Intelligent Reputation Assessment for Participants of Web-based Customer-to-Customer Auctions

Mikołaj Morzy, Marek Wojciechowski, and Maciej Zakrzewicz

Institute of Computing Science
Poznań University of Technology, Piotrowo 3A, 60-965 Poznań, Poland
{Mikolaj.Morzy,Marek.Wojciechowski,Maciej.Zakrzewicz}@cs.put.poznan.pl

Abstract. The Internet witnesses the unprecedent boom of customer-to-customer e-commerce. Most online auction providers use simple participation counts for reputation rating, thus enabling dishonest participants to cheat. In this paper we propose a novel definition of reputation and credibility of C2C e-commerce participants and we present an algorithm for reputation rating estimation. We conduct several experiments on real-world data which prove the feasibility of our algorithm.

1 Introduction

The Internet is quickly becoming an important arena of a novel type of merchandise called electronic commerce, or e-commerce for short. One of the most important models of e-commerce is customer-to-customer commerce representing auctions. We investigate the fundamental property of the C2C model, namely, the credibility of participants. Indeed, trust, fairness, and credibility are perceived by the users as crucial issues in online trading through C2C channels. The anonymity provided by the Internet tempts the participants into dishonest behavior. Unfortunately, currently used reputation reporting mechanisms are not satisfactory and can be easily deceived by malicious participants. Most popular auction sites use a simple participation counter for reputation reporting. Other users are allowed to see this counter along with textual comments and ratings (usually labeled with "negative", "neutral", and "positive"). In order to avoid unfairly high or low ratings only users who truly finalized an auction can mutually post comments and ratings.

In this paper we introduce a novel approach to reputation estimation. We propose to use a data mining technique to analyze the graph of connections between participants to derive knowledge about the credibility of each participant. Our method efficiently discovers most common types of frauds that occur in online auctions. Our accomplishments are twofold. First, we propose a novel definition of the reputation based on credibility of contractors and we present an efficient algorithm to compute it. Second, we empirically prove the practical usability of our algorithm by mining a large volume of real-world data obtained from a leading Polish online auction provider. In addition, we perform a controlled fraud and we show how our method quickly discovers malevolent behavior.

This paper is organized as follows. In Section 2 we present the related work on the subject. Section 3 contains definitions of basic notions used throughout the paper. We describe our algorithm in details in Section 4 and we present the results of the empirical evaluation of the algorithm in Section 5. The paper concludes with future work agenda in Section 6.

2 Related Work

Reputation systems [1] are a practical way of building trust in environments with high anonymity and low trustworthiness of participants. Contemporary Web-based auction systems rely on simple trust models with credibility of participants assessed by counting comments received after each transaction. In [3] Malaga presented a critical analysis of such simple models, identifying several problems that should be solved, including the subjective nature of feedbacks, the credibility of feedbacks, the lack of feedback's context, the lack of differentiation of recent and older feedbacks, and the lack of incentives for a participant to rate the trading partner. Several solutions have been proposed to address at least some of the limitations of current feedback-based models. In [2], the authors introduced a complaint-only trust management method, based on the fact that users are more eager to provide negative comments if they are not satisfied, than to give positive feedback. Another method presented in [4] differentiates comments by taking into account the credibility of the rater, assuming that the rating is of good quality if it is consistent with the majority of ratings. In [5] a novel trust model called PeerTrust was proposed. The presented model includes several trust parameters, such as feedback in terms of satisfaction, number of transactions, credibility of feedback, transaction context, and community context. Credibility of feedback in PeerTrust is assessed differently than in [4]. The idea is to give more weight to feedbacks from more credible participants.

Somewhat related to the problem of reputation assessment in e-commerce systems is the problem of evaluating importance of Web pages. Examples of algorithms for judging the importance of pages are PageRank [7] and HITS [6]. Our method for credibility assessment is somehow similar to the later algorithm. HITS divides the pages into authorities (covering a certain topic) and hubs (directory-like pages linking to authorities). In our method, we apply a similar distinction, dividing auction participants into sellers and buyers, and we use adjacency matrices to recursively compute the credibility of participants.

