Improving User Satisfaction in Agent-Based Electronic Marketplaces by Reputation Modelling and Adjustable Product Quality

Thomas Tran and Robin Cohen
School of Computer Science, University of Waterloo
Waterloo, ON, N2L 3G1, Canada
{tt5tran, rcohen}@uwaterloo.ca

Abstract

In this paper, we propose a market model and learning algorithms for buying and selling agents in electronic marketplaces. We take into account the fact that multiple selling agents may offer the same good with different qualities, and that selling agents may alter the quality of their goods. We also consider the possible existence of dishonest selling agents in the market. In our approach, buying agents learn to maximize their expected value of goods using reinforcement learning. In addition, they model and exploit the reputation of selling agents to avoid interaction with the disreputable ones, and therefore to reduce the risk of purchasing low value goods. Our selling agents learn to maximize their expected profits by using reinforcement learning to adjust product prices, and also by altering product quality to provide more customized value to their goods. This paper focuses on presenting results from experiments investigating the behaviours of buying and selling agents in large-sized electronic marketplaces. Our results confirm that buying and selling agents following the proposed algorithms obtain greater satisfaction than buying and selling agents who only use reinforcement learning, with the buying agents not modelling sellers' reputation and the selling agents not adjusting product quality.

1. Introduction

In this paper, we present a framework for designing an electronic marketplace populated with buying and selling agents that allows the quality of goods being offered by selling agents to vary, over time and which equips buying agents with a reputation modelling mechanism, in order to learn to purchase the goods that are of the greatest value to them.

In particular, we allow our marketplace to be open (agents can freely enter or leave), dynamic (price and quality of goods may be altered by selling agents, in order to meet buying agents' specific needs), uncertain (a buying agent can only examine the quality of the good it purchases after it receives the good from the selected selling agent) and untrusted (there may be dishonest agents in the environment).

The selling agents learn to adjust price and quality in order to maximize profit. In particular, we show how to combine the adjustment of these two factors, to increase user satisfaction.

Buying agents are allowed to have different personal preferences over the goods sold by selling agents (varying relative value of price and quality). The buying agents learn to avoid dishonest sellers and to increase the satisfaction of their users by modelling the reputation of selling agents and focusing their business on those agents with whom they have established a certain degree of trust. In particular, we present a specific strategy for modelling reputation and for updating the reputation ratings of selling agents and show how to integrate this effectively with reinforcement learning. Moreover, our buying agents have the opportunity to periodically explore the marketplace in order to discover new sellers or to allow selling agents that have made adjustments to the quality and price of their goods to be selected for business. This supports the required openness of the environment.

We experimentally measure the value of our proposed algorithms for buying and selling agents, concluding that buying and selling agents using our approach fare better than ones using reinforcement learning alone. We also discuss how our proposed framework compares with other methods for buying and selling agents to improve their performance by modelling each other in the marketplace.

This research therefore provides the basis for incorporating models of trust with variable quality of goods in electronic marketplaces, producing satisfaction for the users of these buying and selling agents.

AAMAS'04, July 19-23, 2004, New York, New York, USA.

2. The Proposed Algorithms

In this section we present our agent market model and propose the learning algorithms for buying and selling agents in electronic marketplaces, based on reputation modelling and reinforcement learning.

2.1. The Agent Market Model

We model the agent environment as an open marketplace which is populated with economically-motivated agents. The nature of an open marketplace allows the economic agents, which we classify as *buyers* and *sellers*, to freely enter or leave the market. Buyers and sellers are self-interested agents whose goal is to maximize their own benefit.

Our market environment is rooted in an information delivery infrastructure such as the Internet, which provides agents with virtually direct and free access to all other agents. The process of buying and selling goods is realized via a *contract-net* like mechanism [2, 6], which consists of three elementary phases:

(i) When a buyer b is in need of some good g, it will announce its request for that good to all the sellers in the market, using multi-cast or possibly broadcast.

(ii) After receiving the request from b, those sellers that have good g available for sales will send a message to b, stating their price bids for delivering the good.

(iii) Buyer b evaluates the submitted bids and selects a suitable seller to purchase good g. Buyer b then pays the chosen seller and receives the good from that seller.

Thus, the buying and selling process can be viewed as an *auction* where sellers play the role of bidders and buyers play the role of auctioneers, and a seller is said to be *winning the auction* if it is able to sell its good to the buyer.

