A Bayesian framework for online reputation systems

Petteri Nurmi
Helsinki Institute for Information Technology HIIT
P.O. Box 68, University of Helsinki, FI-00014, Finland
petteri.nurmi@cs.helsinki.fi

***Abstract*— As the number of online auction sites has increased, interest towards providing reliable summaries, reputations, about the past behaviour of sellers has risen. Existing approaches are often based on heuristics and hence proper evaluation of the systems is difficult. To improve the situation we introduce a game theoretical model for reputation in online auctions. As our second contribution we discuss how reputation systems can use Bayesian model averaging to integrate different information sources. We also discuss practical aspects and evaluate an experimental online reputation system that is based on the proposed framework. Our results indicate that mechanisms based on our framework increase the number of successful transactions significantly (in statistical sense) and result in mechanisms that are robust for changes in the behaviour of the participants.**

I. Introduction

Many online portals such as eBay [1] and Epinions [2] provide their users the opportunity to arrange auctions in a large online community. Unfortunately this kind of systems easily attract selfish users who attempt to utilize the system for their personal benefit. To cope with the situation customers are given summaries, reputations, about the past behaviour of sellers. A requirement for the mechanism giving these summaries is that it discourages selfish behaviour (cheating) and rewards cooperation.

To give practical insights into the problem, we consider the eBay auction site. On eBay the participants are divided into (possibly overlapping) sets of customers (buyers) and sellers. The customers buy items from sellers using the following scheme: first an auction is held and after the price has been fixed the buyer transfers the agreed sum to the account of the seller who then delivers the items that were part of the agreed transaction. However, if the buyers have no additional information about the seller, according to rationality arguments the optimal strategy of the seller is to always cheat, i.e., to keep the money and not to send anything to the buyer or to deliver an item of lower quality. The reputation mechanism used on eBay attempts to overcome cheating by allowing the buyers to give positive (+1), neutral (0) or negative (-1) feedback about the seller after a transaction has been concluded and showing the cumulative sum of feedback to customers. In addition, the buyers can post short comments about the sellers, which are also shown to customers.

As the concept of reputation has various applications in diverse fields [3], also the problem of finding a proper theoretical model has already been addressed. Still, none of the existing models seems to sufficiently explain the behaviour that researchers have observed in, online auction sites such as eBay. We attempt to improve the situation by proposing a new game theoretic model for eBay-like online auctions. In the proposed model each seller maintains an estimate of the reliability of a seller and estimates of multiple sellers are integrated using *Bayesian model averaging* [4]. The model is further discussed in Sections III and IV and the integration of information sources is discussed in Section V.

Although the model we propose is theoretically optimal, it is somewhat insufficient in practice. To boost the performance of implementations that are based on the proposed framework, we introduce in Section VI mechanisms for providing additional information to the buyers.

To further support our arguments, we have implemented an online reputation system that is based on the proposed framework. We have also implemented a system that uses the eBay reputation mechanism and compared the performances of the two approaches. The results of the experiments are discussed in Section VII . Finally, Section VIII concludes the paper and discusses future work. Before going into details of our work, we discuss related work.

II. Related work

Initially the concept of reputation was discussed by Akerlof when he introduced the *market of lemons* problem [5]. In the problem a person is assumed to buy a car from a market. The car can be either a good car or a lemon (a car of bad quality). The buyer must pay for the car in advance, but (s)he will know the quality of the car only afterwards. According to game theoretic arguments, the rational seller should sell a lemon because this will maximize the seller's personal gain. Similarly, if this model is replayed a discrete number of times, the optimal strategy is selling a lemon at each instance of the game (stage). In reality the behaviour of markets of this kind is usually different and thus we need to adopt different optimality conditions for the players.

In modelling reputation the most important optimality condition is called the *sequential equilibrium* and it states that the actions of the buyers should be optimal given the previous game play and the beliefs the buyers have about the behaviour of the seller [6]. Moreover, sequential equilibrium assumes that the beliefs of the buyers are sequentially rational which means that they should eventually converge to represent the correct probability distribution of the types (sell a good car or a lemon). The first proper game theoretic models, introduced by Kreps and Wilson [7] and by Milgrom and Roberts [8], were based on sequential rationality in repeated Bayesian-games.