3 Basic Constructs

Given a set of buyers $B = \{b_1, b_2, \dots, b_n\}$ and a set of sellers $S = \{s_1, s_2, \dots, s_m\}$. Let c denote a comment, $c \in \{-1, 0, 1\}$, where each value represents the "negative", "neutral", and "positive" comment, respectively. Given a set of auctions $A = \{a_1, a_2, \dots, a_p\}$. An auction is a tuple $a_i = \langle b_j, s_k, c\rangle$ where $b_j \in B \wedge s_k \in S$.

Let $S(b_j)$ represent the set of sellers who sold an item to the buyer b_j. We denote the *support* of the buyer b_j as $support(b_j) = |S(b_j)|$. Let $B(s_k)$ represent the set of buyers who bought an item from the seller s_k. We denote the *support* of the seller s_k as $support(s_k) = |B(s_k)|$. According to this formulation, the support of the participant (either buyer or seller) is identical to the reputation measure currently employed by leading online auction providers.

Given a $m \times n$ matrix M_S. Each entry in the matrix represents the flow of support from the seller to the buyer in a finalized auction. Entries in the matrix M_S are initialized as follows.

$$\forall\, i \in \langle 1, m\rangle \;\; M_S[i,j] = \frac{1}{support\,(b_j)} \;\; if\; \langle b_j, s_i, c\rangle \in A,\; 0 \; otherwise$$

Given a $m \times n$ matrix M_B. Each entry in the matrix represents the flow of support from the buyer to the seller in a finalized auction. Entries in the matrix M_B are initialized as follows.

$$\forall\, j \in \langle 1, n\rangle \;\; M_B[i,j] = \frac{1}{support\,(s_i)} \;\; if\; \langle b_j, s_i, c\rangle \in A,\; 0 \; otherwise$$

Given a vector $S_C = [s_1, s_2, \ldots, s_m]$ of seller credibility ratings. Initially, all sellers receive the same credibility of 1. Analogously, given a vector of buyer credibility ratings $B_C = [b_1, b_2, \ldots, b_n]$. Initially, all buyers receive the same credibility of 1. Upon the termination of the algorithm vectors S_C and B_C contain diversified credibility ratings for sellers and buyers, respectively. A reputation rating for a buyer b_j is a tuple $R(b_j) = \langle C_-, C_0, C_+\rangle$. Each component represents the sum of credibilities of sellers participating in transactions with a given buyer and posting a negative, neutral, or positive comment, respectively. Formally, $C_- = \sum_k S_C[k]$ where $\langle b_j, s_k, -1\rangle \in A$, $C_0 = \sum_k S_C[k]$ where $\langle b_j, s_k, 0\rangle \in A$, and $C_+ = \sum_k S_C[k]$ where $\langle b_j, s_k, +1\rangle \in A$. Reputation rating for a seller can be defined analogously.

4 Iterative Reputation Assessment Algorithm

Our method of reputation rating is based on the following recursive definition of credibility. A given buyer is highly credible if the buyer participates in many auctions involving credible sellers. Analogously, a given seller is credible if the seller participates in many auctions involving credible buyers. Since there is no *a priori* estimation of credibility of participants, we assume that initially all participants receive equal credibility. Then, we iteratively recompute the credibility of sellers and buyers in the following way. In each iteration we distribute the current credibility of each buyer among participating sellers. Next, we update the credibility of all sellers by aggregating the credibility collected from participating buyers. After this update we propagate the current credibility of sellers to buyers and we refresh the appropriate ratings. We repeat this procedure several times until the credibility of sellers and buyers converge. Alternatively, the procedure can be repeated a given number of times. After assessing the credibility of all participants the credibility ratings are used together with past comments to derive proper reputation ratings by aggregating the credibility of contractors grouped by the type of the comment issued after the transaction. The intuition behind the algorithm is that the credibility of “good” buyers quickly aggregates in “good” sellers and *vice versa*. Initial ratings consisting of simple participation counts are quickly replaced by the true credibility which reflects the importance of every participant. The outline of the algorithm is presented in Fig. 1.

