputation systems and certain methods of user identification are functional, a lot of fraud takes place in online auctions. The sellers and bidders are not in physical contact, and bidders are not able to even physically see the auctioned item. This situation provides many opportunities for cheating [16–18]. From the perspective of a seller, online auctions bring about certain risks [6, 19–21] particularly as follows.

(i) Bidder does not pay for the goods supplied.

(ii) Bidder wrongly claims that the goods have not been delivered.

From the perspective of the bidder, consider the following [22].

(i) The seller refuses to send the goods [4].

(ii) The description of the auctioned object is false.

(iii) Seller sends different goods or goods of lower quality [23].

(iv) Goods intended for auction are fake or stolen.

(v) Seller can influence the course of the auction (seller can collude with other bidders or can bid on his own items to drive up the price of the item being auctioned) [24]. We denote this as shill behavior.

Fraudulent behavior is relatively widespread in online auctions because an engaging in fraudulent behavior is relatively easy.

(i) Online auction participants are largely anonymous; they frequently act under pseudonyms. Internet auction systems use different methods for verifying the identity of users. These methods, however, may not be sufficiently reliable.

(ii) Online auction houses do not always exhibit full commitment to actively engage in combating fraud.

(iii) Law enforcement is often difficult with regard to differences of legal systems in different countries.

Fraudulent behavior which occurs in online auctions is not easy to detect mainly due to the use of various techniques by fraudsters to camouflage their behavior and due to the pseudonymity of users participating in the auction. The most widely used detection approaches are based on various statistical methods, data mining techniques, methods of analysis of online users behavior, or social networks analysis; see, for example, [6–10] and others. Works that deal with the detection of fraudulent behavior in online auctions focus mostly on certain types of fraudulent behavior, for example, shill behavior [20, 25]. A general approach would be very complicated due to the complexity of the issue. The selection of an appropriate set of attributes is crucial for constructing a detection model. The simplest way to devise an attribute set for fraud detection is to enumerate all the features of tricks that have already occurred. Attributes of fraudulent behaviors are taken directly from statistics related to past transactions. These attributes include the count of positive ratings and negative ratings, the median and the standard deviation, the average of all labeled prices during a specific time period [26], the starting labeled price of a bid, and some Boolean variables [24, 27] which deal with fraudulent behavior in online auctions. They propose an algorithm to detect shill behavior based on comparisons of patterns of behavior in online auctions. Trevathan and Read present in another paper [20] a method for detecting colluding shill users. Chau et al. [28] use methods based on a data mining approach to detect shill behavior. They apply this approach on the user level and on the level of interaction among users. They link these two levels to detect suspicious behavior patterns using Markov random field methods. Xu and Cheng [29] introduce a dynamic auction model for shill detection in real online auctions and use formally specified shilling behavior by the help of linear temporal logic to verify the shill behavior. The other authors [30, 31] propose using Bayesian networks or decision trees to detect fraudulent behavior in online Internet auctions. The use of belief functions to detect shill behavior is presented in the work [19, 25, 32]. The authors indicate a conceptual design framework for calculating the belief functions. They demonstrate the correctness of their approach in an eBay auctions case study. The papers [33, 34] describe some features of shill behavior which are then expressed by belief functions and combined with the aim to classify users into categories of shill, suspect, and trustworthy. Pandit et al. [9] designed and implemented an online auction fraud detection system named NetProbe. The NetProbe system models auction data as a network graph in which sellers and bidders are represented by nodes, and transactions between sellers and bidders are represented by edges. The Markov random field and the belief propagation algorithms are utilized to unearth suspicious trading patterns created by fraudsters and thus to detect possible fraudsters. An online auction fraud detection system was also presented in the works [9, 10, 26, 28, 35]. W. H. Chang and J. S. Chang [36–39] propose the data mining methods for early detection of fraudulent behavior. Forty-four attributes are defined and analyzed in this paper with the aim to build a model for early detection of fraud. Kwan et al. [40] focus on the detection of selling fake products. They define attributes of this fraudulent behavior and use the Bayes approach in their evaluation.

Generally, the detection accuracy is closely related to the suitability of the attributes and the choice of modeling method. It is obvious that previous work in this area has provided good level of progress but some problems still remain. The use of a greater number of measured attributes may not bring substantial improvement. The accuracy of detection can even deteriorate when irrelevant attributes are incorporated into the model. A major improvement in fraud detection based on data from auction portals cannot be expected at the present. We therefore suggest an improvement of fraud detection on the basis of the utilization of additional sources of information available on the Internet.

3. Dempster-Shafer Theory of Evidence

Our model is a particular application of the belief function theory. The belief function theory [11] is designed to deal with the uncertainty and incompleteness of available information. It is a powerful tool for combining evidence and changing prior knowledge in the presence of new evidence. The belief function theory can be considered as a generalization of the Bayesian theory of subjective probability. In the following paragraphs, we give a brief introduction of the basic notions of the belief function theory (frequently called Dempster-Shafer theory or theory of evidence).

Considering a finite set referred to as *the frame of discernment* Θ, a *basic belief assignment* (BBA) is a function $m: 2^{\Theta} \rightarrow [0,1]$ so that

$$\sum_{A \subseteq \Omega} m(A) = 1, \quad (1)$$

where $m(\emptyset) = 0$; see [11]. The subsets of 2^{Θ} which are associated with nonzero values of m are known as focal elements and the union of the focal elements is called the core. The value of $m(A)$ expresses the proportion of all relevant and available evidence that supports the claim that a particular element belongs to the set A but not to a particular subset of A. This value pertains only to the set A and makes no additional claims about any subsets of A. We denote this value also as a degree of belief or basic belief mass, BBM.

Shafer further defined the concepts of belief and plausibility [11] as two measures over the subsets of as follows:

$$\operatorname{Bel}(A) = \sum_{B \subseteq A} m(B). \quad (2)$$

The plausibility of A specifies the likelihood that it is not any other subset in 2^{Θ}. The quantity of plausibility of A is equal to one minus $\operatorname{Bel}(\neg A)$; that is, $\operatorname{Pl}(A) = 1 - \operatorname{Bel}(\neg A)$. For example, the degree of plausibility for fraudulent seller i is $\operatorname{Pl}(\text{seller}_i) = m(\{\text{seller}_i\}) + m(\{\text{seller}_i, \neg\text{seller}_i\})$. According to (3), it is easy to derive that the quantity of plausibility of A is equal to the sum of the masses of B, whose intersection with A is not empty. For all $A \in \Theta$, $\operatorname{Bel}(A)$ forms a lower bound for A that could possibly happen, and $\operatorname{Pl}(A)$ forms an upper bound for A to happen; see (4). Consider

$$\operatorname{Pl}(A) = \sum_{B \cap A = (\emptyset)} m(B) \quad (3)$$

$$\operatorname{Bel}(A) \leq P(A) \leq \operatorname{Pl}(A). \quad (4)$$

Hence, a BBA can also be viewed as determining a set of probability distributions P over Θ so that $\operatorname{Bel}(A) \leq P(A) \leq \operatorname{Pl}(A)$. It can be easily seen that these two measures are related to each other as $\operatorname{Pl}(A) = 1 - \operatorname{Bel}(A)$. Moreover, both of them are equivalent to m. Thus, one needs to know only one of the three functions m, Bel, or Pl to derive the other two. Hence, we can speak about belief function using corresponding BBAs in fact.