The generalization to extensive form games was introduced a few years later by Fudenberg and Levine [9].

In empirical terms, the properties of eBay have been widely studied and there has been both experimental evaluations as well as theoretic approaches that aim to identify weaknesses in eBay's reputation scheme or that attempt to model some peculiarities in the observed behaviour. Of special interest is the empirical work conducted on eBay by Resnick et al. [10], according to whom the marketplace works rather well even though the mechanism itself has many shortcomings such as lack of feedback and easy access to new pseudonyms.

In other online communities the work has mainly concentrated on proposing new mechanisms. For example, [11] uses a Beta-distribution for maintaining reputation values. Other approaches include Sporas [12] and Eigentrust [13].

III. The reputation game

In this section we define the structure of the reputation game the buyers and sellers are playing in an online auction site. Our model differs from previous work in formalizing the way beliefs are derived and updated over time. Before going into details about the model, we define some notation.

Let $\mathcal{M}$ be an online marketplace with a finite set of buyers N and a finite set of sellers M. An arbitrary buyer is indexed using the variable i whereas sellers are indexed with the variable j. The online nature of the market determines the transaction policy of the system which is assumed to be similar as in eBay. Thus after the seller and buyer have reached an agreement about the price (and items), the buyer transfers the agreed amount of money to the seller's bank account and then the seller delivers the agreed items to the buyer. The transferred money amount is called the cost of the transaction and the variable c_{ij} is used to denote it.

We assume that each seller j has a discrete representation for time so that time step t_j^k denotes the k^{th} time j is selling item(s) to the market. Similary we define $t_{j,i}^k$ to be the k^{th} time seller j sells an item to buyer i. The cost of the k^{th} transaction is denoted by $c_j(t_j^k)$ (respectively $c_{j,i}(t_j^k)$). At each time step the seller plays an extensive game against some buyer i. The games played at individual time steps are called *stage games* and we assume that their structure is as in Figure 1. The structure is discussed more thoroughly later in this section. The collection of stage games is a *supergame* [14] and we call this collection the *reputation game* between a set of buyers $N\backslash\{j\}$ and a seller j.

The stage games are assumed to be Bayesian games with asymmetric information, i.e. one of the players has more knowledge about the game than the others. In the online auction model the asymmetry results from the market rules as the buyer does not know whether the seller is *HONEST* (delivers the items) or *DISHONEST* (cheats). Thus the seller can have two modes of behaviour which are called *types* [15] in game theoretic terms. The *HONEST* seller never cheats whereas the *DISHONEST* seller can cheat by not submitting anything or by submitting an item of lower quality. We assume that each seller acts as a *HONEST* merchant in a fraction of the transactions and that this fraction is determined by the seller. We use the variable θ_j to denote the fraction of transactions, where the seller j acts honestly and thus, from the point of view of the system, the probability distribution governing the behaviour of the seller is determined by $p(HONEST) = \theta_j$ and $p(DISHONEST) = 1 - \theta_j$. As the probabilities are determined by the seller, they can be different for different buyers. In this case we use $p(HONEST_{j,i}) = \theta_{j,i}$ to be the probability distribution that determines the fraction of transactions where seller j is honest for buyer i.

If we ignore the potential loss of future income, from seller's point of view, the best value for parameter θ is $\theta_j = 0$. However, if future incomes are taken into account, the selection of the parameters becomes more complicated. The process according to which this parameter is actually chosen is not known and for our purposes it is not even necessary to know the details of this process. As higher values of θ_j result in higher loss-rates at a single stage we call the parameter θ_j the *reciprocity rating* of the seller.

Initially, i.e., at time step t_j^1, the buyers have no information about the reciprocity rating of seller j. As more and more transactions are concluded, the amount of available information increases and this should affect the buying decisions of the customers. We assume sequential rationality and initially assign a belief distribution μ over the possible parameter values and update these beliefs using the Bayes' rule as new information is obtained. According to Bayesian decision theory the best estimate of the seller's honesty at an arbitrary time step t_j^k is given by the mode of the belief distribution, which is

$$\hat{\theta}(t_j^k) = \arg\max_{\theta} \mu(\theta). \tag{1}$$

Assuming (for now) that the underlying true distribution of the types remains fixed, the beliefs of the buyers converge to the true distribution. If we can furthermore prove that the individual stage games have optimal strategies it follows that the model admits a *perfect Bayesian equilibrium*. We next show that the stage games indeed have optimal strategies by discussing their structure more thoroughly. For this we consider the k^{th} transaction taking place between a seller j and some buyer i. For notational simplicity, we drop the time indices (t_j^k) in the rest of this section.