To make our marketplace more realistic and also more interesting, we assume that

- The quality of a good offered by different sellers may not be the same.
- A seller may alter the quality (in addition to the price) of its goods.
- It is possible that some dishonest sellers exist in the market.
- A buyer can examine the quality of the good it purchases only after it receives that good from the selected seller.
- Each buyer has some way to evaluate the good it purchases, based on the price and the quality of the good received.

AAMAS'04, July 19-23, 2004, New York, New York, USA.

2.2. Buying Algorithm

Consider the scenario where a buyer b announces its request for some good g. Let G be the set of goods, P be the set of prices, and S be the set of all sellers in the marketplace. G, P, and S are finite sets.

Buyer b models the reputation of all sellers in the market using function $r^b : S \mapsto (-1, 1)$, which is called the *reputation function* of b. Initially, buyer b sets the *reputation rating* $r^b(s) = 0$ for every seller $s \in S$. After each transaction with a seller s, buyer b will update (increase or decrease) $r^b(s)$ depending on whether or not s satisfies b in the transaction. A seller s is considered *reputable* by buyer b if $r^b(s) \geq \Theta$, where Θ is buyer b's *reputation threshold* ($0 < \Theta < 1$). A seller s is considered *disreputable* by buyer b if $r^b(s) \leq \theta$, where θ is buyer b's *disreputation threshold* ($-1 < \theta < 0$). A seller s with $\theta < r^b(s) < \Theta$ is neither reputable nor disreputable to buyer b. In other words, b does not have enough information to decide on the reputation of s. Let S^b_r and S^b_{dr} be the sets of reputable and disreputable sellers to buyer b respectively, i.e.,

$$S^b_r = \{s \in S \mid r^b(s) \geq \Theta\} \subseteq S, \tag{1}$$

and

$$S^b_{dr} = \{s \in S \mid r^b(s) \leq \theta\} \subseteq S. \tag{2}$$

Buyer b will focus its business on the reputable sellers and stay away from the disreputable ones.

Buyer b estimates the expected value of the goods it purchases using the *expected value function* $f^b : G \times P \times S \mapsto \mathbb{R}$. Hence, the real number $f^b(g, p, s)$ represents buyer b's expected value of buying good g at price p from seller s.

Since multiple sellers may offer good g with different qualities and a seller may alter the quality of its goods, buyer b puts more trust in the sellers with good reputation. Thus, it chooses among the reputable sellers in S^b_r a seller $\hat{s}$ that offers good g at price p with maximum expected value:

$$\hat{s} = \arg\max_{s \in S^b_r} f^b(g, p, s), \tag{3}$$

where $\arg$ is an operator such that $\arg f^b(g, p, s)$ returns s.

If no sellers in S^b_r submit bids for delivering g (or if $S^b_r = \emptyset$), then buyer b will have to choose a seller $\hat{s}$ from the non-reputable sellers, provided that $\hat{s}$ is not a disreputable seller:

$$\hat{s} = \arg \max_{s \in (S - (S^b_r \cup S^b_{dr}))} f^b(g, p, s). \tag{4}$$

In addition, with a small probability ρ, buyer b chooses to explore (rather than exploit) the marketplace by randomly selecting a seller $\hat{s} \in (S - S^b_{dr})$. This gives buyer b an opportunity to discover new reputable sellers. Initially, the value of ρ should be set to 1, then decreased over time to some fixed minimum value determined by b.

After paying seller $\hat{s}$ and receiving good g, buyer b can examine the quality $q \in Q$ of good g, where Q is a finite set of real values representing product qualities. It then calculates the true value of good g using the *true product value function* $v^b : G \times P \times Q \mapsto \mathbb{R}$. For instance, if buyer b considers the quality of good g to be twice more important than its price, it may set $v^b(g,p,q) = 2q - p$.

The expected value function f^b is now incrementally learned in a reinforcement learning framework:

$$\Delta = v^b(g,p,q) - f^b(g,p,\hat{s}), \tag{5}$$

$$f^b(g,p,\hat{s}) \leftarrow f^b(g,p,\hat{s}) + \alpha\Delta, \tag{6}$$

where α is called the *learning rate* ($0 \leq \alpha \leq 1$). Similar to ρ, the learning rate α should initially be set to a starting value of 1 and then reduced over time to a fixed minimum value chosen depending on individual buyers.