Dempster's rule of combination can be used for pooling evidence represented by two belief functions Bel_1 and Bel_2 over the same frame of discernment coming from independent sources of information. The Dempster's rule of combination for combining two belief functions Bel_1 and Bel_2 defined by (equivalent to) BBAs m_1 and m_2 is defined as follows (the symbol $\oplus$ is used to denote this operation):

$$(m_1 \oplus m_2)(A) = \frac{1}{1-k} \sum_{B \cap C = A} m_1(B) \cdot m_2(C), \quad (5)$$

where

$$k = \sum_{B \cap C = (\emptyset)} m_1(B) \cdot m_2(C). \quad (6)$$

Here, k is frequently considered to be *a conflict measure* between two belief functions m_1 and m_2 or a measure of conflict between m_1 and m_2 [11]. Dempster's rule is not defined when $k = 1$, that is, when cores of m_1 and m_2 are disjointed. This rule is commutative and associative; as the rule serves for the cumulation of beliefs, it is not idempotent.

The combination rule is usually denoted as the orthogonal sum of belief values; in other words, the combination of belief from evidence A and belief from evidence B is denoted as $\operatorname{Bel}_{A,B} = \operatorname{Bel}_A \oplus \operatorname{Bel}_B$. Therefore, the global belief of A can be represented as $\operatorname{Bel}(A) = \oplus\operatorname{Bel}_i$, for all pieces of evidence that supports A. To illustrate the concepts, consider a subset $X \subseteq 2^{\Theta}$ and evidence E_1 which yield a set of values represented by $m_{E1}(\{x\})$, $m_{E1}(\{\neg x\})$, and $m_{E1}(\{x, \neg x\})$ (we can also write this last term as $m_{E1}(\Theta)$ for Θ is the frame of discernment which has two elements $\Theta = \{x, \neg x\}$). Suppose that evidence E_1 may provide, in general, some support that X is true, that is, event x occurs, or some support that X is not true, that is, event $\neg x$ occurs. In terms of the mass function, the BBAs for x and $\neg x$ are $m_{E1}(\{x\})$ and $m_{E1}(\{\neg x\})$, respectively. Lack of knowledge about whether x occurs or not is represented by $m_{E1}(\{x, \neg x\})$. The sum of the three values is one. That is, $m_{E1}(\{x\}) + m_{E1}(\{\neg x\}) + m_{E1}(\{x, \neg x\}) = 1$. We can further assume that evidence E_1 is either reliable with probability 0.9 or unreliable with probability 0.1. Now, using our fraudulent seller behavior example, suppose that evidence E_1 supports that seller i selling stolen goods is with 100% certainty. Considering E_1's reliability, E_1 gives 0.9 degree of belief for supporting that bidder i is a fraudulent seller (i.e., $m_{E1}(\{x\}) = 0.9$), but zero degree of belief that seller i is honest (i.e., $m_{E1}(\{\neg x\}) = 0$) because the evidence does not support the belief that seller i is honest. The remaining degree of belief (0.1) is due to the uncertainty; that is, $m_{E1}(\{x, \neg x\}) = 0.1$.

Belief Function Correction. When receiving a piece of information represented by a belief function, some metaknowledge regarding the quality or reliability of the source that provides some information can be available. In the following paragraphs, we describe briefly some possibilities on how to adjust the information according to this metaknowledge.

Discounting. To handle the lower reliability of information sources, a discounting scheme has been introduced by Shafer [11]. It is expressed by the following equations:

$$^{\alpha}m(A) = \begin{cases} (1-\alpha)\, m(A) & \text{if } A \subset \Omega \\ \alpha + (1-\alpha)\, m(\Omega) & \text{if } A = \Omega, \end{cases} \quad (7)$$

where $\alpha \in [0,1]$ is a *discounting factor* and $^{\alpha}m(A)$ denotes the discounted mass of $m(A)$. The larger α is, the more mass $m(A)$ is withdrawn from $A \subset \Theta$ and assigned to the frame of discernment Θ.

Thus, the principle of discounting is transferring parts of basic belief masses (BBMs) of all focal elements which are proper subsets of the frame of discernment to the entire frame. This process is the result of additional information which indicates that the source is not entirely reliable. The transfer of BBMs from a source to the framework reflects an increase of the degree of uncertainty of the data that the source produces.

Dediscounting (reinforcement). In some cases, we need to perform an opposite operation, for example, transfer parts of basic belief mass (BBM) from the entire frame to all focal elements. This can be the result of a situation where we, for example, obtain information in which the source of the information is more reliable than we had anticipated at the beginning. We can then recompute m by reversing the discounting operation [41]. We denote this operation as reinforcement or dediscounting:

$$\begin{aligned} m(A) &= \frac{{}^{\alpha}m(A)}{1-\alpha} \quad \forall A \subset \Omega, \\ m(\Omega) &= \frac{{}^{\alpha}m(\Omega)-\alpha}{1-\alpha}, \end{aligned} \tag{8}$$

where $\alpha \in [0, {}^{\alpha}m(\Omega)]$. We denote here α as a reinforcement coefficient. The result of maximal dediscounting is the totally reinforced belief function. It is noted as ${}^{\mathrm{tr}}m$ and defined as follows:

$${}^{\mathrm{tr}}m(A) = \begin{cases} \dfrac{m(A)}{1-{}^{\alpha}m(\Omega)} & \forall A \subset \Omega \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

The scenarios associated with this idea can often be found in the multievidence pooling systems, where decisions are made based on a set of existing pieces of evidence and the corresponding confidence in or evaluation of these pieces of evidence. Evidence and corresponding confidence may be elicited in different manners, for example, drawn by different experts or based on different viewpoints.

4. Proposed Fraud Detection Method

This paper aims to detect specific Internet auction fraud related to the selling of stolen goods (i.e., goods being stolen and subsequently sold in an online auction). The detection model is based on chosen attributes of this fraud and it also uses contextual information, that is, information found on various public Internet forums and discussions to improve the prediction accuracy of the detection model. The model is based on the belief function theory. The advantage of the use of this theory is the possibility to represent partial knowledge and the possibility to combine pieces of evidence concerning possible fraudulent behavior with pieces of contextual information from the Internet sources.

Our model consists of four steps.

Step 1 (definition of belief functions representing analyzed fraudulent behavior). In the first step, determinative attributes of fraudulent behavior, the selling of stolen goods, are specified: an inadequately low price m_L, goods sold mostly at fixed price m_F, and variety of goods being sold m_V.

Step 2 (definition of the reinforcement coefficient on the basis of contextual information). In this step, certain characteristics of contextual information are used to define the reinforcement coefficient α described in Section 4.2.

Step 3 (assessment of the influence of contextual information on the evidence about fraudulent behavior). In this third step, contextual information is used to eventually increase the value of the belief function expressing our conviction that stolen goods are sold in analyzed online auction. The aim of this step is to assess the effect of “additional” contextual information about stolen goods.