We assume that both the buyer and the seller have a valuation for the item(s) that are part of the transaction. The valuations are denoted by v_i and v_j and they represent how much the item(s) are worth to the participants. We assume that the participants have already agreed on the price c_{ij}, but that the money has not been transferred. In this case we are in the root node of the game tree of Figure 1. According to the market rules, the first one to perform an action is the buyer (capital I in the figure). The buyer can still cancel the transaction by not transferring the money (N) or (s)he can go ahead with the transaction (Y). If the buyer hesitates and cancels the transactions, the resulting situation is the same as in the beginning of the stage and thus the utilities of the participants are zero.

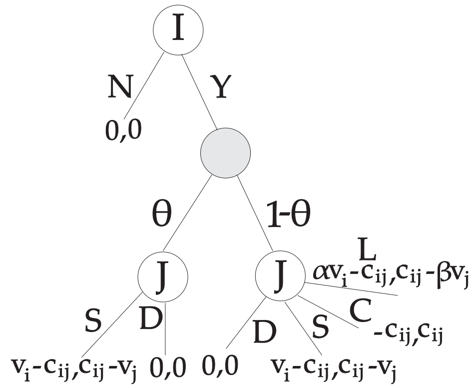

Fig. 1. Single stage of the reputation game.

The more interesting case occurs when the buyer goes ahead with the transaction. If the seller delivers the correct items (S), the buyer loses c_{ij} units of money, but receives the item(s) which have value v_i for the buyer. Thus the utility of the buyer is $v_i - c_{ij}$. Similarly, the seller gains c_{ij} units of money, but loses the item(s) which had the seller the value v_j. Thus the resulting utility for the seller is $c_{ij} - v_j$.

On the other hand, the seller's valuation of the item(s) may change over time so that when (s)he gets the money the difference $c_{ij} - v_j$ is negative and the seller can decline (D) and return the money in which case the utilities of the participants are again zero.

The analysis thus far concludes the behaviour of the honest seller. However, if the seller is dishonest, (s)he has the additional options of not submitting any item(s) (cheating, C) and submitting item(s) of lower quality (L). If the seller decides not to submit anything, the buyer has lost c_{ij} units of money without getting anything. Thus her(his) utility is $-c_{ij}$. Similarly, the seller has obtained c_{ij} units of money without losing anything and thus the seller's utility is c_{ij}.

The more difficult case occurs when the seller decides to submit item(s) of lower quality (L). In this case we can assume that both the buyer and seller have a subjective estimate of the level of quality of the item(s). We use the variable α ($0 \leq \alpha \leq 1$) to denote the buyer's estimate and the variable β ($0 \leq \beta \leq 1$) to denote the seller's estimate. Thus the products of lower quality have a discounted value of αv_i for the buyer. In this case the buyer loses c_{ij} units of money and gains an item with value αv_i which results in an overall utility of $\alpha v_i - c_{ij}$. Similarly, the lower quality item has a value βv_j for the seller and, analogously, the utility of the seller is given by $c_{ij} - \beta v_j$.

Clearly in a single stage game the optimal strategy of the honest seller is to go ahead with the transaction (S) if $c_{ij} - v_j \geq 0$. Similarly, the optimal strategy of the dishonest seller is to cheat whenever $c_{ij} \geq 0$.

The optimal strategy of the buyer is a bit more complicated. If the buyer knows the underlying probability distribution of θ_j, (s)he could act optimally by calculating the expectation of the dominant strategies. However, as the true distribution is unknown, the only possibility for the buyer is to act optimally according to her(his) beliefs. Thus the optimal strategy of the buyer is to go ahead with the transaction if and only if $\hat{\theta}(v_i - c_{ij}) + (1 - \hat{\theta})(-c_{ij}) \geq 0$, i.e., $\hat{\theta} v_i \geq c_{ij}$. As the values of the variables v_i, v_j and c_{ij} are determined at the beginning of the stage, the stage game has equilibrium strategies which are determined according to the discussed rules.

