Control and Cybernetics

vol. **38** (2009) No. 1

Mining online auction social networks for reputation and recommendation*†

by

Mikołaj Morzy[1], **Adam Wierzbicki**[2] **and Apostolos N. Papadopoulos**[3]

[1] Institute of Computing Science, Poznań University of Technology
Piotrowo 2, 60-965 Poznań, Poland

[2] Polish-Japanese Institute of Information Technology
Koszykowa 86, 02-008 Warszawa, Poland

[3] Department of Informatics, Aristotle University
54124 Thessaloniki, Greece

e-mail: Mikolaj.Morzy@put.poznan.pl, adamw@pjwstk.edu.pl, papadopo@csd.auth.gr

Abstract: Online auctions are quickly becoming one of the leading branches of e-commerce. Unfortunately, online auctions attract many fraudulent activities. Reputation systems are crucial for guaranteeing fairness of trade and reliability of service. Currently used reputation systems offer little protection from malevolent contractors. In this paper we present a new method for mining the reputation of sellers in online auctions. We devise two independent measures that assess reliability and questionability of sellers in parallel, leading to the concept of positive and negative reputation. To compute these measures we construct an S-graph which reflects the social linkage between sellers and buyers. We use both explicit and implicit feedbacks provided by auction participants, carefully identifying missing feedbacks that have been purposefully left out. Based on reputation estimates the community of online auction participants can detect misbehaving contractors and counteract fraud. Thus, the application of social information about reputation of contractors can be perceived as recommendations. Experimental evaluation of our proposal proves the feasibility and usefulness of the presented approach.

Keywords: online auctions, reputation, recommendations.

---

*Research supported by the Polish Ministry of Science grants N N516 4307 33, 69/N-SINGAPUR/2007/0, and by the 2005-2007 Joint Research and Technology Program between Poland and Greece funded by GSRT, Greek Ministry of Development.

†Submitted: June 2008; Accepted: October 2008.

1. Introduction

E-commerce growth is far from slowing down. According to Forrester Research (Forrester, 2007), in 2006 online sales in the USA rose 25% year on year to $219 billion. Even if we exclude travel sales, the growth reaches $146 billion (29% year on year). This year e-commerce keeps growing at an astonishing pace of 18%, surging by $259 billion (and $174 billion excluding travel). The report estimates that online sales represent 6% of total retail sales in 2006. Today, 63% of online population have already engaged in e-commerce activities. Among them, 27% of online consumers have participated in auctions on the Web, leaving plenty of room for growth. Forrester Research reports that online auction markets generated $49 billion in sales in 2006, which represents 25% of online sales, but predicts further development. To quote the report: "*we expect online auction sales to reach steady average growth of 40%*". Auctions, which are one of the very first forms of economic activity known to mankind, are experiencing a triumphant come-back in the electronic form. The model describing online auctions is called customer-to-customer, and its validity and practical usability is well proved by the popularity of over 250 currently active online auction sites, such as `www.ebay.com`, `www.ubid.com`, `www.onsale.com`, and many others.

The numbers behind the biggest players are incredible. According to marketing research company Nielsen, eBay, the global leader in online auction market, attracts 50 million unique visitors per month. An article of clothing is sold on eBay every 3 seconds, a car is sold every 90 seconds, and 30 000 pieces of luxury jewelry are sold daily. The volume of annual transactions reached $23 billion and 400 000 people make their living through eBay. Over 95 millions of registered eBay users perform 5 million transactions each week, with 12 million items posted on eBay at any given point in time.

Apart from offering new and unprecedent possibilities, online auctions provide opportunities for dishonest participants to commit fraud. The lack of physical contact between trading partners significantly affects the level of trust. This can be observed in the results of a recent Eurobarometer poll, which reveals that 73% of customers refraining from e-commerce are motivated by the concerns about the security of payment. In addition, 37% of consumers are concerned about warranty terms and the quality of service. The number of complaints regarding online auctions has grown rapidly over the last few years. American Federal Trade Commission reports that 48% of all e-commerce related complaints involved fraud committed in online auctions, and the total loss caused by fraud was as high as $437 million in one year. National Consumers League reveals that 63% of complaints about Internet fraud concerned online auctions, and the average loss of $478 per person. Recently, FBI Internet Crime Complaint Center presented its annual report on Internet crime (IC3, 2007). In 2006, FBI IC3 received 207 492 complaints, among which 45% concerned online auctions. Total loss of internet crime victims, who reported, augmented to $200 million, with the average loss of a victim reaching \$724. The average loss of a victim of an online auction theft was slightly lower, reaching \$602 (for instance, the average loss of a victim who fell for the infamous Nigerian Letter scam was \$5100). The report clearly proves that online auction fraud exceeds by far other types of Internet crime, such as non-delivered merchandise and payment (19% of complaints) or check fraud (5% of complaints).

In this paper, we shall assume that no trusted third parties are present to aid auction participants. This assumption holds true in most online auction environments. In the absence of a trusted third party, online auctions involve a high level of risk. Participants are uncertain about the behavior of their peers. Reputation systems are among the few measures that can be used in such environment to aid the decision making of auction participants (Resnick, 2000) and to create incentives for fair behavior. Trust and fairness of the competition are perceived by auction participants as fundamental issues in developing a successful customer-to-customer market (Resnick, 2002). Furthermore, the reputation of sellers has an economically observable and statistically significant effect on the price of items (Houser, 2001). Based on reputation ratings the community of online auction participants can detect misbehaving contractors and counteract fraud. Thus, the application of social information about reputation of contractors can be perceived as recommendations.

Most information for reputation computation comes from mutual feedbacks provided by contractors upon the closing of a transaction. Yet, due to the nature of these feedback, currently used reputation systems suffer from severe limitations. Most systems calculate the so-called "indirect propagated reputation" (Mui, 2003) that relies on feedback from untrusted third parties. This feedback can be incorrect or missing, which is frequently the case in online auctions. Feedback can also be misused by malicious adversaries who execute coalition attacks (see Melnik, 2002, and Resnick, 2004) or Sybil attacks (Douceur, 2002) to artificially increase their reputation.

The successful use of reputation systems in online auctions requires a careful design of reputation algorithms. Currently used algorithms are ill-suited to handle complex characteristics of online auctions. Particularly, these algorithms calculate the reputation of all auction participants in the same way. Yet, in online auctions, buyers and sellers are exposed to very different levels of risk. Sellers usually receive advance payments, and are therefore exposed to little financial risk. On the contrary, buyers, who may not receive purchased goods (or receive goods of poor quality), are exposed to substantial risk. The quality of a reputation system for online auctions is therefore determined by the correctness of estimation of seller's reputation. It can also be observed that in an online auction system, buyers and sellers form a social network that can be exploited in the estimation of reputation.