Step 4 (categorization of sellers). In the last step, sellers are categorized according to the resulting belief functions representing the selling of stolen goods into three categories: proper seller, suspect seller, and fraudulent seller. Based on statistical analysis, we defined thresholds η and ξ for making decisions on a seller classification, where $\eta < \xi$. The first element η is the belief value threshold for determining whether a seller is a proper seller. If the value of $\mathrm{Bel}(\mathrm{stolen}_i)$ is below η, the seller i will be classified as a proper seller. The second element ξ is the belief value threshold for determining seller who sells stolen goods. If the value of $\mathrm{Bel}(\mathrm{stolen}_i)$ exceeds ξ, the seller will be classified as a seller who sells stolen goods. If the value of $\mathrm{Bel}(\mathrm{stolen}_i)$ is between η and ξ, the respective seller will be classified as suspect.

4.1. Belief Functions Representing Illegal Behavior. We explored in detail data about auctions related to five prosecuted cases out of 54 complaints of stolen goods being sold in Internet auctions. Our aim was to determine the characteristics of online auction fraud related to the sale of stolen goods. The cases were the result of complaints lodged with the Czech Trade Inspection [42]. Because detailed judgments were not available for the prosecuted cases, examiners who worked on the cases were interviewed to elicit some additional information. The following attributes of the online auction offering stolen goods were specified.

(1) Stolen goods were sold at inadequately low prices (at least about 20% below the price of legitimate goods).

(2) Fraudsters prefer to sell stolen goods for a fixed price.

(3) A variety of goods were sold via fraudulent accounts (such as car accessories, footwear, sporting goods, etc.).

(4) Life span of such fraudulent accounts was very short (often less than twelve days).

(5) In most cases, the goods were sold within several days of creating the account. In most cases (31 out of 54); the goods were sold within six days of account creation.

(6) Starting price of an auction where a seller sells stolen goods is usually less than the average starting price.

(7) Fraudsters had accounts on multiple auction systems, and the value of their reputational score is low.

Our aim was to define belief functions corresponding to the chosen attributes and then to combine these belief functions to assess whether the respective seller sells stolen goods or not. Based on our analysis, we have chosen the following attributes as indicators that stolen goods are being sold (the other ones were too difficult to verify or express mathematically):

(1) goods sold at inadequately low prices,
(2) goods sold mostly at fixed prices,
(3) a variety of goods being sold,
(4) low starting price.

As we mentioned earlier, the illegal behavior classification process employs a mathematical theory, Dempster-Shafer theory of belief functions, to represent the uncertainties of evidence pertaining to different hypotheses. We denoted the frame of discernment concerning the analyzed fraudulent behavior $\Theta = \{\text{stolen}, \neg\text{stolen}\}$. Here, stolen represents the hypothesis that stolen goods are sold in the analyzed online auction; ¬stolen represents the hypothesis that the analyzed online auction is conducted properly. The power set of the set Θ (the set of all subsets) 2^{Θ} has three elements (we do not consider the empty set here): $2^{\Theta} = \{\{\text{stolen}\}, \{\neg\text{stolen}\}, \{\text{stolen}, \neg\text{stolen}\}\}$, where $\{\text{stolen}, \neg\text{stolen}\} = \Theta$ denotes our ignorance. That means that we are not able to assess whether stolen goods are sold in the online auction or not. The basic mass assignment expressing our belief concerning single evidences concerning the examined behaviour (i.e., goods sold at inadequately low prices; goods sold mostly at fixed prices; a variety of goods being sold) is described in the next paragraphs.

Inadequate Low Price. This attribute shows that the seller i sells stolen goods for lower prices than the average price of the legitimate goods. The belief functions have the following forms:

$$\begin{aligned} m_L(\{\text{stolen}_i\}) &= v_L \frac{\overline{P} - P_i}{\overline{P}}, \\ m_L(\{\neg\text{stolen}_i\}) &= 0, \\ m_L(\Theta_i) &= 1 - v_L \frac{\overline{P} - P_i}{\overline{P}}, \end{aligned} \tag{10}$$

for $P_i \leq \overline{P}$, and

$$\begin{aligned} m_L(\{\text{stolen}_i\}) &= 0, \\ m_L(\{\neg\text{stolen}_i\}) &= v_{La} \frac{P_i - \overline{P}}{P_i}, \\ m_L(\Theta_i) &= 1 - v_{La} \frac{P_i - \overline{P}}{P_i}, \end{aligned} \tag{11}$$

for $P_i > \overline{P}$, where v_L is the weight of this evidence. We can intuitively read this weight as a reliability of this evidence; the same is valid for v_{La}; P_i is the price at which the seller i sells certain goods. $\overline{P}$ is the average price of the same goods offered through online auction system.

With this equation, we have expressed our belief that the lower the price of goods offered by seller i compared to the average price of respective goods is, the higher the suspicion that the seller offers stolen goods. We also assume that the equation reflecting the offering of legitimate goods does not show that the seller does not offer "stolen" goods; that is, $m_L(\{\neg\text{stolen}_i\}) = 0$.

Goods Sold Mostly at Fixed Price. The sellers (fraudsters) want to sell their stolen goods as quickly as possible. They want to dispose of them quickly and easily. Therefore, they prefer to sell the goods at a fixed price (Internet auction systems have the option "*buy now*"). It is the fastest way to sell goods on an online auction. When a bidder purchases goods at a fixed price, the auction ends, and the seller does not have to wait for the end of the auction. The belief functions have the following forms:

$$\begin{aligned} m_F(\{\text{stolen}_i\}) &= v_F \frac{N_{Fi}}{N_i}, \\ m_F(\{\neg\text{stolen}_i\}) &= 0, \\ m_F(\Theta_i) &= 1 - v_F \frac{N_{Fi}}{N_i}, \end{aligned} \tag{12}$$

where N_{Fi} is the number of goods sold by seller i for the fixed price; N_i is the total number of goods sold by this seller. It is valid that the higher the number of goods sold at fixed price, compared to the total number of goods sold by this seller i, the higher the suspicion that this seller sells "stolen" goods. Therefore, we also assume that the presented equation does not indicate that the seller does not sell stolen goods; that is, $m_F(\{\neg\text{stolen}\}) = 0$. The parameter v_F is in these equations the weight of evidence. We can intuitively interpret this parameter as the reliability of respective evidence.

A Variety of Goods Being Sold. Let us suppose that the seller sells stolen goods in online auctions. He sells any kinds of goods that he "gets." The variety of goods being sold is then higher than that of the average proper seller. The belief functions of this attribute have the following forms:

$$\begin{aligned} m_V(\{\text{stolen}_i\}) &= v_V \frac{V_i - \overline{V}}{V_i}, \\ m_V(\{\neg\text{stolen}_i\}) &= 0, \\ m_V(\Theta_i) &= 1 - v_V \frac{V_i - \overline{V}}{V_i} \end{aligned} \tag{13}$$

for $V_i \geq \overline{V}$ and

$$\begin{aligned} m_V(\{\text{stolen}_i\}) &= v_{Va} \frac{\overline{V} - V_i}{\overline{V}}, \\ m_V(\{\neg\text{stolen}_i\}) &= 0, \\ m_V(\Theta_i) &= 1 - v_{Va} \frac{\overline{V} - V_i}{\overline{V}} \end{aligned} \tag{14}$$

for $V_i < \overline{V}$, where V_i is the amount of different types of goods sold by seller i and $\overline{V}$ is the average amount of different types of goods sold by proper sellers in a respective category. The v_V and v_{Va} parameters are the weights of evidence. We can intuitively interpret this weight as the reliability. It is valid that the higher the variety of goods the seller i sells, compared to the average types of goods sold by proper seller, the higher the suspicion that this seller i sells stolen goods. Therefore, we assume that the given equation does not indicate that the seller does sell stolen goods; that is, $m_V(\{\neg\text{stolen}_i\}) = 0$.