IV. Sequential rationality in the reputation game

The theoretical model discussed in the previous section assumes that we maintain an estimate of the honesty of the seller, but gives no practical insights on how to do this. The solution is to look at the problem from a statistical viewpoint and to consider the sequence of transactions as if it was a sample from a binomial distribution, i.e. a sequence of independent coin tosses, with a bias parameter θ. Thus we can estimate the reciprocity rating using Bayesian parameter estimation of the binomial distribution. In addition, if we use a smoothed Beta distribution as the prior distribution for the binomial density, the Bayesian posterior estimate coincides with the maximum likelihood estimate.

Instead of assuming each transaction as a binary outcome, we take into account the cost of the transaction. Thus instead of estimating the proportion of successful transactions, we estimate the proportion of the cost involved in successful transactions. This way the punishments and rewards of the system make exploiting the system more difficult.

Let $\mathcal{A}_j$ be the set of all transactions where the seller has been j. Of these a subset $\mathcal{C}_j$ has been successful, i.e. the seller was honest. We define the variable $\hat{\theta}_j$ to be the ratio of successful transactions and all transactions, i.e.

$$\hat{\theta}_j = \frac{\sum_{\mathcal{C}_j} c_{ij}}{\sum_{\mathcal{A}_j} c_{ij}}. \tag{2}$$

When we use Bayesian parameter estimation, we assign a Beta-prior on the values of θ_j, which results in a posterior distribution that is also Beta-distributed. The Beta-distribution is a two parameter distribution, whose parameters are denoted by a and b. The parameter a measures the number of successes and, analogously, b measures the number of failures. The bias estimate of the underlying binomial distribution is given by the mode of the Beta distribution which is given by

$$\frac{a-1}{a+b-2}. \tag{3}$$

As our initial parameter values for a and b we use $a = b = 1$. If the k^{th} transaction was successful, we update the parameter a by setting $a^{k+1} = a^k + c_{ij}$ and, if the seller cheated, we update the parameter b by setting $b^{k+1} = b^k + c_{ij}$. Substituting the resulting values into Equation 3 results in the following Equation for the mode of the posterior distribution

$$\begin{aligned} \arg\max_{\hat{\theta}_j} &= \frac{(\sum_{\mathcal{C}_j} c_{ij} + 1) - 1}{(\sum_{\mathcal{C}_j} c_{ij} + 1) + ((\sum_{\mathcal{A}_j} c_{ij} - \sum_{\mathcal{C}_j} c_{ij}) + 1) - 2} \\ &= \frac{\sum_{\mathcal{C}_j} c_{ij}}{\sum_{\mathcal{A}_j} c_{ij}}, \end{aligned} \tag{4}$$

which is the same as the maximum likelihood estimate θ_j. The same formula can also be used to derive an estimate for $\theta_{j,i}$, i.e. for the fraction of times seller j cooperates with an individual buyer i.

V. Integrating Reputations

In online reputation systems the feedback about the behaviour of a seller is derived from interactions of various customers. However, as the reliability of feedback from others is hard to estimate, a reasonable assumption is that individual buyers base their decisions mainly on personal experience and consider the information from others as a confirmation of their own beliefs. Current reputation systems fail to properly aggregate the different information sources so that a personalized reputation rating could be calculated for the individual buyers. Our proposal to this task is to use Bayesian model averaging, in which the overall beliefs are represented as a sum of weighted "component" beliefs. The model components represent the different sources of information and the weights represent the importance of a source. In mathematical terms, the aggregate density of the information sources is

$$p(\theta_j) = \sum_i \frac{\pi_i}{\sum_i \pi_i} p_i(\theta_j), \tag{5}$$

where π_i is the weight of the i^{th} model component and p_i is the underlying probability density of that component. The component weights are intrepreted so that each weight term is an estimate about the reliability of information from that particular source. The values are required to lie in the interval $[0, 1]$.

Each of the component densities p_i maintains an estimate of the honesty of a seller. From a theoretical point of view the internal structure of the component densities is not important and they can be, e.g., individual binomial distributions, *hierarchical models* or a *Bayesian network* with a more complex structure. In addition, the number of components can be quite freely selected. However, for our purposes it suffices to consider a model that consists of two components. Namely, in online markets such as eBay personal experience and experiences of the others are the only information sources that are readily accessible. Another possible source of information are the relationships in the social network of a customer. In the next section we also consider incorporating an estimate of short term behaviour into the model.