Current trust management research has identified another concept that can be used to enhance reputation algorithms for online auctions. This is the concept of distrust (Marsh, 2005). Distrust is not an absence of trust, but a negative trust. It reflects the expectation that the distrusted agent will behave in a way that will adversely affect the welfare of the distrusting agent. In this work, we present one of the first attempts to use distrust in online auction reputation systems.

The contribution of this paper is threefold. First, a new method for calculating seller's reputation based on a social network of buyers and sellers is described. Second, the social network is enhanced by adding information hidden in feedbacks that were purposely left out. Third, buyers are allowed to use distrust and trust simultaneously in online auctions, because the system calculates two reputation values: positive and negative reputation. We test our reputation measure using a large body of the real world data acquired from the online auction site. The experiments prove the feasibility of our approach.

In this paper we present the seller rank, a novel reputation measure designed specifically for sellers in online auctions. In our research we draw inspiration from social network analysis. We transform the graph of seller-buyer connections into a directed graph, where nodes represent sellers and edge weights represent the degree of trust and distrust present between the sellers. Edge weights are determined based on feedbacks provided by buyers. During transformation we identify missing feedbacks and we decide, based on the history of a given user feedbacks, if the missing feedback has been omitted deliberately. Next, we apply the random walk procedure that allows us to determine the true importance of each node, which is the basis for computing the seller rank of each node. We present an algorithm based on the random walk paradigm. The first iteration of the algorithm computes the degree of trust of each seller, whereas the second iteration of the algorithm computes the degree of distrust of each seller. Our original contribution includes the definition of the seller rank, the method of using missing feedbacks as indicators of negative performance of sellers, the formulation of the algorithm for the seller rank computation, and the experimental evaluation of the proposal.

This paper is organized as follows. In Section 2 we present the related work on the subject. Section 3 introduces basic definitions used throughout the paper and presents the transformation of the original seller-buyer network. We use the resulting S-graph as the input data structure for the random walk method presented in Section 4. We report on the results of the experimental evaluation of our algorithm in Section 5. We conclude the paper in Section 6 with a brief summary of the future work agenda.

2. Related work

There are several ways to fight online auction fraud. Physical identification of existence of participants and items can increase the trust assigned to an identity or an auction. Buyers can be protected by fraud protection programs and insurances. Online auction sites encourage users to use secure payment mechanisms, such as credit cards or specialized secure payment companies, where the payment can be tracked and disputed. Another option is to use a trusted third party as an escrow service, which is preferred for high value transactions. Most online auction sites provide complaint centers and online dispute resolution services. Finally, sellers and buyers can use reputation systems based on feedback ratings and trust mark seals issued by a trusted party.

Most auction sites use the reputation system developed by eBay, where the credibility is expressed as the number of positive feedbacks minus the number of negative feedbacks received by the user (see Houser, 2001, and Resnick, 2002). A user can also receive neutral feedbacks. This simple mechanism suffers from several deficiencies, as pointed out in Malaga (2001). Feedbacks issued by users are subjective, lack transactional and social context, contain highly asymmetric information. Neutral feedbacks are very rare, the spectrum for positive feedbacks is very broad, and negative feedbacks occur only when the quality of service becomes unacceptable, otherwise users refrain from posting a negative feedback in the fear of retaliation.

In recent years several new solutions have been proposed that aim at overcoming at least some of the deficiencies of feedback-based models. An interesting proposal was formulated in Aberer and Despotovic (2001) where the authors present a complaint-only trust model. Although originally developed for the peer-to-peer environment, this highly decentralized model can be successfully used in online auctions. Another model originating from the peer-to-peer environment is PeerTrust (Xiong and Liu, 2003). PeerTrust is a complex model consisting of many parameters, such as feedback in terms of satisfaction, number of transactions, credibility of feedback, transaction context, and community context. The solution presented in Chen and Jaswinder (2001) prunes false feedbacks and accepts only feedbacks that are consistent with the majority of feedbacks received by a given user. The need for a trusted third party is advocated in Ba, Whinston and Zhang (2003). The authors propose to introduce a trusted judge that could authorize, identify, and manage the reputation of auction participants. An efficient method for assessing the level of trust between any two individuals, based on a small amount of explicit trust and distrust statements per individual is presented in Guha et al. (2004). A thorough survey of reputation management systems can be found in Ruohamaa, Kutvonen and Koutrouli (2007).

The problem of evaluating the importance of Web pages by Web search engines can be regarded as similar to the problem of reputation assessment in online auctions. The method for reputation estimation presented in Morzy, Wojciechowski and Zakrzewicz (2005) is based on the HITS algorithm (Kleinberg, 1998). It divides the participants of the online auction into two disjoint sets of sellers and buyers, and uses a recursive definition of credibility to estimate the reputation of each participant. Our current research is strongly influenced by another algorithm for Web page scoring, called PageRank (Page, 1998). Similarly to PageRank, we perform a random walk over the network of seller connections and we compute the stationary distribution of the Markov chain resulting from the transition matrix. The main difference is that the graph of seller connections is undirected and all edges between nodes are weighted.

3. Basic definitions

The main drawback of all feedback-based reputation systems is the fact that the reputation of a user can be strongly influenced by the behavior of other users who can submit incorrect feedbacks or withhold feedback completely. This influence is most visible when attackers form a coalition in order to maliciously modify the reputation of a selected user by submitting false positive feedback using forged identities. This inherent vulnerability of feedback-based reputation systems stems from the fact that reputation is computed using direct polling. Direct polling can result in an overly optimistic evaluation of reputation (Cialdini, 2000), because it collects excessive positive feedback as compared to negative feedback. The reason for this is that many users refrain from posting negative feedback because they are too inexperienced or have no comparison on which to base their opinion. One should not also neglect a psychological tendency to avoid conflicts that arise when negative feedbacks are issued. Therefore, instead of using direct polling, we choose to compute the reputation of a given user solely based on the reputation of other, relevant users.

3.1. Motivation

To better explain our approach, we will use the following example: consider a hypothetical shop "*Joe's*". Customers who shop only at "*Joe's*" could be biased in favor of "*Joe's*". It is possible to evaluate the reputation of a shop in another way. Other shop owners could be asked about what their customers think of the shop "*Joe's*". This approach would evaluate the reputation of "*Joe's*" based on opinions of users who have experience with other shops, and therefore avoid the overestimation of reputation. By considering only opinions of customers who go to more than one shop, the danger of receiving biased opinions is diminished. Furthermore, the owner of the "*Joe's*" has no direct impact on the set of customers being questioned about her reputation.