Low Starting Price. If the starting price of an auction is lower than normal, it may indicate that a seller (fraudster) wants to sell their stolen goods as quickly as possible. They want to dispose of them quickly and they want surely to sell their goods. For the same reason, if the starting price of an auction is higher than normal, the auction may have less chance of selling stolen goods. The belief functions have the following forms:

$$\begin{gathered} m_P\left(\{\text{stolen}_i\}\right) = v_P \frac{\overline{\text{SP}} - \text{SP}_i}{\overline{\text{SP}}}, \\ m_P\left(\{\neg\text{stolen}_i\}\right) = 0, \\ m_P\left(\Theta_i\right) = 1 - v_P \frac{\overline{\text{SP}} - \text{SP}_i}{\overline{\text{SP}}} \end{gathered} \tag{15}$$

for $\text{SP}_i \leq \overline{\text{SP}}$ and

$$\begin{gathered} m_P\left(\{\text{stolen}_i\}\right) = 0, \\ m_P\left(\{\neg\text{stolen}_i\}\right) = v_{Pa} \frac{\text{SP}_i - \overline{\text{SP}}}{\text{SP}_i}, \\ m_P\left(\Theta_i\right) = 1 - v_{Pa} \frac{\text{SP}_i - \overline{\text{SP}}}{\text{SP}_i} \end{gathered} \tag{16}$$

for $\text{SP}_i > \overline{\text{SP}}$.

Here, v_P is the weight of this evidence. We can intuitively read this weight as a reliability of this evidence; the same is valid for v_{Pa}; SP_i is the price at which the seller i starts to sell certain goods (starting price). $\overline{\text{SP}}$ is an average starting price of the same goods.

With this equation, we have expressed our belief that the lower the staring price of goods offered by seller i compared to the average price of respective goods is, the higher the suspicion that the seller offers stolen goods (see (15)). If an auction's starting price is higher than average, this might indicate that the seller wants to get a good price for her goods and that they are able to wait in order to get a buyer who is willing to pay a good price or maybe even to repeat the auction. Therefore, the higher the starting price, the more possibility the seller acts as a proper seller (see (16)). In contrast, if an auction's starting price is lower than average, it is possible that the seller sells stolen goods.

Combination of Characteristic Signs (Evidences) of Fraudulent Behavior. A single characteristic is not enough to identify fraudulent behavior. Thus, once we have obtained more belief functions expressing our belief regarding fraudulent behavior, we combine them in a consistent manner to get a more complete assessment of what the whole group of evidences indicates. The combination of belief functions is done with the help of the Dempster combination rule (4). We express the assumption that a given seller i sells stolen goods with the help of belief function $m(\{\text{stolen}_i\})$. We calculate the value $\text{belief}(\{\text{stolen}_i\}) = m(\{\text{stolen}_i\})$ using the combination of single belief functions expressing appropriate evidence:

$$m\left(\{\text{stolen}_i\}\right) = \left(m_L \oplus m_F \oplus m_V \oplus m_P\right)\left(\{\text{stolen}_i\}\right). \tag{17}$$

The operator $\oplus$ is the Dempster's rule of belief function combination (see (4)). We perform the combination of multiple evidences according to the Dempster rule; first, we combine two belief functions $\text{bel}(\cdot)$; then we combine the result with the third belief function, fourth belief function, and so forth. For example, the following rules combine the first and second belief functions:

$$\begin{aligned} &\left(m_L \oplus m_F\right)\left(\{\text{stolen}_i\}\right) \\ &\quad = \frac{1}{K}\left[m_L\left(\{\text{stolen}_i\}\right) m_F\left(\{\text{stolen}_i\}\right)\right. \\ &\qquad + m_L\left(\{\text{stolen}_i\}\right) m_F\left(\Theta\right) \\ &\qquad \left. + m_L\left(\Theta\right) m_F\left(\{\text{stolen}_i\}\right)\right], \\ &\left(m_L \oplus m_F\right)\left(\{\neg\text{stolen}_i\}\right) = \frac{1}{K}\left[m_L\left(\Theta\right) m_E\left(\Theta\right)\right], \end{aligned} \tag{18}$$

where $K = 1 - (m_L(\{\neg\text{stolen}_i\})m_F(\{\text{stolen}_i\}) + m_L(\{\text{stolen}_i\})m_F(\{\neg\text{stolen}_i\}))$.

The value of $\text{bel}(\text{stolen}_i)$ indicates the degree of credibility of stolen_i.

The Definition of the Reinforcement Coefficient α. We explored various sources of information on the Internet (on Internet forums, discussions, etc.) that could potentially serve as a source of contextual information and could be used to improve the detection of fraud. We extracted pieces of information from these information sources and examined them in detail within the context of online auctions conducted on the Aukro auction portal [2]. For example, we found complaints discussed in various Internet forums and discussions. For example, some users on these forums complained that somebody had stolen from them a certain object and this object appeared later on an Internet auction. An example is the internet forum [43]. Here, a user mentions the theft of car navigation device Columbus. At the same forum [43], other users inform him that they recently saw this navigation device offered for sale on the Internet auction portal Aukro. Further conversation on this forum relates to whether it could be the stolen device mentioned by the first user. Some users say that it is for certain the stolen device because (a) this particular navigation is factory-fitted, (b) the price in the online auction is suspiciously low, and (c) the online auction in which this device has been sold started few hours after the theft. This case belonged to the five mentioned prosecuted cases from the complaints of selling stolen goods in Internet auctions.

We performed analysis of about 3,841 different Internet forums and found out that similar conversations occurred in 84 cases (see Section 5.1) which we tried to connect to Internet auctions conducted on the Aukro online auction portal (see Section 5.1). It is clear that every theft is not mentioned on an Internet forum. However, when we find out information about the theft in Internet forums or discussions, we consider this information as contextual information that can help identify fraudulent behavior, selling stolen goods in the internet auction.

We have defined the reinforcement coefficient α as follows: we find out that the stolen object appears in an online auction a short time (some hours) after the theft occurred. The thieves want to dispose of stolen goods as soon as possible. We have expressed the reinforcement coefficient α as a function of time:

$$\alpha = \begin{cases} Ke^{-kt} & \text{if time of the theft is specified} \\ 0 & \text{if time of the theft is not specified,} \end{cases} \tag{19}$$

where the values of K and k are coefficients that are to be determined on the basis of statistical analysis and t is the difference between the start time of an online auction with goods which have been discussed as stolen on some Internet forum and the time of publication of the complaint on the respective Internet forum. We found out that the person who reports the theft usually provides the approximate time information as well. In the event that the time information about the theft is not specified, the α value is set to 0. The α parameter has no effect on respective information about selling stolen goods in the internet auction in this case.