Finally, as probabilities are not very intuitive for the user, we need to map them into more meaningful values. This can be accomplished by applying an arbitrary (weakly) monotonic function $f(\cdot)$, which is continuous on the interval $[0, 1]$, on the probabilities. The final reputation values are then given by $f(p(\theta))$. If different portals use a similar scheme, but with different mappings f, the reputation values are comparable across different web-portals.

VI. Extensions

From a theoretical perspective, the generic framework that we have discussed is sufficient for modelling the problem. However, there are some remaining aspects of the model which are more ad hoc by nature and need to be solved using heuristics. First of all, we need to look at how to set the weights and how to adapt them as new information arrives. Secondly, sequential rationality in the model is based on the assumption that the underlying dynamics are stationary which does not hold in practice. Thus we need to discuss the effect of dynamics and how to give estimates about the dynamics of the behaviour.

Setting component weights

Our solution for setting the component weights uses exponential growth for adaptating the weights. We initialize the values of π_i to constant values λ_i and require that the constants sum up to one, i.e., $\sum_i \lambda_i = 1$. These constants are chosen so that magnitudes of the values reflect prior knowledge about the reliability of the different sources.

After having initialized the weights, we derive an update term α_i for weight π_i using the equation

$$\lambda_i \alpha_i^{t_i} = 1.0, \tag{6}$$

where t_i is a positive integer. Updates are then performed by multiplying the current weight π_i with the update term α_i each time new information becomes available.

Intuitively the value of t_i should be smaller for the component representing personal information than for the component representing information from others. In principle these values could be set by a domain expert, but in practice it suffices to use common sense, i.e. to set a small (a magnitude of tens) value for t_1 and much bigger value (a magnitude of thousands) to t_2. The experimental setting uses $t_1 = 10$ and $t_2 = 1000$ and thus, in the experimental setting, personal experience is weighted much more than the information gathered by others.

Optimizing performance under non-stationary dynamics

A drawback of the theoretical model is that, if the behaviour of a seller abruptly changes, the convergence of the reputations to a value that correctly represents the new behaviour can be slow. The convergence speed depends on various factors, such as the size of transaction history and the cost of the transactions. We can overcome some of these problems by giving the users as additional information a short term estimate of the behaviour of the seller.

For deriving the short term estimator we weight the costs $c(t)$ of the most recent transactions using the function $\exp(-n/\mu)$, where the variable $n = 0, 1, 2, \cdots$ represents the n^{th} most recent transaction and μ is a fixed parameter. Otherwise the estimator is derived using the mode of a Beta distribution where the parameters have been estimated from a different subset of informations. Formally the parameter estimate in this case is given by

$$\tilde{\theta}_j = \frac{\sum_{n,\mathcal{C}_j} \exp(-n/\mu) c(t_{k-n}) - 1}{\sum_{n,\mathcal{A}_j} \exp(-n/\mu) c(t_{k-n}) - 2}. \tag{7}$$

VII. Experiments

To provide empirical evidence about the performance of our approach, we implemented a web-based test application called $bHonest$. The application is accessible through the web-site `http://db.cs.helsinki.fi/u/ptnurmi/cgi-bin/bHonest/index.php`. In addition, we implemented a test site that uses the eBay reputation scheme. This version is accessible through `http://db.cs.helsinki.fi/u/ptnurmi/cgi-bin/ebay/index.php`.

In the rest of the section we first discuss details about the implemented systems, after which we present the experimental setting and the results of our experiments.

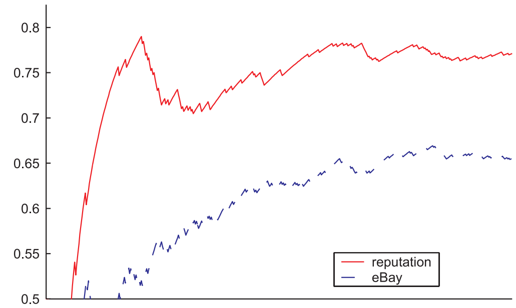

Fig. 2. Ratio of successful transactions when bHonest (solid) and when eBay system are used (dashed).