3.2. S-Graph

Assume we are given a set of sellers $S = \{s_1, s_2, \ldots, s_m\}$. Two sellers s_i and s_j are linked if there are at least *min_buyers* who committed an auction with sellers s_i and s_j, and the closing price for each auction was at least *min_value*. The set of such buyers is denoted B_{ij} and the number of such buyers is called the *strength* of the link, denoted $l_{ij} = |B_{ij}|$. The *neighborhood* $N(s_i)$ of the seller s_i consists of the set of sellers $\{s_j\}$, such that the seller s_i is linked with s_j, given user-defined thresholds *min_buyers* and *min_value*. The cardinality of the neighborhood $N(s_i)$ is called the *density* of the neighborhood.

The rationale behind user-defined thresholds is the following: *min_ buyers* selects sellers with a significant number of sales, and *min_ value* prunes transactions with low value. The density can be interpreted as follows: a buyer who buys from sellers s_i and s_j acknowledges the quality of both sellers. Unexperienced buyers are unlikely to link many sellers, rather, sellers are linked by experienced buyers. In this way, we discard unreliable information from unexperienced buyers. The fact that two sellers are linked indicates that either they trade similar and popular goods (e.g., music or books), or that their offer is complementary (e.g., bicycles and bicycle add-ons). Obviously, a link between two sellers may be purely coincidental. Nevertheless, high density of a seller is a good indicator of seller's trustworthiness. High density of a seller implies both a large number of sales, and the fact that the seller attracts many experienced buyers (buyers, who buy from other reliable sellers). In this way, density takes into consideration both the number of auctions (as traditional feedback counters do), and the quality of buyers. Intuitively, if two sellers are linked, they mutually share a part of their reputation and the link between the sellers may be considered as an endorsement of each other's reputation (Morzy, 2005).

Let $G = (S, E)$ be a directed graph with the set of nodes S and the set of edges E. An edge exists between the nodes s_i and s_j if sellers s_i and s_j are linked, given user-defined thresholds of *min_ buyers* and *min_ value*. We refer to the graph G as the *sellers graph*, or S-graph for short. Let f_{ij} denote the feedback provided by the buyer b_i to the seller s_j (for the sake of simplicity of presentation we assume that each buyer commits at most one auction with a given seller). Each feedback can be "positive", "neutral", "negative", or "missing". A feedback is considered "missing" if it has been deliberately left out, e.g., as the result of a hardly satisfying transaction, which does not qualify as bad, but deserves no praise for the quality of service. For the purpose of future computation of the positive and negative reputation we assign the following numerical values to different feedbacks:

$$f_{ij}^{+} = \begin{cases} 0.2 & if \quad f_{ij} = ''neutral'' \\ 0.8 & if \quad f_{ij} = ''positive'' \\ 0 & otherwise \end{cases}$$

$$f_{ij}^{-} = \begin{cases} 0.1 & if \quad f_{ij} = ''neutral'' \\ 0.2 & if \quad f_{ij} = ''missing'' \\ 0.7 & if \quad f_{ij} = ''negative'' \\ 0 & otherwise \end{cases}$$

Based on these values we define weights of edges in the S-graph. Each edge e_{ij} between nodes s_i and s_j has two weights representing the sums of "positiveness" and "negativeness" between the sellers s_i and s_j. The positive weight is given by $w_{ij}^{+} = \sum_{b_n \in B_{ij}} f_{nj}^{+}$. The positive weight captures praise expressed toward the seller s_j by the buyers acknowledged by the seller s_i. Analogously, the negative weight is given as $w_{ij}^{-} = \sum_{b_n \in B_{ij}} f_{nj}^{-}$ and it represents the amount of complaints expressed toward the seller s_j by the buyers acknowledged by the seller s_i. The positive weight is defined in terms of explicit feedbacks only. The negative weight is defined in terms of explicit feedbacks (”neutral”,”negative”) as well as implicit feedbacks (”missing”).

An interesting question arises of how to differentiate between an auction that has not been commented on purpose (this is the type of missing comment that we refer to as an implicit semi-negative comment) and an auction for which the feedback is not present due to other reasons (e.g., a buyer is an unexperienced user who does not know how to post a feedback). Below we propose a simple method for identifying missing implicit comments. Let $F(s_i) = \langle f_1, f_2, \ldots, f_n\rangle, f_i \in \{0, 1\}$ be a chronologically ordered list of feedback flags issued by the user s_i, where $f_k = 0$ denotes the fact that the k-th auction of the user s_i has not been commented and $f_k = 1$ denotes the fact that the k-th auction of the user s_i has an explicit feedback. We arbitrarily assume that the effect of experience from each auction (either positive or negative) influences next two auctions of a given user. $F(s_i)$ can be transformed into an ordered list of trigrams $T(s_i) = \langle t_1, t_2, \ldots, t_{n-2}\rangle$, where $t_i = f_i f_{i+1} f_{i+2}$ is a binary concatenation of feedback flags for the i-th auction with feedback flags for the consecutive two auctions. There are $2^3 = 8$ possible trigrams represented by binary numbers ranging from 000 (three consecutive auctions do not have a feedback) to 111 (three consecutive auctions have a feedback). Thus, $T(s_i)$ can be represented as a vector $\bar{T}(s_i) = \left[t_i^0, \ldots, t_i^7\right]$, where t_i^n is the number of occurrences of the n-th trigram in $T(s_i)$. We regard $\bar{T}(s_i)$ as a condensed representation of feedback habits of the user s_i. Having transformed the original history of user feedbacks into an 8-dimensional vector we can compare this vector to a template vector representing a user who almost never provides a comment for her auctions (in our experiments we have used the template vector $\bar{T}(0) = [1, 0.1, 0.1, 0.01, 0.1, 0.01, 0.01, 0]$, where three consecutive auctions without a comment have the weight 1, two missing comments have the weight 0.1, and one missing comment has the weight 0.01). Let the k-th auction of the seller s_i have no feedback. First, we build $F(s_i) = \langle f_1, f_2, \ldots, f_k\rangle$, which is transformed into $T(s_i) = \langle t_1, t_2, \ldots, t_{k-2}\rangle$, and the resulting list $T(s_i)$ is transformed into the vector $\bar{T}(s_i)$. Next, we compute the Ochini coefficient (cosine similarity measure) between $\bar{T}(s_i)$ and $\bar{T}(0)$ as follows

$$Ochini(\bar{T}(u_i), \bar{T}(0)) = \frac{\sum_{k=0}^{7} t_i^k * t_0^k}{\sqrt{\sum_{k=0}^{7} (t_i^k)^2 * \sum_{k=0}^{7} (t_0^k)^2}} .$$

If $Ochini(\bar{T}(u_i), \bar{T}(0)) < \beta$, where β is a user-defined threshold, we conclude that the two vectors are similar and the omission of a feedback should not be regarded as an implicit feedback. This procedure is flexible and allows for inclusion of temporary changes in posting habits. The choice of three consecutive auctions as the range of psychological influence of an auction outcome is arbitrary and can be freely changed.