4.2. Influence of Contextual Information on the Evidence about Fraudulent Behavior. Let us suppose that the chosen attributes of an online auction have been examined. The calculated belief assignment $m(\{\text{stolen}_i\})$ using (17) of this auction indicates that the seller may be selling a stolen good. At the same time, we find out that one of the users of an Internet forum complained that the same good was stolen from him. We consider this information as additional context information that reinforces our belief that stolen goods are sold within this auction. This reflects the fact that part of BBM is transferred from the total frame of discernment $m(\Theta)$, denoting our uncertainty, into the element $m(\{\text{stolen}_i\})$. This operation reflects our belief related to the sale of stolen goods (we denote this operation as dediscounting or reinforcement; see (8)). We can calculate the resulting belief about fraudulent behavior $m_R(\text{stolen}_i)$ by using the following equation:

$$m_R(\text{stolen}_i) = \frac{m(\text{stolen}_i)}{1-\alpha}. \tag{20}$$

We have calculated the belief concerning the certain fraudulent online auction (stolen goods are sold within this auction) $m_R(\{\text{stolen}_i\})$ with the use of additional contextual information available from Internet sources. This additional information reinforces our confidence that stolen goods are sold within this auction. On the other side, our uncertainty concerning analyzed online auction that stolen goods are sold here will decrease.

4.3. Categorization of Sellers according to the Resulting Belief Function Representing the Behavior of Selling of Stolen Goods. We will divide users into categories according to the degree of belief that a certain user i sells stolen goods, that is, $\text{bel}_R(\{\text{stolen}_i\})$. These categories are "proper seller," "suspect seller," and "seller selling stolen goods."

We define two thresholds η and ξ. The first threshold η is the threshold for determining whether a seller i is a proper seller. If the value of $\text{bel}_R(\{\text{stolen}_i\})$ is below η, the seller i will be considered a proper seller. The second threshold ξ is the threshold for determining that a seller i sells stolen goods. If the value of $\text{bel}_R(\{\text{stolen}_i\})$ exceeds ξ, the respective seller i will be considered a seller selling stolen goods. If the value of $\text{bel}_R(\{\text{stolen}_i\})$ is between η and ξ, we will consider the seller "only" suspect of selling stolen goods. The thresholds η and ξ will be qualified on the basis of statistical evaluations of analyzed auctions.

The Schema of the Framework Based on Proposed Model. Major task of this framework is to identify sellers who sell stolen goods and identify proper sellers. Figure 1 depicts the schema of framework based on proposed model.

Sellers are evaluated mathematically using a data fusion method that combines information from different information sources on the Internet and auction-level features.

The threshold ξ of certifying sellers as seller selling stolen goods should be fairly high to reduce the number of false positives. For the sellers that are certified as suspect, the values of their $\text{bel}_R(\{\text{stolen}_i\})$ must be lower than ξ but greater than the values of threshold η (see Table 1). This means that the evidence is not sufficient enough to support a conclusion that a seller sells stolen goods, even though the seller behaved more like a seller selling stolen goods than an honest proper seller. As a result, the seller is considered suspect. When new additional independent evidence is presented, the suspect certification will be revalidated. If a seller's certification changes, the new certification is committed to the database. If a seller's certification is labeled as seller selling stolen goods, the seller is subject to further investigation and possible punishment, but this step is outside the scope of this paper.

5. Case Study and Result Analysis

The original motivation for our study was an actual case with which we were familiar: a valuable specific brand-name radio was stolen from a car. The radio owner mentioned the theft including the information about the time of the event on an Internet forum [43]. Others members of this forum reported, as reaction to this information, that an auction selling the same type of radio was taking place at the same time. The radio owner took legal steps then and the perpetrator of the theft was apprehended as the police cooperated with the operator of the online auction portal. Based on the experience from this case, we started a detailed examination of various Internet forums and online auction site aukro.cz acting in the Czech Republic.

5.1. Description of the Data Set. We created a simple crawler and searched systematically the Czech Internet forums and

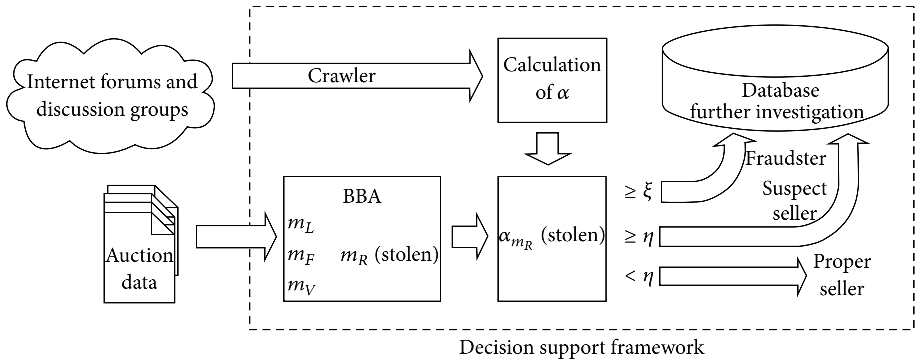


Figure 1: Schema of framework based on proposed model.

Table 1: Classification assignment rules.

$\mathrm{bel}_R(\{\mathrm{stolen}_i\}) \leq \eta$	Proper seller
$\eta < \mathrm{bel}_R(\{\mathrm{stolen}_i\}) < \xi$	Suspect seller
$\mathrm{bel}_R(\{\mathrm{stolen}_i\}) \geq \xi$	Seller selling stolen goods

discussion groups starting on January 15, 2013, and ending April 18, 2014. The aim was to find notices about theft on these forums. We explored a total of 3,841 different Internet forums and discussions groups. We found out that word theft and similar expressions (i.e., equivalents like thievery, burglary, abstraction, pilferage, and robbery) occurred at 416 of them. We performed a detailed analysis of them and we chose the 84 (including complaints examined by Czech Trade Inspection, see Section 4.1) of them where we assumed, based on reaction of members of Internet forums or Internet discussion groups, that the stolen goods may be offered for sale on online auctions.

Along with the analysis of Internet forums and discussions, we downloaded systematically all the information on the auctions that were posted on the aukro.cz site. We specially carried out detailed analyses of online auctions in which goods mentioned on scanned Internet forums and discussion groups were being sold. In total, we got information from 30,059 auctions. Each auction had a unique seller and one or more bidders. We had information on unique ID numbers of all sellers and bidders in all scanned auctions (see sample data in Table 3 and the example of data in Table 2).

5.2. Data Processing. We were interested in activities of individual sellers. For each auction we knew its bidding history of auctions, prices, goods sold within these auctions, and history of seller's transactions. Based on information from auctions, we constructed average prices of monitored goods or average amount of goods sold by sellers in respective category. We investigated auctions hosted by chosen sellers during the monitored period and other auctions as well with the aim to find out the number of types of sold goods in various categories. The historical pieces of information were statistically processed.

The basic masses assigned for evidence, in which stolen goods were being sold, were calculated using (10)–(17). The result of calculations are presented in Table 5. The weights of evidence v_L, $v_{La}v_F$, v_V, v_{Va}, v_P, and v_{Pa} were set 0.9, 0.9, 0.7, 0.8, 0.8, 0.85, and 0.85 in agreement with our experiments. We consider the character "Inadequate low price" as the most predicative. The character "Goods sold at fixed prices" is the less reliable in determining a seller selling stolen goods.

The value t is the difference between the time when the online auction selling goods which have been mentioned as stolen on some Internet forum began and the time of publication of this complaint on the respective Internet forum. The value of k (19) was set on the basis of our experiments at 0.65.