Thus, if $\Delta = v^b(g,p,q) - f^b(g,p,\hat{s}) \geq 0$ then $f^b(g,p,\hat{s})$ is updated with the same or a greater value than before. This means that seller $\hat{s}$ has a good chance to be chosen by buyer b again if it continues offering good g at price p in the next auction. Conversely, if $\Delta < 0$ then $f^b(g,p,\hat{s})$ is updated with a smaller value than before. This implies that seller $\hat{s}$ may not be selected by buyer b in the next auction if it continues selling good g at price p.

In addition to updating the expected value function, the reputation rating $r^b(\hat{s})$ of seller $\hat{s}$ also needs to be updated. Let $\vartheta^b(g) \in \mathbb{R}$ be the product value that buyer b demands for good g. In other words, the demanded product value $\vartheta^b(g)$ is buyer b's threshold for the true product value $v^b(g,p,q)$. We use a reputation updating scheme motivated by that proposed in [11] as follows:

If $v^b(g,p,q) - \vartheta^b(g) \geq 0$, that is, if seller $\hat{s}$ offers good g with value greater than or equal to the value demanded by buyer b, then its reputation rating $r^b(\hat{s})$ is increased by

$$r^b(\hat{s}) \leftarrow \begin{cases} r^b(\hat{s}) + \mu(1 - r^b(\hat{s})) & \text{if } r^b(\hat{s}) \geq 0, \\ r^b(\hat{s}) + \mu(1 + r^b(\hat{s})) & \text{if } r^b(\hat{s}) < 0, \end{cases} \tag{7}$$

where μ is a positive factor called the *cooperation factor*[1] ($\mu > 0$).

Otherwise, if $v^b(g,p,q) - \vartheta^b(g) < 0$, that is, if seller $\hat{s}$ sells good g with value less than that demanded by buyer b, then its reputation rating $r^b(\hat{s})$ is decreased by

$$r^b(\hat{s}) \leftarrow \begin{cases} r^b(\hat{s}) + \nu(1 - r^b(\hat{s})) & \text{if } r^b(\hat{s}) \geq 0, \\ r^b(\hat{s}) + \nu(1 + r^b(\hat{s})) & \text{if } r^b(\hat{s}) < 0, \end{cases} \tag{8}$$

where ν is a negative factor called the *non-cooperation factor*[2] ($\nu < 0$).

1 Buyer b will consider seller $\hat{s}$ as being *cooperative* if the good $\hat{s}$ sells to b has value greater than or equal to that demanded by b.

2 Buyer b will consider seller s as being *non-cooperative* if the good $\hat{s}$ sells to b has value less than that demanded by b.

The set of reputable sellers to buyer b now needs to be updated based on the new reputation rating $r^b(\hat{s})$, as in one of the following two cases:

- If $(\hat{s} \in S_r^b)$ and $(r^b(\hat{s}) < \Theta)$ then buyer b no longer considers $\hat{s}$ as a reputable seller, i.e.,

$$S_r^b \leftarrow S_r^b - \{\hat{s}\}. \tag{9}$$

- If $(\hat{s} \notin S_r^b)$ and $(r^b(\hat{s}) \geq \Theta)$ then buyer b now considers $\hat{s}$ as a reputable seller, i.e.,

$$S_r^b \leftarrow S_r^b \cup \{\hat{s}\}. \tag{10}$$

Finally, the set of disreputable sellers also needs to be updated:

- If $(\hat{s} \notin S_{dr}^b)$ and $(r^b(\hat{s}) \leq \theta)$ then buyer b now considers $\hat{s}$ as a disreputable seller, i.e.,

$$S_{dr}^b \leftarrow S_{dr}^b \cup \{\hat{s}\}. \tag{11}$$

2.2.1. Setting μ and ν The co-operation and non-cooperation factors, μ and ν, are used to adjust the reputation ratings of sellers once the buyer has examined the quality of the good purchased.

To protect itself from dishonest sellers, buyer b may require $|\nu| > |\mu|$ to implement the traditional assumption that reputation should be difficult to build up, but easy to tear down. Moreover, buyer b may vary μ and ν as increasing functions of the true product value v^b to reflect the common idea that a transaction with higher value should be more appreciated than a lower one (i.e., the reputation rating of a seller that offers higher true product value should be better increased and vice versa).

In particular, we propose the following equations for the calculation of μ and ν. If $v^b(g,p,q) - \vartheta^b(g) \geq 0$, we define the cooperation factor μ as

$$\mu = \begin{cases} \dfrac{v^b(g,p,q) - \vartheta^b(g)}{\Delta v^b} & \text{if } \frac{v^b(g,p,q) - \vartheta^b(g)}{\Delta v^b} > \mu_{min}, \\ \mu_{min} & \text{otherwise}, \end{cases} \tag{12}$$

where $\Delta v^b = v^b_{max} - v^b_{min}$ with v^b_{max} and v^b_{min} being the maximum and minimum value of the true product value function $v^b(g,p,q)$[3]. We prevent μ from becoming zero when $v^b(g,p,q) = \vartheta^b(g)$ by using the value μ_{min}.