To prove the existence of the implicit feedback we begin by investigating the distribution of the numbers of missing feedbacks per user (in this experiment we include only buyers). The results of the experiment are shown in Fig. 1. Interestingly, there are only few buyers with more than 20 missing feedbacks. This might indicate that most of the missing feedbacks are in fact purposeful omissions, thus turning the missing feedbacks into implicit feedbacks. When we have constrained our search to buyers who had participated in at least 10 auctions, the average percentage of missing feedbacks dropped to 11.6%, which indicates that experienced users are even less likely to omit a feedback.

Fig. 2 presents user selectivity depending on the value of the Ochini coefficient. Recall that the Ochini coefficient represents similarity between a given user's feedback vector and the template vector of a hypothetical 'I-don't-do-feedbacks' user, with the values closer to 1 representing high similarity and the values closer to 0 representing high dissimilarity. The figure presents the percentage of users who would be considered as generally not providing feedbacks, given the value of the Ochini coefficient threshold. For reasonable values of the Ochini coefficient threshold (i.e., 0.5 and above) less than 10% of buyers are regarded as reluctant to provide feedbacks, which means that their missing feedback would not be considered as implicit feedback. Again, this result proves that for the majority of buyers a missing feedback is an important, yet unspoken, assessment of business partner's performance.

4. Random walk

This section introduces two measures of seller reputation: the positive seller rank and the negative seller rank. These two measures could be combined, but we believe that they are independent. The two measures express trust and distrust toward a seller. Theoretically, trust and distrust can be independent; for example, a large volume of sales could lead to a high trust in the seller, but a low quality of service could increase the distrust simultaneously. In the S-graph, positive and negative seller ranks correspond to the "positive charge" and "negative charge" of each node. The intuition behind using the random walk is the following. Each edge e_{ij} in the S-graph represents the evaluation of the service provided by the seller s_j by buyers endorsed by the seller s_i (and *vice versa*). The positive weight w_{ij}^+ represents the trust that the seller s_i expresses (indirectly) with respect to the user s_j, and the negative weight w_{ij}^- represents the distrust that the seller s_i expresses with respect to the user s_j. Using the random walk method we can combine these binary relationships of trust and distrust into a global view, where each seller is assigned two measures of reputation.

Imagine a surfer who uses the online auction website to find the most reliable sellers by simply following links between auctions. Let us assume that the original network of buyer-seller connections has been already transformed into an S-graph. The surfer randomly picks a seller s_i as the starting point. Next,

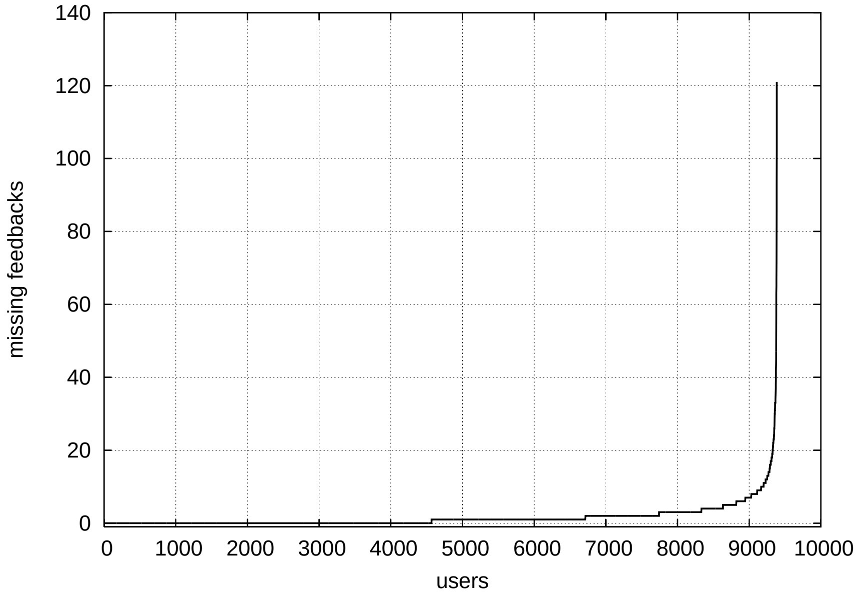

Figure 1. Missing feedback distribution.

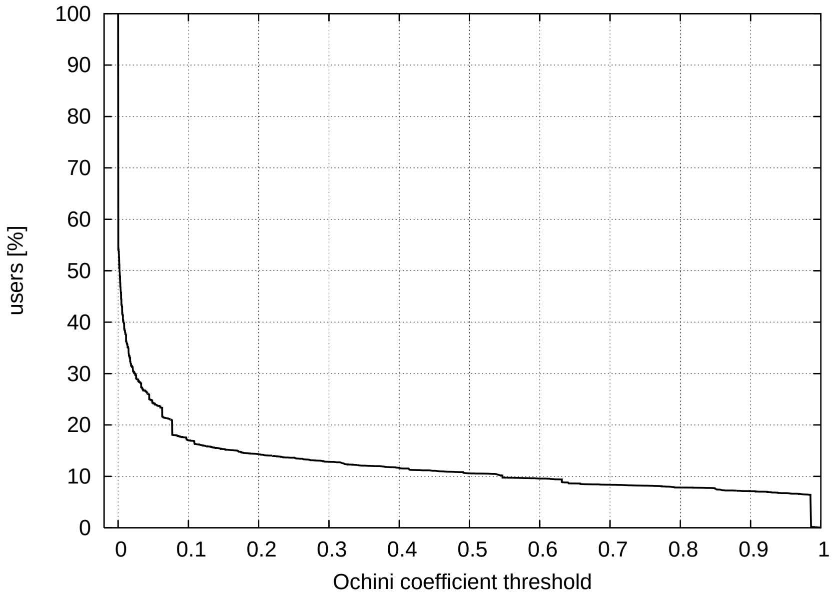

Figure 2. Selectivity of the Ochini coefficient.