Require: $A = \{a_1, a_2, \ldots, a_p\}$, the set of finalized auctions
Require: $B = \{b_1, b_2, \ldots, b_n\}$, the set of buyers
Require: $S = \{s_1, s_2, \ldots, s_m\}$, the set of sellers
Require: M_S,M_B, matrices representing the structure of the inter-participant network
Require: S_C,B_C, vectors representing the credibility of participants

Initialize matrices M_S,M_B and vectors S_C,B_C appropriately
repeat
  for all $s_k \in S$ **do**
    $S_C[s_k] = \sum_{j=1}^{n} M_S[j,k] * B_C[b_j]$
  end for
  for all $b_j \in B$ **do**
    $B_C[b_j] = \sum_{k=1}^{m} M_B[j,k] * S_C[s_k]$
  end for
until vectors S_C and B_C converge
Output S_C and B_C as credibility ratings
Compute reputation ratings $R(b_j)$, $R(s_k)$ $\forall b_j \in B$, $\forall s_k \in S$ using S_C, B_C, and A

Fig. 1. Iterative Reputation Assessment Algorithm

5 Empirical Results

Here we present the results achieved on real datasets. The datasets have been acquired from `www.allegro.pl`, the leading Polish provider of online auctions. The datasets contain information on 400 000 participants and over 2 000 000 terminated auctions. All experiments are conducted on Pentium IV 2.4 GHz with 480 MB RAM. Data are stored and preprocessed using Oracle 9i database.

Figures 2 and 3 present the scaling of the algorithm. We differentiate the number of users from 1 000 to 100 000. The performance of the algorithm is satisfactory even for large user communities. We attribute the performance of the algorithm to the delegation of the most computationally expensive parts of the algorithm to the database engine and replacing the procedural processing with recursive SQL processing. The second test verifies the scalability of the algorithm with respect to the number of auctions. As can be seen, the execution time of the algorithm is almost linear.

Figures 4 and 5 present the convergence of credibility computed by our algorithm. Figure 4 depicts the changes of credibility in subsequent iterations for a selected subset of sellers. We choose the sellers with the highest standard deviation to include in the figure, so the figure presents only the most atypical sellers.

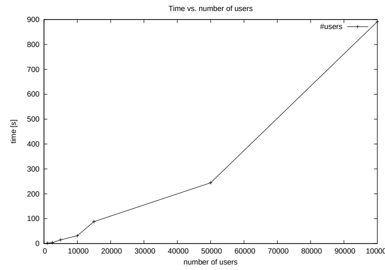

Fig. 2. Time vs. #users

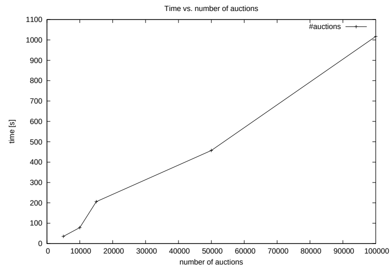

Fig. 3. Time vs. #auctions

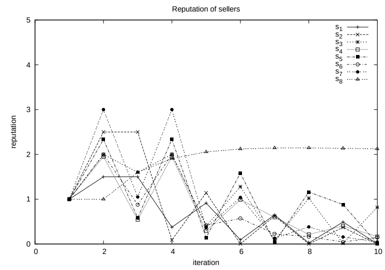

Fig. 4. Credibility of sellers

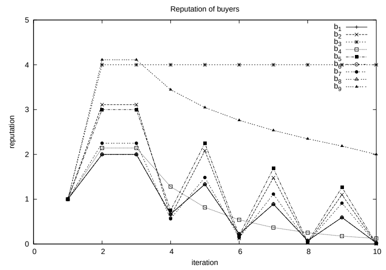

Fig. 5. Credibility of buyers

For the vast majority of sellers the changes in credibility are much smoother and the final credibility estimation stabilizes after a few iterations. The results of a similar selection for buyers are depicted in Fig. 5. Again, the estimation of credibility quickly converges and the rating stabilizes after only a few iterations.