The values α and $m_R(\{\mathrm{stolen}_i\})$ in Table 6 were calculated using (2)–(4). The value $m_R(\{\mathrm{stolen}_i\})$ expresses our belief that the seller sells stolen goods. On the other side, values $\mathrm{Pl}_R(\{\neg\mathrm{stolen}_i\})$ represent our belief that the seller is the proper seller including our uncertainty or rather ignorance concerning the classification whether the seller i sells legal goods or the seller i sells stolen goods. Hence, in this case, we gave our doubt (uncertainty) concerning the seller's honesty on the side of their innocence.

The value of threshold ξ was set on the basis of our experiments at 0.85 and the value of threshold η at 0.75. The threshold ξ was set sufficiently high to reduce the number of false positives. These classification results are assigned in accordance with the following rules:

if $\mathrm{Bel}_R(\{\mathrm{stolen}_i\}) \leq 0.75$, then we consider seller as proper seller;

if $0.75 < \mathrm{Bel}_R(\{\mathrm{stolen}_i\}) < 0.85$, then we consider seller as suspect seller;

if $\mathrm{Bel}_R(\{\mathrm{stolen}_i\}) \geq 0.85$, then we consider seller as seller who sells stolen goods.

5.3. Analysis and Discussion. We now analyze the auction data and the certification results by considering three levels of certification. As we mentioned above, we explored a total of 30,059 auctions and 3,841 different Internet forums and discussions groups. We performed a detailed analysis of 416 of them and then chose 84 for more detailed examination. In this paper, we present a sample of our results in Tables 3, 4, 5, and 6. We now discuss various sellers with different certifications.

Table 2: Sample record.

(a)

RECORD Radio car brand Alpine CDE-133BT								
Auction	Category	Starting price (CZK)	Number of bids	Number of bidders	Start time	End time	Winning price (CZK)	Fixed price (Y/N)
	Audio (radio)	500	5	3	January 21, 2013 21:08	January 23, 2013 18:35	1500	Y

(b)

Historical statistics	Average price of these goods	Average number of bids	Average feedback	Note
	2525	4		

(c)

Seller info	Seller ID	Num. of auctions	Bids/Auction	Num. of categories	Feedback (percentage)	Active since	Note
	D***r	2	2	2	2 (100%)	Jan-5-13	

(d)

Contextual information		
Link	Time of publication of the complaint on the respective Internet forum	Note
http:/kradene.cz/ukradene/88/Autoradio-Alpine-CDE-133BT	Jan-20-13 15:00 (found out via telephone inquiry)	

Table 3: Sample data collected from Czech online auction Aukro.

Seller *i*	Price at which seller *i* sold certain goods (in CZK)	Average price of these goods throughout the online auction system (in CZK)	Number of goods which the seller *i* sells at fixed price	Total number of goods sold by seller *i*	Average starting price of certain goods (in CZK)	Starting price of these goods sold by seller *i* (in CZK)	Number of different type of goods offered by seller *i*	Average number of different types of sold goods sold by sellers in respective category
D***r	1500	2525	2	2	650	450	2	2
O***2	700	1850	1	1	300	300	1	2
m***k	1250	1600	1	2	700	500	2	2
d***l	650	750	2	6	500	500	6	2
2***j	1420	1540	1	7	800	400	5	2
b***s	1200	1450	0	8	750	500	5	2
k***J	800	1050	1	3	500	450	3	2
D***r	950	1100	3	11	600	650	2	2
s***m	750	600	0	2	100	150	3	2
b***n	1320	1500	1	8	750	550	4	2
n***k	1800	2200	2	3	1000	850	2	2
n***2	890	850	1	7	600	500	4	2

Seller Selling Stolen Goods. According to our model, the sellers D***r, O***2, and d***l are sellers who sell stolen goods. Using (10)–(17), the degrees of belief from available evidence are combined to obtain the joint belief of shill. The influence of contextual information is expressed by (20). The belief function Bel of a seller selling stolen goods for D***r is 0.894767, which is greater than $\xi(0.85)$; thus, the model evaluates the certification of D***r to a seller who sells stolen goods. The same conclusion applies to the seller O***2. We also had confirmations in these cases. These cases were addressed by the Czech Police. The investigation confirmed that these sellers actually sold stolen goods. Similarly, the seller d***l was certified as a seller who sells stolen goods. In this case, we do not have some explicit acknowledgment. But we can see the selling pattern of this seller in Table 3. Similarly, like sellers D***r and O***2, the aim of this seller was to sell as much stolen goods as fast as possible. He preferred fix price and the starting price was also set very low. Also, the connection between the complaints on the Internet and auctions was very tight in this case.

Table 4: The basic masses assigned to single "selling stolen goods" characteristics (10), (11), (12), and (13).

Seller i	$m_L(\{\text{stolen}_i\})$	$m_L(\Theta)$	$m_F(\{\text{stolen}_i\})$	$m_F(\Theta)$	$m_V(\{\text{stolen}_i\})$	$m_V(\Theta)$	$m_P(\{\text{stolen}_i\})$	$m_P(\Theta)$
D***r	0.365347	0.634653	0.7	0.3	0	1	0.261538	0.738462
O***2	0.559459	0.440541	0.7	0.3	0	0.6	0	1
m***k	0.196875	0.803125	0.35	0.65	0	1	0.242857	0.757143
d***l	0.12	0.88	0.233333	0.766667	0.533333	0.466667	0	1
2***j	0.07013	0.92987	0.1	0.9	0.48	0.52	0.425	0.575
b***s	0.155172	0.844828	0	1	0.48	0.52	0.283333	0.716667
k***J	0.214286	0.785714	0.233333	0.766667	0.266667	0.733333	0.085	0.915
D***r	0.122727	0.877273	0.190909	0.809091	0	1	0	0.934615
s***m	0	0.82	0	1	0.266667	0.733333	0	0.716667
b***n	0.108	0.892	0.0875	0.9125	0.4	0.6	0.226667	0.773333
n***k	0.163636	0.836364	0.466667	0.533333	0	1	0.1275	0.8725
n***2	0	0.959551	0.1	0.9	0.4	0.6	0.141667	0.858333

Table 5: The influence of contextual information on belief function (19) and (20).

Seller i	$m(\{\text{stolen}_i\})$	$m(\{\neg\text{stolen}_i\})$	$m(\Theta)$	t (hours)	α	$m_R(\{\text{stolen}_i\})$	$m_R(\{\neg\text{stolen}_i\})$	$m_R(\Theta)$
D***r	0.8594	0	0.1406	28	0.039527	0.894767	0	0.105233
O***2	0.797566	0.080974	0.121461	21	0.079597	0.866539	0.087976	0.045484
m***k	0.604748	0	0.395252	299	$6.72E-14$	0.604748	0	0.395252
d***l	0.685156	0	0.314844	12	0.195776	0.851946	0	0.148054
2***j	0.749772	0	0.250228	38	0.014541	0.760835	0	0.239165
b***s	0.685161	0	0.314839	42	0.009747	0.691905	0	0.308095
k***J	0.595802	0	0.404198	28	0.039527	0.620322	0	0.379678
D***r	0.276478	0.047307	0.676215	148	$2.43E-07$	0.276478	0.047307	0.676215
s***m	0.176071	0.339733	0.484196	26	0.048278	0.185003	0.356967	0.45803
b***n	0.622327	0	0.377673	12	0.195776	0.773823	0	0.226177
n***k	0.610812	0	0.389188	18	0.107444	0.684341	0	0.315659
n***2	0.526218	0.019164	0.454617	22	0.072022	0.567059	0.020652	0.412289

Suspect Seller. For suspects, the evidence is not sufficient enough to support the sellers as sellers selling stolen goods, but it is still more sufficient than evidence that supports ¬stolen (we express this support here by the value $\text{Pl}_R(\{\neg\text{stolen}_i\})$). In the case study, the system gives the certification of suspect seller to two sellers. They are sellers 2***j and b***n. We first justify bidder 2***j. This seller has very low starting price and also the connection between the complaints on the internet and auctions was very tight. Hence, behavior of this seller matches significant pattern of selling stolen goods. Similarly, the characteristics of the seller b***n are not so important after all. Here, the close connection between the complaints on the Internet and auctions of this seller is what mainly makes this seller suspect.