in each iteration, the surfer chooses an edge e_{ij} to follow, with probability proportional to the positive weight of the edge, w^+_{ij}. Due to the existence of small clusters and cliques, irregularly, the surfer jumps randomly to another seller s_k and starts the walk again. Let $n = |S|$ be the number of nodes in the S-graph. Let β denote the probability of continuing the walk (as opposed to randomly jumping to a node). When averaging over a sufficient number of iterations, the probability that at any point in time the surfer is visiting node s_i is given by the formula:

$$P(s_i) = \frac{(1-\beta)}{n} + \beta \sum_{s_j \in N(s_i)} \frac{P(s_j) * w^+_{ij}}{\sum\limits_{s_k \in N(s_j)} w^+_{jk}} . \tag{1}$$

Contrary to the original PageRank formulation (Page et al., 1998), all edges in the S-graph are bidirectional. As the result, there are no dead-ends in the graph and no modifications of the graph structure are necessary[1] The *base rank* of a seller s_i is defined as the probability that a random surfer finishes her random walk in the i-th node.

Since Equation 1 is recursive, it must be computed iteratively until convergence. The base rank is the eigenvector of the transition matrix M^+ defined as

$$M^+ = (1-\beta) * \left[\frac{1}{n}\right]_{n \times n} + \beta A^+ \tag{2}$$

where

$$A^+_{ji} = \begin{cases} \dfrac{w^+_{ij}}{\sum\limits_{s_k \in N(s_j)} w^+_{jk}} & \textit{if } e_{ij} \in E \\ 0 & \textit{otherwise} \end{cases} . \tag{3}$$

Let x^t denote the vector of base ranks of sellers at time t, and let x^t_i denote the base rank of the seller s_i during iteration t. Initially, $x^0_i = 1$, i.e., all base ranks are initialized to 1. One iteration of Equation 1 corresponds to computing $x^{t+1} = M^+ x^t = (M^+)^t x^0$, and after a sufficient number of iterations x^T converges to the eigenvector of the transition matrix M^+, so we have $x^{T+1} = x^T$. The sequence $X = x^0, x^1, \ldots, x^t$ is a Markov chain. According to Panconesi (2005), the Markov chain converges to a unique stationary limit for any starting distribution if it is irreducible and aperiodic. Since the matrix M^+ allows to randomly jump to any node at any iteration, the Markov chain X is certainly irreducible, i.e., the underlying S-graph is strongly connected. Let p^n_{ii} denote the probability that the random surfer starting from the i-th node reaches this node again after n iterations. The Markov chain is periodic, if there exists a

[1] In the PageRank, for pages which have no outward-links, a link to all other pages in the Web is added to uniformly redistribute the rank sinking in dead-ends.

node i such that the greatest common denominator $gcd\{n \geq 1 : p_{ii}^n > 0\} > 1$. The Markov chain X of seller ranks is aperiodic, because the transition matrix M^+ does not contain any cycles.

After convergence, the values of base ranks are very low. In order to be able to meaningfully compare the base ranks assigned to sellers, we define the positive seller rank measure of the seller s_i as

$$SR^+(s_i) = \left\lceil \log_2\left(\frac{P(s_i)}{\min\limits_{s_j \in S}\{P(s_j)\}}\right) \right\rceil . \tag{4}$$

The positive seller rank measure projects the ratings assigned to sellers onto a small discrete numerical domain (in our experiments $Dom(SR^+) = \{1, 2, \ldots, 12\}$). Instead of taking the logarithm one can use linear scaling, as the choice of the scaling function is arbitrary. The scaling should be performed for practical reasons, to give the users of an online auction site an understandable and easy to use estimation of seller's positive reputation.

The computation of the negative reputation of each seller is identical in all steps except that instead of using the positive weight w_{ij}^+ of each edge, the negative weight w_{ij}^- is used. The probability that, when averaging over a sufficient number of iterations, the random surfer guided by negative weights of edges in the S-graph is visiting the node s_i is given by the formula:

$$P'(s_i) = \frac{(1-\beta)}{n} + \beta \sum_{s_j \in N(s_i)} \frac{P'(s_j) * w_{ij}^-}{\sum\limits_{s_k \in N(s_j)} w_{jk}^-} . \tag{5}$$

Again, in order to solve this recursive equation, we compute the eigenvector of the transition matrix M^- defined as

$$M^- = (1-\beta) * \left[\frac{1}{n}\right]_{n \times n} + \beta A^- \tag{6}$$

where

$$A_{ji}^- = \begin{cases} \dfrac{w_{ij}^-}{\sum\limits_{s_k \in N(s_j)} w_{jk}^-} & \textit{if } e_{ij} \in E \\ 0 & \textit{otherwise} \end{cases} . \tag{7}$$

The properties of the transition matrix M^- are identical to the properties of the matrix M^+ defined in Equation 3, so we deduce that, analogously to M^+, the Markov chain defined by the transition matrix M^- has a unique stationary distribution for any starting distribution. The base rank values in the eigenvector of the transition matrix M^- are very low, so we scale these values to obtain final negative seller rank defined as

$$SR^{-}(s_i) = \left\lceil \log_2\left(\frac{P'(s_i)}{\min\limits_{s_j \in S}\{P'(s_j)\}}\right) \right\rceil . \tag{8}$$

In the next section we present the results of experimental evaluation of the proposed solution.

5. Experiments

The data have been acquired from `www.allegro.pl`, Polish leader of online auctions. The dataset consists of 440,000 participants, 400,000 auctions, and 1,400,000 bids. The number of participants is greater than the number of auctions, because for each participant the highest bid is stored in the database, whether it was the winning bid or not. Therefore, we have data on some participants who did not win any auction. The dataset has been created using the following procedure: 10,000 sellers have been randomly picked, and for this seed set all their auctions from the period of six months have been collected. Next, all buyers participating in these auctions have been added to the dataset. Analogously, 10,000 buyers have been randomly picked and a similar procedure has been applied to this seed set. Altogether, complete information on 20,000 participants is available. Data are stored and preprocessed using Oracle 10g database.