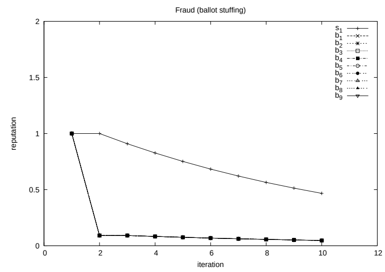

Fig. 6. Ballot stuffing

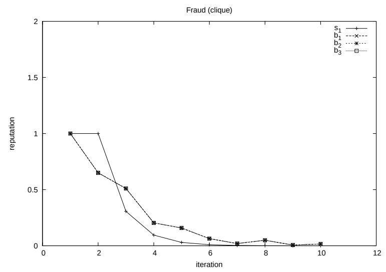

Fig. 7. Clique

In the next experiment we are simulating ballot stuffing. A dishonest seller s_1 decides to create dummy buyers $b_1, \ldots, b_{10}$ to inflate his/her reputation rating. Additionally, the seller s_1 participates in auctions with buyers from outside the group. Figure 6 presents the credibility estimation for the group of participants involved in cheating. Estimates for all involved buyers are exactly the same, because those buyers are indistinguishable from the point of view of the topology of relationships. Already in the second iteration the algorithm discovers that buyers $b_1, \ldots, b_{10}$ are not credible, since they do not receive any feedback from sellers other than s_1. The seller s_1 initially aggregates all credibility from the buyers $b_1, \ldots, b_{10}$, but the credibility of the seller diminishes over time, causing the credibility of the buyers $b_1, \ldots, b_{10}$ to drop even further. The credibility of the seller s_1 slowly evaporates through the relationship with buyers from outside the group and no new credibility flows in from other sellers or buyers.

The next experiment represents a plot to form a clique. A dishonest seller tries to outwit the system by creating a set of virtual buyers and interconnects them to form a clique (one can easily imagine registering few users and finalizing low cost auctions between them to make them pretend as credible and active participants of auctions). Figure 7 presents the credibility ratings for seller s_1 and a group of buyers b_1, b_2, and b_3 involved in a clique plot. The algorithm discovers the fraud and determines that the real credibility of participants is low. Therefore, after a few iterations the deceiving group receives a low credibility rating. This result is probably even more desirable than the previous one, because the clique cheating is more dangerous to honest auction participants and harder to discover using manual analysis methods.

6 Conclusions

In this paper we have presented a novel algorithm for reputation rating of on-line auction participants, which evaluates the reputation based on the network of inter-participant relationships using a recursive definition of credibility. The experiments prove the practical usability of the solution. Our future work agenda includes extension of the algorithm to safeguard against artificial lifting of bids by dummy buyers created by dishonest sellers. We also plan to scale the algorithm to allow for real-time analysis of huge amounts of data.

References

1. P. Resnick, R. Zeckhauser, E. Friedman, K. Kuwabara. *Reputation systems.* Communications, of the ACM, 43(12), 2000.
2. K. Aberer and Z. Despotovic. *Managing trust in a peer-to-peer information system.* In Proc. of the ACM CIKM, pp. 310-317, 2001.
3. R. A. Malaga. *Web-based reputation management systems: Problems and suggested solutions.* Electronic Commerce Research, 1(4), 2001.
4. M. Chen and J. P. Singh. *Computing and using reputations for Internet ratings.* In 3rd ACM Conference on Electronic Commerce, 2001.
5. L. Xiong, L. Liu. *A Reputation-Based Trust Model for Peer-to-Peer eCommerce Communities.* In IEEE Conf. on Electronic Commerce (CEC'03), Newport Beach, USA, June, 2003.
6. J.Kleinberg, *Authoritative Sources in a Hyperlinked Environment.* Proc. 9th Ann. ACM-SIAM Symp. Discrete Algorithms, ACM Press, New York, 1998.
7. L. Page, S. Brin, R. Motwani, T. Winograd, *The PageRank Citation Ranking: Bringing Order to The Web.* Stanford Digital Library Technologies Project, 1998.