Proper Seller. We show what kind of sellers is considered honest. Without much doubt, D***r and s***m are not sellers who would sell stolen goods. The starting prices of goods offered by these sellers are even higher than the average starting price. The plausibility $\text{Pl}_R(\{\neg\text{stolen}_i\})$ reaches values higher than 0.7. The values of $\text{Bel}_R(\{\text{stolen}_i\})$ of these sellers are much smaller than the threshold value $\eta(0.75)$. Now, we consider remaining proper sellers. The values of $m_R(\{\text{stolen}_i\})$ of the sellers m***k, b***s, k***J, n***k, and n***2 are less than the threshold η. They are considered to be proper sellers. We found nothing to suggest that their behavior on the auction does not correspond to the average behavior pattern.

The system would operate in real conditions as follows: when value of $m_R(\{\text{stolen}_i\})$ for some seller is greater than the threshold ξ, then we consider him/her as seller who sells stolen goods. This seller must then be subject to further investigation to confirm or refute our belief. If value of $m_R(\{\text{stolen}_i\})$ for some seller is greater than the threshold η but less than threshold ξ, we "only" suspect this seller that they sell stolen goods. It is further recommended to monitor the behavior of these sellers. If the value of $m_R(\{\text{stolen}_i\})$ for some seller is less than the threshold η, then this seller is considered to be proper seller. Their behavior on the auction corresponds to the average behavior pattern.

6. Conclusion and Future Work

Based on the conceptual framework of Dempster-Shafer theory, a practical approach for detection of a specific type of fraudulent behavior (selling stolen goods) has been proposed. This method in essence takes into consideration evidence found from different information sources, in this case, from

Table 6: Categorization of sellers according to the resulting belief function.

Seller i	$\mathrm{Bel}_R(\{\mathrm{stolen}_i\})$	$\mathrm{Pl}_R(\{\neg\mathrm{stolen}_i\})$	Result
D***r	0.894767	0.105233	Seller selling stolen goods
O***2	0.866539	0.133461	Seller selling stolen goods
m***k	0.604748	0.395252	Proper seller
d***l	0.851946	0.148054	Seller selling stolen goods
2***j	0.760835	0.239165	Suspect seller
b***s	0.691905	0.308095	Proper seller
k***J	0.620322	0.379678	Proper seller
D***r	0.276478	0.723522	Proper seller
s***m	0.185003	0.814997	Proper seller
b***n	0.773823	0.226177	Suspect seller
n***k	0.684341	0.315659	Proper seller
n***2	0.567059	0.432941	Proper seller

online auction system and from Internet information sources. Pieces of knowledge about auctions including bidding behavior were processed and quantified. Using the Dempster rule of combination, we combined evidence that enforces each other and resolved the conflicts between different pieces of evidence. We also took into account contextual information from various Internet sources. We consider this information as a factor that can influence (reinforce) our belief concerning the analyzed fraudulent behavior. The case study shows that our proposed approach is quite accurate and practical for real world deployment.

We are convinced that the presented approach represents a promising line of research. Similar methods and systems can be used in particular by auction portals that can monitor suspicious auctions and then provide warnings to users to simply pay attention to what they buy. In addition, law enforcement may use the system to investigate reports of fraud.

In future research, we plan to design a fraud detection agent with self-learning capability so that the information from various sources used in our approach can be gained and integrated automatically. We believe that the belief function theory can provide a practical approach and enhance system performance for fraud detection in online auctions.

7. Limitations of Our Study

We have mentioned our motivation when creating the mathematical model described above. We researched specific complaint, which was described on an Internet forum [43]. Here, the user mentioned the theft of a car navigation device Columbus. We traced this case to its end solution. Indeed, it was confirmed during police investigation that the device was really sold on the Aukro online auction. The case resulted in indictment of seller. Furthermore, we know about some other similar cases.

These events motivated us to build a mathematical model for the detection of sale of stolen goods on online auctions. However, the exact validation of suggested model is very difficult. Generally, it is not easy to ascertain whether a particular item is actually stolen or not. Thus, the detection abilities of our model are hereby limited. The model can give only indication that some goods, sold by a particular seller on online auction, may be stolen. It means that this indication (result) of the model must be verified some other way. This may include some thorough investigation or monitoring of the respective user. Hence, this way, the model can be beneficial for auction operators whose aim is to ensure honest behavior of all users on online auctions.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

References

[1] eBay, 2013, http://www.ebay.cz/.

[2] Aukro, 2013, http://aukro.cz/.

[3] B. Gavish and C. L. Tucci, "Fraudulent auctions on the Internet," *Electronic Commerce Research*, vol. 6, no. 2, pp. 127–140, 2006.

[4] D. G. Gregg and J. E. Scott, "A typology of complaints about eBay sellers," *Communications of the ACM*, vol. 51, no. 4, pp. 69–74, 2008.

[5] D. Harrison McKnight, V. Choudhury, and C. Kacmar, "The impact of initial consumer trust on intentions to transact with a web site: a trust building model," *Journal of Strategic Information Systems*, vol. 11, no. 3-4, pp. 297–323, 2002.

[6] D. G. Gregg and J. E. Scott, "The role of reputation systems in reducing on-line auction fraud," *International Journal of Electronic Commerce*, vol. 10, no. 3, pp. 95–120, 2006.

[7] Y. Hu and B. Panda, "Design and analysis of techniques for detection of malicious activities in database systems," *Journal of Network and Systems Management*, vol. 13, no. 3, pp. 269–291, 2005.

[8] Y. Ku, Y. Chen, and C. Chiu, "A proposed data mining approach for internet auction fraud detection," in *Intelligence and Security Informatics*, vol. 4430 of *Lecture Notes in Computer Science*, pp. 238–243, Springer, New York, NY, USA, 2007.

[9] S. Pandit, D. H. Chau, S. Wang, and C. Faloutsos, "Netprobe: A fast and scalable system for fraud detection in online auction

networks," in *Proceedings of the 16th International World Wide Web Conference*, pp. 201–210, Banf, Canada, May 2007.

[10] B. Zhang, Y. Zhou, and C. Faloutsos, "Toward a comprehensive model in internet auction fraud detection," in *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS '08)*, p. 79, IEEE, Waikoloa, Hawaii, USA, January 2008.

[11] G. Shafer, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, USA, 1976.