Experimental system

Both reputation mechanisms were tested using a similar system. In the implemented system anonymous buyers can buy products from anonymous sellers using virtual money. The sellers are added to the system by the administrator whereas the buyers are real humans. The market rules of the system are the same as has been described earlier, i.e. the buyer first submits the money after which the decision of the seller is made. Currently the only supported actions for the seller are cheating and delivering the (virtual) items.

The decision making process of the seller is, by default, based on stochastic rules which are governed by a parameter θ. The parameter θ serves as the baseline cooperation value, i.e. the smaller the value the more often the seller cheats. The values of θ are otherwise fixed, but the administrator can change the values at any time. The stochastic rules can be overruled using defect and coordinate lists which are determined by the administrator. These lists tell to whom the sellers should always sell and whom to always cheat. In normal reputation systems the identities of the buyers and sellers are unknown to each other, but we added this functionality to simulate relationships in social networks.

Experimental setting

Our test setting consisted of two tests. In both tests a fixed number (6) of persons were required to perform a set of transactions in a simulated market. Three of the users used the $bHonest$ system and the three others used the eBay system. The goal of the test subjects was to minimize the sum of costs in unsuccessful transactions.

In the first test, the buyers were required to perform 50 transactions in the market. This setting was repeated three times so that the resulting data set consisted of 900 measurements, of which 450 were from $bHonest$ and 450 from eBay. From each transaction we recorded the selected seller, the price of the transaction and the actual outcome (success vs. failure). During a single test run the marketplace had ten sellers with different parameter values θ_j. These values were initialized at the beginning of each test run to random values after which they were kept fixed during the test run. Both systems were, of course, initialized to use the same parameter values.

In the second test, we gave the reputation values of the buyers' to the test persons and changed randomly the parameters governing the behaviour of the sellers. We applied this test setting two times for two sets of sellers. In the first set, we decreased the parameter values for seven users and increased them for three users. In the second set we reversed these quantities and decreased the values for three sellers and increased for seven sellers. The goal of this test was to see the change in the buyers' performance as the behaviour of the sellers changes.

As a baseline we used the eBay reputation scheme because, although it is not optimal, it works reasonably. In the first test we used two evaluation metrics: the number of successful transactions and the proportion of sum of successful transactions. For the second test we used only the number of successful transactions.

Results

First of all, we measured the ratio of successful transactions in the 450 first transactions (test one). The results are illustrated in Figure 2. Clearly, our system converges much faster and achieves better overall performance. To further support our claim, we performed standard significance testing and calculated the $z-score$ for the difference of the mean performance of the two systems. The test-statistics are shown in Table I and, according to the resulting P-value, the performance increase is statistically significant. Overall, the end performance of the systems differed by 11 percentage units in favour of *bHonest*.

SYSTEM	**Variance**	**Mean**
eBay	31.6667	30.6667
bHonest	34.8095	36.6667
z-score	-2.8501	
p-value	0.0022	

TABLE I

STATISTICS AND TEST-SCORES FOR THE EXPERIMENTS.

In real systems, the users are willing to use higher amounts of money only if they trust the sellers and the system. To measure this, we used the formula in Equation 2, i.e. the proportion of costs in successful transactions as the second evaluation metric. Figure 3 illustrates the results of the first

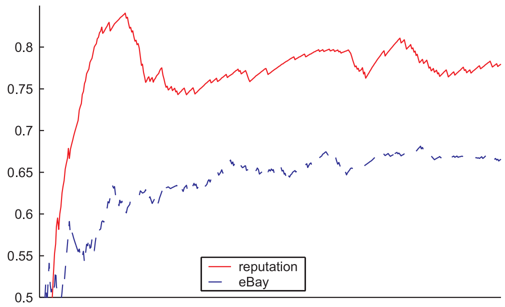

Fig. 3. Ratio of money involved in successful transactions when bHonest is used (solid) and when eBay system is used (dashed).