However, if $v^b(g,p,q) - \vartheta^b(g) < 0$, we define the non-cooperation factor ν as

$$\nu = \lambda\left(\frac{v^b(g,p,q) - \vartheta^b(g)}{\Delta v^b}\right), \tag{13}$$

where λ is called the *penalty factor* ($\lambda > 1$) to implement the above mentioned idea that $|\nu|$ should be greater than $|\mu|$.

3 $v^b max$ and v^b_{min} are derived from the maximum and minimum elements of the finite sets P and Q.

AAMAS'04, July 19-23, 2004, New York, New York, USA.

If applying equation (8) using ν as defined in (13) results in the updated value $r^b(\hat{s}) \leq -1$, that is, seller $\hat{s}$ is so non-cooperative, then buyer b will place $\hat{s}$ in the disreputable set S^b_{dr} by setting $r^b(\hat{s}) = \theta$.

2.3. Selling Algorithm

Consider the scenario where a seller $s \in S$ has to decide on the price to sell some good g to a buyer b. Let B be the (finite) set of buyers in the marketplace, and let function $h^s : G \times P \times B \mapsto \mathbb{R}$ estimate the expected profit for seller s. Thus, the real number $h^s(g,p,b)$ represents the expected profit for seller s if it sells good g at price p to buyer b. Let $c^s(g,b)$ be the cost of seller s to produce good g for buyer b. Note that seller s may produce various versions of good g, which are tailored to meet the needs of different buyers. Seller s will choose a price $\hat{p}$ greater than or equal to cost $c^s(g,b)$ to sell good g to buyer b such that its expected profit is maximized:

$$\hat{p} = \arg \max_{\substack{p \in P \\ p \geq c^s(g,b)}} h^s(g,p,b), \tag{14}$$

where in this case $\arg$ is an operator such that $\arg h^s(g,p,b)$ returns p.

The expected profit function h^s is learned incrementally using reinforcement learning:

$$h^s(g,p,b) \leftarrow h^s(g,p,b) + \alpha(\phi^s(g,p,b) - h^s(g,p,b)), \tag{15}$$

where $\phi^s(g,p,b)$ is the actual profit of seller s if it sells good g at price p to buyer b, and is defined as follows:

$$\phi^s(g,p,b) = \begin{cases} p - c^s(g,b) & \text{if seller } s \text{ wins the auction,} \\ 0 & \text{otherwise.} \end{cases} \tag{16}$$

Thus, if seller s does not win the auction then $(\phi^s(g,p,b) - h^s(g,p,b))$ is negative, and by (15), $h^s(g,p,b)$ is updated with a smaller value than before. This means that price $\hat{p}$ will probably not be chosen again to sell good g to buyer b in future auctions, but rather some lower price will. Conversely, if seller s wins the auction then price $\hat{p}$ will probably be re-selected in future auctions.

If seller s succeeded in selling good g to buyer b once, but subsequently fails for a number of auctions, say for m consecutive auctions (where m is seller s specific constant), then it may not only be because s has set a too high price for good g, but probably also because the quality of g does not meet buyer b's expectation. Thus, in addition to lowering the price via equation (15), seller s may optionally add more value (quality) to g by increasing its production cost[4]:

$$c^s(g,b) \leftarrow (1 + Inc)c^s(g,b), \tag{17}$$

4 This supports the common assumption that it costs more to produce high quality goods.

where Inc is seller s specific constant called the *quality increasing factor*.

In contrast, if seller s is successful in selling good g to buyer b for n consecutive auctions, it may optionally reduce the quality of good g, and thus try to further increase its future profit:

$$c^s(g,b) \leftarrow (1 - Dec)c^s(g,b), \tag{18}$$

where Dec is seller s specific constant called the *quality decreasing factor*.