The thresholds for building the S-graph were set as follows: $min\_buyers = 1$ and min_price = \$15. The minimum price threshold prunes 5% of sellers who do not offer any items above the threshold price. For the above mentioned thresholds 57% of sellers are represented in the S-graph. One might argue that the *min_price* threshold is set too prohibitively, but the average price of items in the mined dataset is close to \$30, so the threshold rather prunes insignificant auctions. Fig. 3 presents the distribution of the positive seller rank. Interestingly, the random walk procedure does not constrain the highest positive seller ranks to the sellers with the highest density. In our dataset the highest density described 14% of the sellers, whereas the highest seller ranks of 11 and 12 cover almost 25% of the sellers. The highest value of 12 covers slightly more than 1% of sellers. This result is particularily encouraging because it suggests that the positive seller rank is not only merely a function of the volume of sales. The convergence of the positive base rank is presented in Fig. 4. In this experiment the positive base ranks of sellers have all been initialized to $x_i^0 = 1$. As can be noticed, positive base ranks converge quickly and after 40 iterations the stationary distribution of the Markov chain $X = x^0, x^1, \ldots, x^t$ is found.

Two clusters of sellers are visible, the majority of sellers have density in the range $\langle 1, 100\rangle$, 14% of sellers have density in the range $\langle 1400, 1500\rangle$ and the average density is 38.6. We interpret this result in the following way: the smaller cluster of dense sellers represents the group of most important and

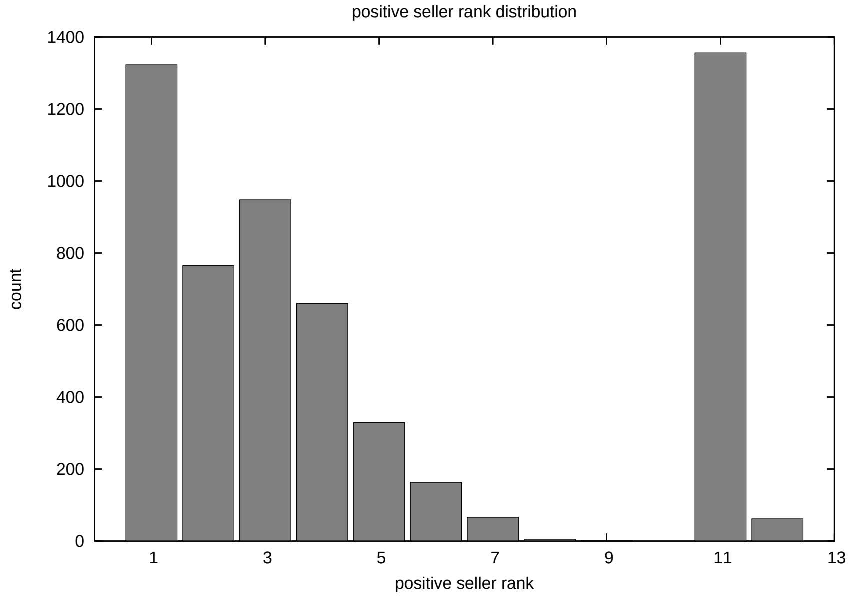

Figure 3. Positive seller rank distribution.

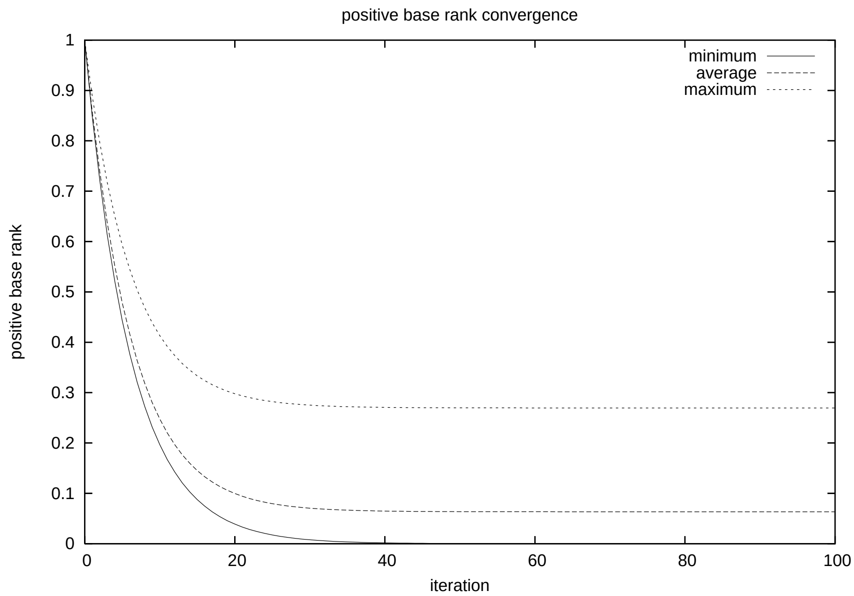

Figure 4. Positive base rank convergence.

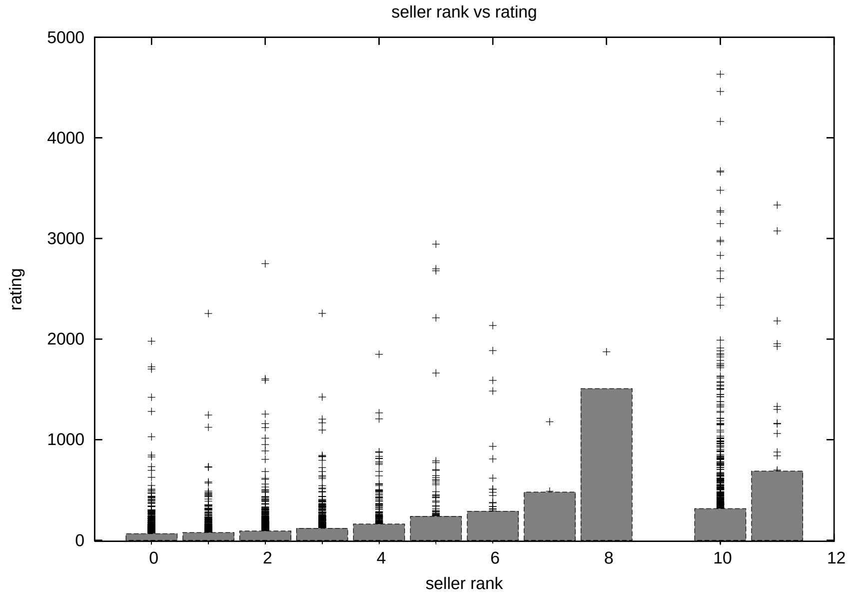

Figure 5. Positive seller rank vs rating.

credible sellers, mostly big online retailers. The bigger cluster represents both casual sellers (with density in the range $\langle 1, 10\rangle$), and small online shops (with density in the range $\langle 10, 100\rangle$).