[12] L. Beranek and J. Knizek, "The use of contextual information to detection of fraud on online auctions," *The Journal of Internet Banking and Commerce*, vol. 18, no. 3, pp. 1–18, 2013.

[13] Bryant and C. Colledge, "Trust in electronic commerce business relationships," *Journal of Electronic Commerce Research*, vol. 3, no. 2, pp. 32–39, 2002.

[14] D. L. Hoffman, T. P. Novak, and M. Peralta, "Building consumer trust online," *Communications of the ACM*, vol. 42, no. 4, pp. 80–85, 1999.

[15] L. C. Hsu and C. H. Wang, "A study of e-trust in online auctions," *Journal of Electronic Commerce Research*, vol. 9, pp. 310–321, 2008.

[16] B. Lee, H. Cho, M. Chae, and S. Shim, "Empirical analysis of online auction fraud: credit card phantom transactions," *Expert Systems with Applications*, vol. 37, no. 4, pp. 2991–2999, 2010.

[17] C. Chua and J. Wareham, "Self-regulation for online auctions: an analysis," in *Proceedings of the 23rd International Conference on Information Systems (ICIS '02)*, pp. 115–125, December 2002.

[18] Y. Sakurai and M. Yokoo, "A false-name-proof double auction protocol for arbitrary evaluation values," in *Proceedings of the 2nd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS '03)*, pp. 329–336, July 2003.

[19] 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.

[20] 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. Gupta and R. Sharman, Eds., pp. 446–470, Information Science Reference, 2009.

[21] J. Trevathan and W. Read, "Secure online english auctions," in *E-Business and Telecommunication Networks*, vol. 9 of *Communications in Computer and Information Science*, pp. 119–133, Springer, Berlin, Germany, 2008.

[22] J. Trevathan, "Security, anonymity and trust in electronic auctions," *Crossroads*, vol. 11, pp. 3–9, 2005.

[23] V. Da Silva Almendra and D. Schwabe, "Analysis of fraudulent activity in a Brazilian auction site," in *International World Wide Web Conference (WWW '09)*, ACM, Madrid, Spain, 2009.

[24] J. Trevathan and W. Read, "Detecting collusive shill bidding," in *Proceedings of the 4th International Conference on Information Technology-New Generations (ITNG '07)*, pp. 799–808, IEEE Computer Society, Las Vegas, Nev, USA, April 2007.

[25] 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.

[26] D. H. Chau and C. Faloutsos, "Fraud detection in electronic auction," in *Proceedings of EWMF'05: European Web Mining Forum (ECML/PKDD '05)*, pp. 87–97, October 2005.

[27] J. Wang and C. Q. Chiu, "Detecting online auction inflated-reputation behaviors using social network analysis," in *Proceedings of North American Association for Computational Social and Organizational Science*, Notre Dame, Indiana, 2005.

[28] D. H. Chau, S. Pandit, and C. Faloutsos, "Detecting fraudulent personalities in networks of online auctioneers," in *Knowledge Discovery in Databases: PKDD 2006*, vol. 4213 of *Lecture Notes in Computer Science*, pp. 103–114, Springer, Berlin, Germany, 2006.

[29] H. Xu and Y. Cheng, "Model checking bidding behaviors in internet concurrent auctions," *Computer Systems Science & Engineering*, vol. 22, no. 4, pp. 179–191, 2007.

[30] B. J. Ford, H. Xu, and I. Valova, "Identifying suspicious bidders utilizing hierarchical clustering and decision trees," in *Proceedings of the International Conference on Artificial Intelligence (ICAI '10)*, vol. 2, pp. 195–201, July 2010.

[31] A. Goel, H. Xu, and S. M. Shatz, "A multi-state Bayesian network for shill verification in online auctions," in *Proceedings of the 22nd International Conference on Software Engineering and Knowledge Engineering (SEKE '10)*, pp. 279–285, July 2010.

[32] 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.

[33] L. Beranek, P. Tlusty, and R. Remes, "The usage of belief functions for an online auction reputation model," in *Proceedings of the 28th Conference on Mathematical Method in Economy*, pp. 49–54, September 2010.

[34] L. Beranek and J. Knizek, "The usage of contextual discounting and opposition in determining the trustfulness of users in online auctions," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 7, no. 1, pp. 34–50, 2012.

[35] W.-H. Chang and J.-S. Chang, "A novel two-stage phased modeling framework for early fraud detection in online auctions," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11244–11260, 2011.

[36] W.-H. Chang and J.-S. Chang, "A Multiple-Phased modeling method to identify potential fraudsters in online auctions," in *Proceedings of the 2nd International Conference on Computer Research and Development (ICCRD '10)*, pp. 186–190, Kuala Lumpur, Malaysia, May 2010.

[37] W.-H. Chang and J.-S. Chang, "Using clustering techniques to analyze fraudulent behavior changes in online auctions," in *Proceedings of the International Conference on Networking and Information Technology (ICNIT '10)*, pp. 34–38, Manila, Philippines, June 2010.

[38] W.-H. Chang and J.-S. Chang, "An online auction fraud screening mechanism for choosing trading partners," in *Proceedings of the 2nd International Conference on Education Technology and Computer (ICETC '10)*, vol. 5, pp. V5-56–V5-60, Shanghai, China, June 2010.

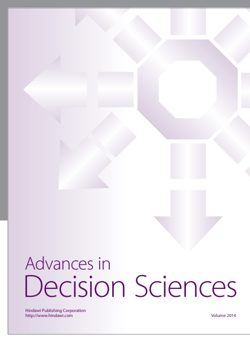
[39] W. H. Chang and J. S. Chang, "An effective early fraud detection method for online auctions," *Electronic Commerce Research and Applications*, vol. 11, no. 4, pp. 346–360, 2012.

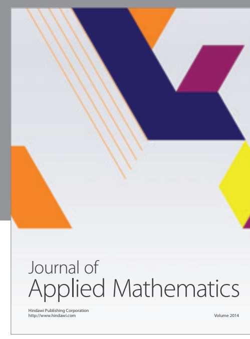
[40] M. Y. K. Kwan, R. E. Overill, K. P. Chow et al., "Evaluation of evidence in internet auction fraud investigations," in *Advances in Digital Forensics VI*, vol. 337 of *IFIP Advances in Information and Communication Technology*, pp. 121–132, Springer, Berlin, Germany, 2010.

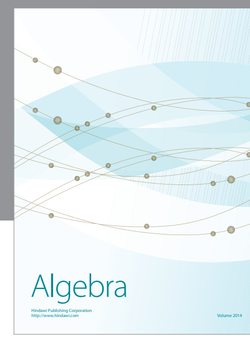
[41] P. Smets, "Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem," *International Journal of Approximate Reasoning*, vol. 9, no. 1, pp. 1–35, 1993.

[42] Czech e-shops inspection, Kontrola e-shop v roce 2012: Vedou nekalé obchodn praktiky, Oversito.cz, 2013, (Czech), http://www.oversito.cz/aktualne/kontrola-e-shopu-v-roce-2012-vedou-nekale-obchodni-praktiky-92/.

[43] Stolen Columbus, "Odcizeni Columbuse," Forum.Skodahome.cz., (Czech), 2013, http://forum.skodahome.cz/topic/114051-odcizeni-columbuse/.







Hindawi

Submit your manuscripts at
http://www.hindawi.com