3. Experimental Results

We have performed extensive experimentation to measure the value of our model on both microscopic and macroscopic levels. On the micro level, we were interested in examining the individual benefit of agents, particularly their level of satisfaction. Our experimental results confirm that in both modest and large-sized marketplaces, buyers and sellers following our proposed algorithms achieve better satisfaction than buyers and sellers who only use reinforcement learning, with the buyers not modelling sellers' reputation and the sellers' not adjusting product quality [7]. On the macro level, we studied how a marketplace populated with our buyers and sellers would behave as a whole. Our results show that such a marketplace can reach an equilibrium state where the agent population remains stable (as some sellers who repeatedly fail to sell their goods will decide to leave the market), and that this equilibrium is beneficial for the participant agents [7].

Due to the page limit, this paper only focuses on presenting the micro experimental results of large-sized marketplaces. We simulate a large marketplace populated with 160 sellers and 120 buyers using Java 2. The seller population is divided into four groups:

- Group A consists of seller s_0, s_1, ..., and s_{39} that offer goods with quality chosen randomly from the interval $[32.0, 42.0]$.
- Group B consists of seller s_{40}, s_{41}, ..., and s_{79}. These are dishonest sellers who try to attract buyers with high quality goods ($q = 45$) and then cheat them with really low quality ones ($q = 1$).
- Group C consists of seller s_{80}, s_{81}, ..., and s_{119} that offer goods with fixed quality $q = 39.0$.
- Group D consists of seller s_{120}, s_{121}, ..., and s_{159} that also offer goods with an initial quality of 39.0. However, these sellers follow the proposed selling algorithm to improve the quality of their goods.

The buyer population is divided into two groups:

- Group I consists of buyer b_0, b_1, ..., and b_{59}. These buyers use reinforcement learning alone and do not model sellers' reputation.
- Group II consists of buyer b_{60}, b_{61}, ..., and b_{119}. These buyers follow the proposed buying algorithm.

Other parameters are as follows:

- Quality is chosen equal to cost to support the common assumption that it costs more to produce high quality goods. The true product value function $v^b(p,q) = 3.5q - p$, and the demanded product value $\vartheta^b(g) = 100$. Thus, even when a seller has to sell at cost, it must offer goods with quality of at least 40 in order to meet the buyers' requirement[5].
- If $v^b - \vartheta^b \geq 0$, we define the cooperation factor μ as in equation (12), where $\mu_{min} = 0.005$, $v^b_{max} = 3.5q_{max} - p_{min}$, $v^b_{min} = 3.5q_{min} - p_{max}$, $q_{max} = p_{max} = 49.0$, and $q_{min} = p_{min} = 1.0$. In this definition, we vary μ as an increasing function of v^b to reflect the idea that the reputation rating of a seller that offers higher product value should be better increased. We prevent μ from becoming zero when $v^b = \vartheta^b$ by using the value of μ_{min}.
- If $v^b - \vartheta^b < 0$, we define the noncooperation factor ν as in equation (13), where we set the penalty factor $\lambda = 3$. In this definition, we also vary ν as an increasing function of v^b to support the idea that the lower product value a seller offers, the more its reputation rating should be decreased. The use of factor $\lambda > 1$ indicates that a buyer will penalize a non-cooperative seller λ times greater than it will award a cooperative seller. This implements the traditional assumption that reputation should be difficult to build up, but easy to tear down.
- ρ and α are both initially set to 1 and decreased over time (by factor 0.9997) down to $\rho_{min} = \alpha_{min} = 0.1$. $\Theta = 0.5$, $\theta = -0.9$, $m = n = 10$, and $Inc = Dec = 0.05$.

It should be obvious from the settings that a successful buyer should focus its business on group D of sellers and try to keep away from the other groups.

3.1. Buyers' Satisfaction

We compare the satisfaction of a buyer following the proposed algorithm with that of a buyer not modelling sellers' reputation by looking at the numbers of purchases they make to the four groups of sellers, and alternatively the histograms of true product values that they obtain. We are also interested in seeing how much better the buyer following the proposed algorithm is able to avoid interaction with the group of dishonest sellers. The experimental results reported here are based on the average taken over the respective populations of the two groups of buyers.

Table 1 shows the number of purchases made to the four groups of sellers by the buyer not modelling sellers' reputation (labelled as b_I), and the buyer following the proposed algorithm (labelled as b_{II}). As showed in the table, buyer b_{II} makes about 315, 490, and 406 fewer purchases (or 33.6%, 75.4%, and 33.9% fewer) from group A, B, and C of sellers respectively, compared to the number of purchases made to these three groups of sellers by buyer b_I. In addition, buyer b_{II} makes approximately 1210 more purchases (or 54.6% more) from group D of sellers, compared to the number of purchases made to that group by buyer b_I. In other words, buyer b_{II} focuses its business on the best group of sellers (group D) and stays away from the undesired ones, and therefore obtains better satisfaction.