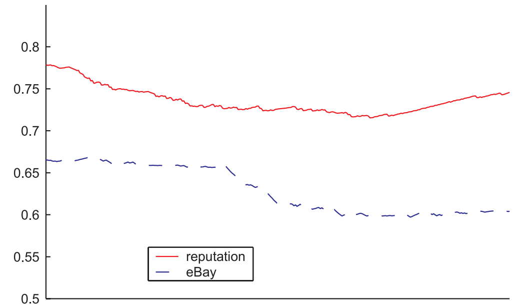

Fig. 4. Ratio of successful transactions after the behaviour of the sellers' abruptly changed. The solid line represents bHonset and the dashed line represents eBay.

test, when this evaluation metric was used. The end performance of *bHonest* was 78% whereas for eBay it was below 67%. Thus again, our system was around 11 percentage units better. Although the test did not result in higher differences, it showed that in more realistic situations the performance of our system is very good.

For evaluating the second test we used only the ratio of successful transactions. The results of this are shown in Figure 4. In this case, the difference in performance varied a lot and the minimum difference was around six percentage units whereas the end performance differed actually by 14 percentage units. The main reason for the "poor" initial performance of our system in this test was that users were not completely familiar with our system. Thus as they became more familiar, the system clearly outperformed eBay.

VIII. Conclusions and Future Work

In this paper we have presented a theoretical model for modelling reputation in online auctions. In addition, we presented a framework that uses Bayesian model averaging for integrating information from multiple sources. Finally, we evaluated our system against the binary reputation scheme used on eBay and, as the results indicate, our system usually outperforms eBay by more than 10% percent.

The work conducted in this paper offers directions for both theoretical and practical future work. Currently we ignored such options as the buyers giving intentionally bad feedback and the effect of discount factors on the selection of the stragegy of submitting an item of lower quality. In practical settings our first goal is to apply the proposed framework to ad hoc and sensor networks. In the field of online auctions, our goal in the future is to consider privacy and security related aspects of our model.

References

[1] (2004) eBay. http://www.ebay.com/. [Online]. Available: http://www.ebay.com/

[2] (2004) Epinions. http://www.epinions.com/. [Online]. Available: http://www.epinions.com/

[3] L. Mui, M. Mohtashemi, and A. Halberstadt, "Notions of reputation in multi-agents systems: a review," in *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems*. ACM Press, 2002, pp. 280–287.

[4] J. A. Hoeting, D. Madigan, A. E. Raftery, and C. T. Volinsky, "Bayesian model averaging: A tutorial," *Statistical Science*, vol. 4, pp. 382 – 417, 1999.

[5] G. A. Akerlof, "The market for lemons: Quality uncertainty and the market mechanism," *The Quarterly Journal of Economics*, vol. 84, no. 3, pp. 488–500, Aug 1970.

[6] D. M. Kreps and R. Wilson, "Sequential equilibria," *Econometrica*, vol. 50, no. 4, pp. 863 – 894, July 1982.

[7] ——, "Reputation and imperfect information," *Journal of Economic Theory*, vol. 27, pp. 253–279, 1982.

[8] P. Milgrom and J. Roberts, "Predation, reputation and entry deterrence," *Journal of Economic Theory*, vol. 27, pp. 280–312, 1982.

[9] D. Fudenberg and D. K. Levine, "Reputation and equilibrium selection in games with a patient player," *Econometrica*, vol. 57, no. 4, pp. 759–778, July 1989.

[10] P. Resnick, R. Zeckhauser, J. Swanson, and K. Lockwood, "The value of reputation on eBay: A controlled experiment," Working paper, Michigan School of Information, Harvard University and eBay, March 2004. [Online]. Available: http://www.si.umich.edu/~presnick/papers/postcards/index.html

[11] L. Mui, M. Mohtashemi, C. Ang, P. Szolovits, and A. Halberstadt, "Ratings in distributed systems: A bayesian approach," in *Proceedings of the Workshop on Information Technologies and Systems*, 2001.

[12] G. Zacharia, A. Moukas, and P. Maes, "Collaborative reputation mechanisms in electronic marketplaces," in *Proceedings of the 32nd Hawaii International Conference on System Sciences*. IEEE Computer Society, 1999.

[13] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, "The EigenTrust algorithm for reputation management in P2P networks," in *Proceedings of the 12th International Conference on World Wide Web (WWW)*. ACM Press, 2003, pp. 640 – 651.

[14] R. Selten, "The chain-store paradox," *Theory and Decision*, vol. 9, pp. 127–159, 1978.

[15] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, Massachusetts: MIT Press, 1991.