Because the Markov chain X is both irreducible and aperiodic, the static distribution does not depend on the initial distribution x^0. This result can be easily noticed in Fig. 4, where convergence of the base rank is depicted for the base ranks of sellers initialized to sellers ratings (traditional feedback counts). Please note that the maximum base rank line is plotted using second y-axis. The minimum base rank line starts below 0 because the dataset contains examples of sellers dominated by negative feedbacks.

In order to compare the seller rank to the currently used feedback counts we performed the next two experiments. Fig. 5 shows the projection of the positive seller rank onto seller rating (traditional feedback count). Solid boxes represent average ratings for each value of the positive seller rank. Unusually high rating values for $SR^+ = 8$ and $SR^+ = 9$ are outliers caused by a small sample size (less than 1% of sellers have positive seller ranks of 8 or 9). It follows from the figure that the average rating increases with the increase of the positive seller rank. Importantly, the positive seller rank measure is more discriminatory that the traditional rating. Fig. 5 shows several examples of sellers with similar ratings scattered across different positive seller rank buckets. An interesting feature of the seller rank measure is the fact, that the highest values are also being assigned to sellers with relatively low ratings.

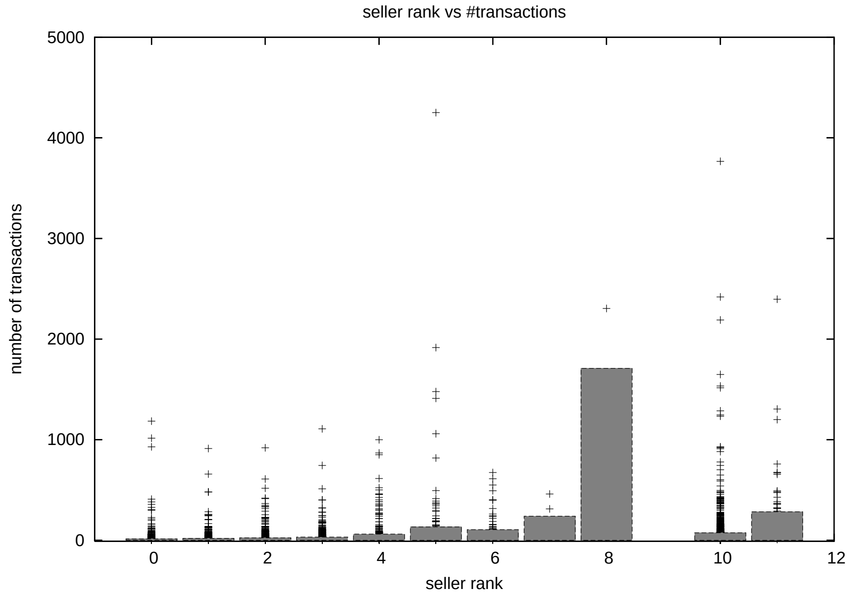

Figure 6. Positive seller rank vs number of auctions.

Fig. 6 presents the projection of the positive seller rank measure onto the average number of auctions of sellers. We perform this experiment in order to see if there is a correlation between the reliability of the seller (as measured by the positive seller rank measure) and the number of auctions. As in previous figure solid boxes represent the average number of auctions of sellers who have been assigned a given positive seller rank value. For smaller values of the positive seller rank measure the differences in the average number of auctions are negligible, but starting with the seller rank value of 4 one can easily notice a growing pattern, with the notable exception of $SR^+ = 6$, which covers 1.2% of all sellers. In general, however, a greater value of the positive seller rank indicates higher average number of auctions for a given seller. The results shown in Figs. 5 and 6 are not comparable, because seller ratings represent the ratings collected by sellers during the entire lifespan of each seller identity, and the number of auctions for each seller has been computed based on the available dataset covering the period of six months.

The analysis of the properties of the negative reputation measure is presented in Figs. 7 and 8. These experiments were conducted on the S-graph created with the following values of thresholds: $min\_buyers = 20$ and $min\_price = 0$. Fig. 7 depicts the distribution of the negative seller rank. As expected, the majority of sellers are characterized with the lowest negative reputation. However, starting with the negative seller rank value of 4 the measure captures those sellers, whose quality of service is questionable. We are able to discover this fact mostly by

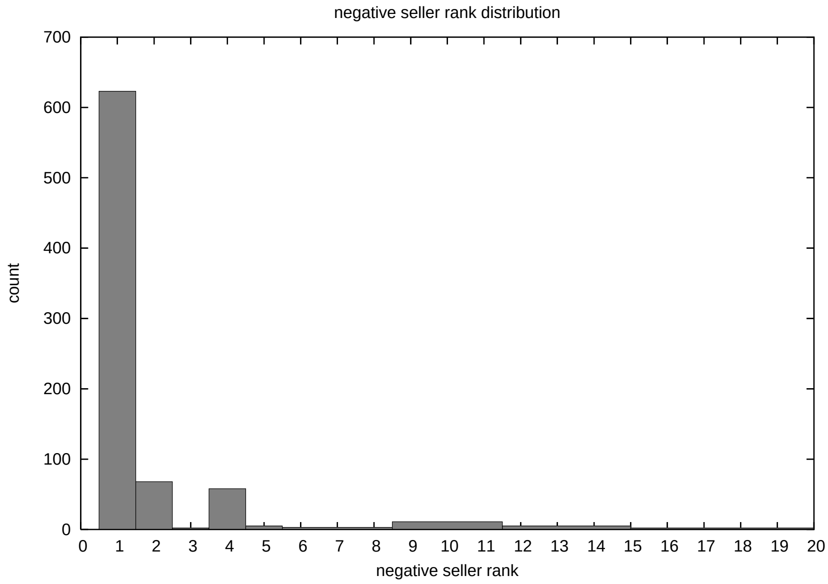

Figure 7. Negative seller rank distribution.