	Group A	Group B	Group C	Group D
b_I	937.0	650.2	1196.0	2216.8
b_{II}	622.2	160.0	790.3	3427.5

Table 1. Number of purchases made to four groups of sellers by the buyer not modelling sellers' reputation (b_I), and the buyer following the proposed algorithm (b_{II}).

As an alternative view, Figure 1(a) and (b) below present the histograms of product values obtained by a buyer not maintaining reputation ratings of sellers, and by a buyer following the proposed buying algorithm, respectively. The histograms clearly show that the buyer following the proposed algorithm receives fewer goods with low values (65 - 105) and more goods with high value (110), and is therefore better satisfied. In particular, the buyer following the proposed algorithm makes about 2400 more purchases with high mean product value of 110 (or about 16 times greater) than those made by the buyer not modelling sellers' reputation.

We are also interested in seeing how much better the buyer using the proposed algorithm is able to avoid interaction with the group of dishonest sellers (i.e., group B), compared to the buyer not modelling sellers' reputation. Figure 2(a) and (b) show the profits made by the dishonest sellers from the buyer not modelling sellers' reputation and from the buyer following the proposed algorithm, respectively.

5 Because $v^b(p,q) = 3.5q - p$ and $3.5(40) - 40 = 100$.

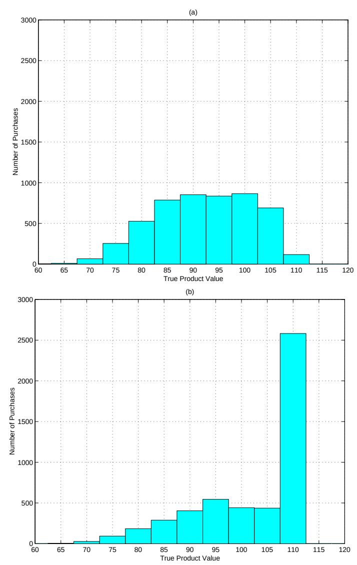

Figure 1. Histograms of true product values obtained by a buyer not modelling sellers' reputation (a), and by a buyer following the proposed buying algorithm (b).

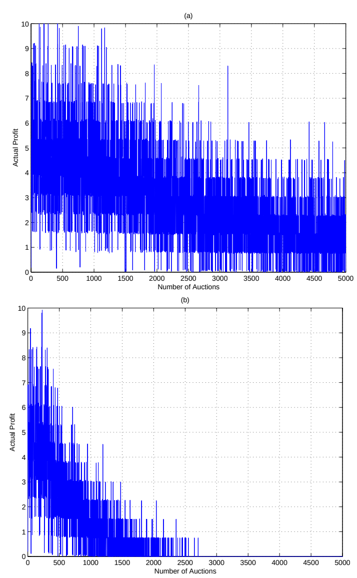

Figure 2. Graphs of profit values made by the dishonest sellers from a buyer not modelling sellers' reputation (a), and from a buyer following the proposed buying algorithm (b).

We can see that graph (a) is higher than graph (b), indicating that the dishonest sellers are able to make more profit from those buyers that do not model sellers' reputation. Moreover, the profit in graph (b) is reduced to zero after about 2700 auctions, implying that from that point on the dishonest sellers are not able to make any profit from the buyer following the proposed algorithm, since they are considered as disreputable sellers and therefore no longer chosen by the buyer.

3.2. Sellers' Satisfaction

We compare the satisfaction level of the four groups of sellers by examining their sales and profits made to a buyer. The results shown are based on the average taken over the population of all buyers in the market.

Table 2 presents the number of sales made by the four groups of sellers to a buyer.

Group A	Group B	Group C	Group D
779.6	405.1	993.2	2822.1

Table 2. Number of sales made by four groups of sellers to a buyer.

Group D is able to make the most number of sales. In particular, the number of sales made by this group is approximately 3.6 times greater than that made by group A, 7 times greater than that made by group B, and 2.8 times greater than that made by group C, respectively.

Alternatively, Figure 3(a), (c), and (d) show the graphs of profit values made from a buyer by group A, C, and D of sellers, respectively. The goods offered by sellers in group A usually do not meet the buyers' need since their quality is chosen randomly. As a result, this group of sellers receives low profit (graph (a)). The dishonest sellers in group B attract buyers with high quality goods and then cheat them with really low quality ones to make big profit. Consequently, their sales are on and off (mostly made to the group of buyers that do not model sellers' reputation as shown in Figure 2), resulting in greatly fluctuating profit (graph (b)). Group C of sellers offers goods with fixed quality and is able to make relatively high profit in the first 1500 auctions. However, as the sellers in group D improve the quality of their goods, the sellers in group C start losing their sales in the long run. Graph (c) shows that their profit begins to go down after about 1500 auctions, and reaches the mean of about 0.25 after 3500 auctions. Although sellers in group D start with relatively low quality goods, they consider improving the quality of their goods according to the proposed selling algorithm. As a result, they make more and more sales and their profit increases substantially after 1500 auctions, reaching the mean of about 1.25 (graph (d)), which is five times greater than that of group C.

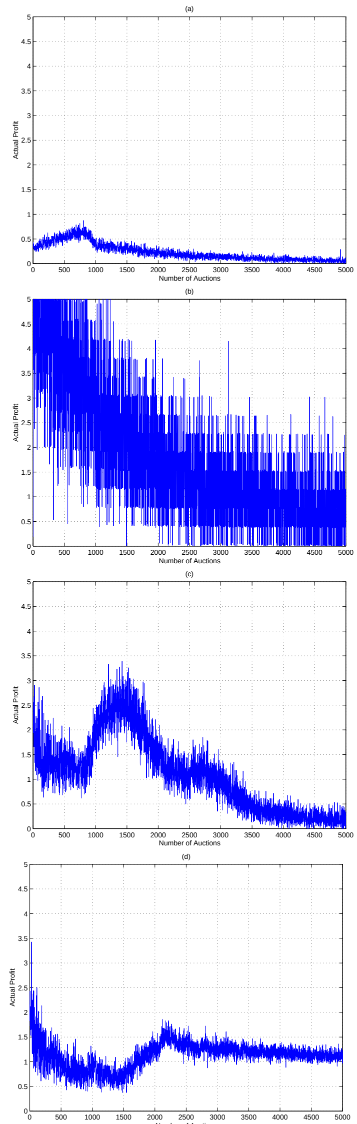

Figure 3. Graphs of actual profit made by group A (a), group B (b), group C (c), and group D of sellers (d).

In summary, our experimental results show that in large-sized marketplaces, buyers and sellers following our proposed algorithms achieve better satisfaction than buyers and sellers who only use reinforcement learning but the buyers do not model sellers' reputation and the sellers do not consider adjusting product quality. Similar results are also confirmed in experiments of modest-sized marketplaces, which are omitted here due to the paper's page limit.

4. Discussion and Conclusion

This research provides an effective mechanism for modelling reputation in electronic marketplaces where sellers may alter the quality of their goods, over time. A number of researchers have investigated the modelling of reputation. Yu and Singh develop a general model for trust, focusing on acquiring information from other agents in an agent community [11]. Their scheme to update the trust rating of agents uses constant factors and does not take into consideration the extent to which an agent has (or has not) cooperated. That is, a greatly cooperative agent and a slightly cooperative agent will receive the same increasing amount in their trust ratings. Similarly, a greatly disappointing agent and a slightly disappointing agent will receive the same decreasing amount in their trust ratings. In contrast, we have variable cooperative and non-cooperative factors, to allow for agents who greatly disappoint to be more seriously penalized. In addition, we propose a specific formula for setting the co-operation and non-cooperation factors, to provide protection from potentially dishonest sellers.

We can in fact prove that with μ, ν defined according to formulas (12) and (13), a buying agent cannot be infinitely harmed by a dishonest seller if it sets the penalty factor λ properly[6]. This reinforces the importance of our particular approach in untrusted environments.

Esfandiari and Chandrasekharan [3] introduce a model for trust acquisition and trust propagation. However, their definition of trust does not make a clear distinction between distrust and lack of knowledge about trust. In contrast, our approach uses a reputation function that maps sellers to the range $(-1, 1)$, allowing a buyer to assign the neutral value of zero to new sellers that it has no experience about. In addition, in contrast to [3], our proposed reputation mechanism enables buying agents to quickly identify the reputable sellers while avoiding the disreputable ones. This is achieved by introducing the reputation and disreputation thresholds and incorporating the reputable and disreputable sets into the buying algorithm.

Our research also compliments that of Breban and Vassileva [1], that examines the value of modelling trust in order to form coalitions in multi-agent environments. Although their focus is on how best to organize agents into groups, the work is consistent with ours in that it allows for an evolution of trust over time and advocates the value of reaching equilibrium within agent societies.