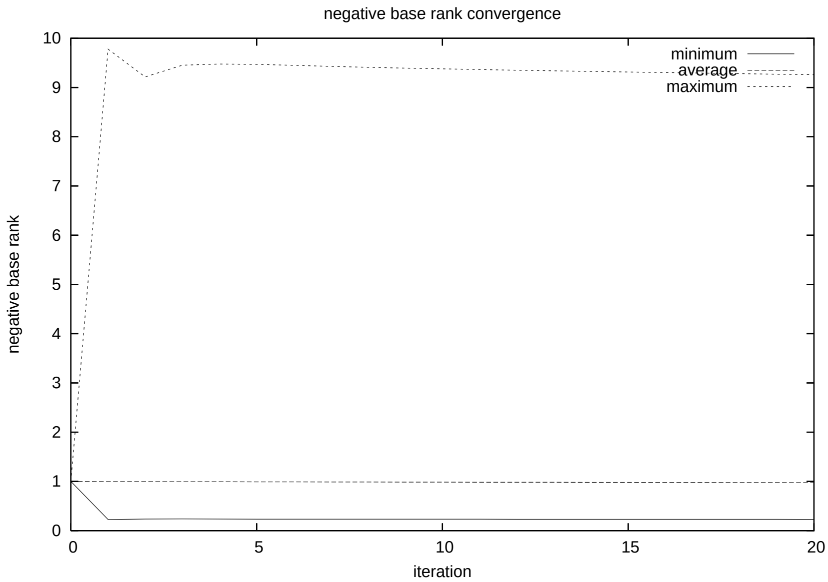

Figure 8. Negative seller rank convergence.

using implicit disapproval hidden in missing comments. Fig. 8 presents the convergence of the negative base rank (minimum, average, and maximum values) during first 20 iterations. As can be seen, the eigenvector of the transition matrix containing the negative base rank stabilizes very quickly, after only few iterations.

6. Conclusions

In this paper we have introduced a novel reputation measure for sellers in online auctions. Our measure considers the network of seller-buyer connections and mines the topology of the network to derive useful knowledge about sellers. All computations are performed on the transformed S-graph, where information from explicit and implicit feedbacks are aggregated. Application of random walk procedures to the S-graph reveals further interesting patterns and properties of the data. In particular, the stationary distributions of the vectors of base seller ranks, defined on the transition matrices can be used as the measure of seller's positive and negative reputation. We believe that the positive seller rank can be successfully used as an indicator of seller's reliability, and the negative seller rank can be used as an indicator of seller's questionability. In other words, the positive seller rank measures the trust, while the negative seller rank measures the distrust. Main advantages of the proposed solution include resistance to manipulation, ability to discover complex fraudulent activities, and practical usability.

References

Aberer, K. and Despotovic, Z. (2001) Managing trust in a peer-2-peer information system. *CIKM'01: Proc. of the 10th International Conference on Information and Knowledge Management*, New York, NY, USA. ACM Press, 310–317.

Ba, S., Whinston, A.B. and Zhang, H. (2003) Building trust in online auction markets through an economic incentive mechanism. *Decision Support Systems* **35** (3), 273–286.

Chen, M. and Jaswinder, P.S. (2001) Computing and using reputations for internet ratings. *EC'01: Proc. of the 3rd ACM Conference on Electronic Commerce*, New York, NY, USA. ACM Press, 154–162.

Cialdini, R.B. (2000) *Influence: Science and Practice.* Allyn & Bacon.

Douceur, J. (2002) The Sybil Attack. *IPTPS'02: Proc. of the 1st International Workshop on Peer-to-Peer Systems*, Cambridge, MA, USA. **LNCS 2429**, Springer Verlag, 251–260.

Forrester (2007) Forrester Research, State of retailing online 2007. `http://www.shop.org/soro07`

Guha, R., Kumar, R., Raghavan, P. and Tomkins, A. (2004) Propagation of trust and distrust. *WWW'04: Proc. of the 13th Intern. Conf.*

on World Wide Web, New York, NY, USA. ACM Press, 403–412.
Houser, D. and Wooders, J. (2001) Reputation in auctions: Theory, and evidence from eBay. Technical Report, University of Arizona.
IC3 (2007) 2006 IC3 Annual Report. FBI Internet Crime Complaint Center `http://www.ic3.gov/media/annualreport/2006_IC3Report.pdf`
Kleinberg, J. (1998) Authoritative sources in a hyperlinked environment. *SODA'98: Proc. of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms*, Philadelphia, PA, USA. Society for Industrial and Applied Mathematics, 668–677.
Malaga, R.A. (2001) Web-based reputation management systems: Problems and suggested solutions. *Electronic Commerce Research* **4** (1).
Marsh, S. and Dibben, M. (2005) Trust, untrust, distrust and mistrust – an exploration of the dark(er) side. *iTrust'05: Proc. of the 3rd International Conference iTrust, Rocquencourt, France.* **LNCS 3477**, Springer Verlag, 17–33.
Melnik, M.I. and Alm, J. (2002) Does a seller's e-commerce reputation matter? Evidence from eBay auctions. *The Journal of Industrial Economics* **50** (3), 337–347.
Morzy, M. (2005) Density-based measure of reputation of sellers in online auctions. *ADMKD'05: Proc. of the 1st ADBIS Workshop on Data Mining and Knowledge Discovery*, Tallinn, Estonia. Publishing House of Poznan University of Technology, 65–75.
Morzy, M., Wojciechowski, M. and Zakrzewicz, M. (2005) Intelligent reputation assessment for participants of web-based customer-to-customer auctions. *AWIC'05: Proc. of the 3rd International Atlantic Web Intelligence Conference*, Lodz, Poland. **LNAI 3528**, Springer Verlag, 320–326.
Mui, L. (2003) Computational models of trust and reputation: Agents, evolutionary games, and social networks. Ph.D. Dissertation, Massachusetts Institute of Technology.
Page, L., Brin, S., Motwani, R. and Winograd, T. (1998) The PageRank Citation Ranking: Bringing Order to the Web. Technical Report, Stanford University. http://ilpubs.stanford.edu:8090/422/
Panconesi, A. (2005) The stationary distribution of a Markov chain. Unpublished note, Sapienza University of Rome, http://www.dis.uniroma1.it/~leon/didattica/webir/pagerank.pdf.
Resnick, P., Zeckhauser, R., Friedman, E. and Kuwabara, K. (2000) Reputation systems. *Communications of the ACM* **43** (12), ACM Press.
Resnick, P. and Zeckhauser, R. (2002) Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system. In: M.R. Baye, ed., *The Economics of the Internet and E-Commerce. Advances in Applied Microeconomics*, **11**, Elsevier Science.
Resnick, P., Zeckhauser, R., Swanson, J. and Lockwood, K. (2004) The value of reputation on eBay: a controlled experiment. Technical Report, School of Information, University of Michigan.

Ruohamaa, S., Kutvonen, L. and Koutrouli, E. (2007) Reputation management survey. *ARES'07: Proc. of the 2nd International Conference on Availability, Reliability and Security*, Washington, DC, USA. IEEE Computer Society, 103–111.

Xiong, L. and Liu, L. (2003) A reputation-based trust model for peer-to-peer e-commerce communities. *EC'03: Proc. of the 4th ACM conference on Electronic Commerce*, New York, NY, USA. ACM Press, 228–229.