This research also provides an alternative to the design of buying and selling agents in electronic marketplaces,

6 The proof is omitted here, for lack of space. See [7] for details.

in comparison with the work of Vidal and Durfee [8, 9]. Their approach relies on recursive modelling in order to make effective purchasing decisions. In particular, they examine when an agent benefits from having deeper models of others, including 0-level agents that learn from their observations of the environment and from any environmental rewards they receive, 1-level agents that model others as 0-level agents, 2-level agents that model others as 1-level agents, etc. As pointed out in [8, 9], the agents with deeper recursive models of others suffer from the computational costs associated with maintaining these deep models. The use of a reputation model, combined with reinforcement learning, in our work, offers an alternative mechanism for buying agents to avoid sellers who are likely to disappoint. It also contrasts with the model of Vidal and Durfee, in that it uses the disreputable set to explicitly identify and subsequently ignore the dishonest agents of the marketplace. In addition, Vidal and Durfee's model does not allow sellers to alter the quality of their goods. In contrast, we provide this as an option for selling agents to increase the satisfaction of their users, while still allowing buying agents to operate effectively in such a dynamic marketplace[7].

Our work is also related to other efforts that show the merit of reinforcement learning for multi-agent systems [4, 5]; our focus has been on applying this form of learning in application domains where the agents are economically motivated and act in open market environments.

In conclusion, we have developed a framework for the design of electronic marketplaces that allows selling agents to alter both the price and the quality of their goods and allows buying agents to adjust their purchasing decisions, based on a combination of reinforcement learning and reputation modelling. This demonstrates that a model can effectively integrate reputation with reinforcement learning and that the adjustment of product quality can indeed be supported. The experimental results provided here serve to confirm the value of this approach in providing effective satisfaction for users of buying and selling agents in these electronic marketplaces.

References

[1] S. Breban, and J. Vassileva. Using Inter-agent Trust Relationships for Efficient Coalition Formation. In *Proceedings of the Fifteenth Conference of the Canadian Society for Computational Studies of Intelligence*, pages 221-236, May 2002.

[2] R. Davis, and R. G. Smith. Negotiation as a Metaphor for Distributed Problem Solving. In *Artificial Intelligence*, 20(1): 63-109, January 1983.

[3] B. Esfandiari and S. Chandrasekharan. On How Agents Make Friends: Mechanisms for Trust Acquisition. In *Proceedings of the Fifth International Conference on Autonomous Agents Workshop on Deception, Fraud and Trust in Agent Societies*, pages 27-34, 2001.

[4] T. W. Sandholm, and R. H. Crites. Multi-Agent Reinforcement in the Iterated Prisoner's Dilemma. In *Biosystems*, 37: 147-166, 1995.

[5] S. Sen, M. Sekaran, and J. Hale. Learning to Coordinate without Sharing Information. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, pages 426-431, 1994.

[6] R. G. Smith. The Contract Net Protocol: High Level Communication and Control in a Distributed Problem Solver. In *IEEE Transactions on Computers*, C-29(12): 1104-1113, December 1980.

[7] Thomas Tran. Reputation-Oriented Reinforcement Learning Strategies for Economically-Motivated Agents in Electronic Market Environments. PhD Thesis, School of Computer Science, University of Waterloo, 2003.

[8] J. M. Vidal, and E. H. Durfee. The Impact of Nested Agent Models in an Information Economy. In *Proceedings of the Second International Conference on Multi-Agent Systems*, pages 377-384, 1996.

[9] J. M. Vidal. Computational Agents That Learn About Agents: Algorithms for Their Design and A Predictive Theory of Their Behavior. PhD Thesis, Department of Computer Science & Engineering, University of Michigan, 1998.

[10] P. R. Wurman, M. P. Wellman, and W. E. Wash. The Michigan Internet AuctionBot: A Configurable Auction Server for Humans and Software Agents. In *Proceedings of the Second International Conference on Autonomous Agents*, pages 301-308, 1998.

[11] B. Yu, and M. P. Singh. A Social Mechanism of Reputation Management in Electronic Communities. In M. Klusch and L. Kerschberg, editors, *Cooperative Information Agents IV*, Lecture Notes in Artificial Intelligence, Vol. 1860, pages 154-165. Springer-Verlag, Berlin, 2000.

7 In [7], we present further comparison with the model of Vidal and Durfee, including results to show the relative advantage of our model with respect to both user satisfaction and computational cost.

AAMAS'04, July 19-23, 2004, New York, New York, USA.